

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2004-508812

(P2004-508812A)

(43) 公表日 平成16年3月25日(2004.3.25)

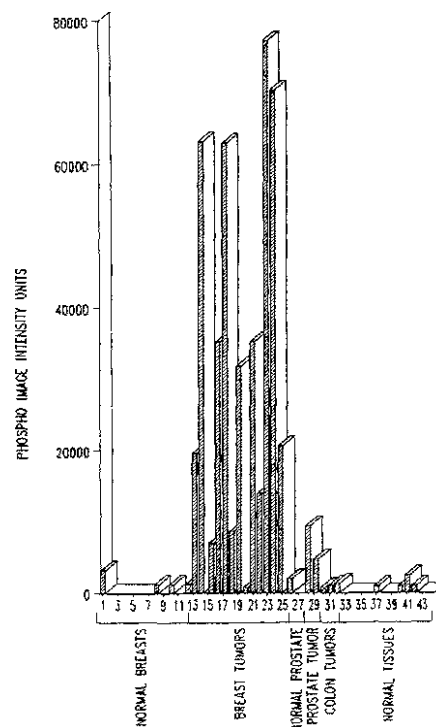
(51) Int.Cl. ⁷	F I	テーマコード (参考)
C 1 2 N 15/09	C 1 2 N 15/00	Z N A A 4 B O 2 4
A 6 1 K 31/7088	A 6 1 K 31/7088	4 B O 6 3
A 6 1 K 35/14	A 6 1 K 35/14	Z 4 B O 6 5
A 6 1 K 38/00	A 6 1 K 39/395	E 4 C O 8 4
A 6 1 K 39/395	A 6 1 K 39/395	T 4 C O 8 5
	審査請求 未請求 予備審査請求 有	(全 333 頁) 最終頁に続く

(21) 出願番号	特願2001-586963 (P2001-586963)	(71) 出願人	397069329
(86) (22) 出願日	平成13年5月22日 (2001.5.22)		コリクサ コーポレーション
(85) 翻訳文提出日	平成14年11月21日 (2002.11.21)		アメリカ合衆国 ワシントン 98104
(86) 国際出願番号	PCT/US2001/016776		, シアトル, コロンビア ストリート
(87) 国際公開番号	W02001/090152		1124, スイート 200
(87) 国際公開日	平成13年11月29日 (2001.11.29)	(74) 代理人	100078282
(31) 優先権主張番号	09/577, 505		弁理士 山本 秀策
(32) 優先日	平成12年5月24日 (2000.5.24)	(74) 代理人	100062409
(33) 優先権主張国	米国 (US)		弁理士 安村 高明
(31) 優先権主張番号	09/590, 583	(74) 代理人	100113413
(32) 優先日	平成12年6月8日 (2000.6.8)		弁理士 森下 夏樹
(33) 優先権主張国	米国 (US)	(72) 発明者	フルダキス, トニー エヌ.
(31) 優先権主張番号	09/699, 295		アメリカ合衆国 フロリダ 34243,
(32) 優先日	平成12年10月26日 (2000.10.26)		サラソタ, ブロードムーア バインズ
(33) 優先権主張国	米国 (US)		ブルバード 7937
			最終頁に続く

(54) 【発明の名称】 乳癌の治療および診断のための組成物および方法

(57) 【要約】

癌（特に、乳癌）の治療および診断のための組成物および方法が開示される。例示的な組成物は、1つ以上の乳癌ポリペプチド、その免疫原性部分、このようなポリペプチドをコードするポリヌクレオチド、このようなポリペプチドを発現する抗原提示細胞、ならびにこのようなポリペプチドを発現する細胞に特異的なT細胞を含む。開示された組成物は、例えば、疾患（特に、乳癌）の診断、予防および／または処置において有用である。特定の好ましい実施形態において、本発明のポリペプチドおよび／またはポリヌクレオチドは、免疫原性である。すなわち、これらは、免疫応答（特に、体液性および／または細胞性免疫応答）を誘発し得る。



【特許請求の範囲】

【請求項 1】

単離されたポリヌクレオチドであって、該ポリヌクレオチドは、以下：

(a) 配列番号 1、3～86、142～298、301～303、307、313、314、316、317、325 および 327～330 に提供される配列；

(b) 配列番号 1、3～86、142～298、301～303、307、313、314、316、317、325 および 327～330 に提供される配列の相補体；

(c) 配列番号 1、3～86、142～298、301～303、307、313、314、316、317、325 および 327～330 に提供される配列の少なくとも 20 個連続する残基からなる配列；

(d) 配列番号 1、3～86、142～298、301～303、307、313、314、316、317、325 および 327～330 に提供される配列に中程度にストリンジェントな条件下でハイブリダイズする配列；

(e) 配列番号 1、3～86、142～298、301～303、307、313、314、316、317、325 および 327～330 の配列に少なくとも 75 % の同一性を有する配列；

(f) 配列番号 1、3～86、142～298、301～303、307、313、314、316、317、325 および 327～330 の配列に少なくとも 90 % の同一性を有する配列；ならびに

(g) 配列番号 1、3～86、142～298、301～303、307、313、314、316、317、325 および 327～330 に提供される配列の縮重改変体、からなる群より選択される配列を含む、単離されたポリヌクレオチド。

10

20

【請求項 2】

単離されたポリペプチドであって、該ポリペプチドは、以下：

(a) 配列番号 299、300、304～306、308～312、314、326 および 331～337；

(b) 請求項 1 に記載のポリヌクレオチドによってコードされる配列；

(c) 請求項 1 に記載のポリヌクレオチドによってコードされる配列に少なくとも 70 % の同一性を有する配列；ならびに

(d) 請求項 1 に記載のポリヌクレオチドによってコードされる配列に少なくとも 90 % の同一性を有する配列、からなる群より選択されるアミノ酸配列を含む、単離されたポリペプチド。

30

【請求項 3】

発現制御配列に作動可能に連結された、請求項 1 に記載のポリヌクレオチドを含む、発現ベクター。

【請求項 4】

請求項 3 に記載の発現ベクターで形質転換またはトランスフェクトされた、宿主細胞。

【請求項 5】

請求項 2 に記載のポリペプチドに特異的に結合する、単離された抗体またはその抗原結合フラグメント。

40

【請求項 6】

患者における癌の存在を検出するための方法であって、該方法は、以下の工程：

(a) 該患者から生物学的サンプルを得る工程；

(b) 請求項 2 に記載のポリペプチドに結合する結合因子と該生物学的サンプルを接触させる工程；

(c) 該結合因子に結合するポリペプチドの量を該サンプル中で検出する工程；

(d) ポリペプチドの量を、所定のカットオフ値と比較し、そして、それから、該患者における癌の存在を決定する工程、を包含する、方法。

【請求項 7】

50

請求項 2 に記載の少なくとも 1 つのポリペプチドを含む、融合タンパク質。

【請求項 8】

配列番号 1、3～86、142～298、301～303、307、313、314、316、317、325 および 327～330 に示される配列に中程度にストリンジェントな条件下でハイブリダイズする、オリゴヌクレオチド。

【請求項 9】

腫瘍タンパク質に特異的な T 細胞を刺激および／または拡大するための方法であって、該方法は、以下：

- (a) 請求項 2 に記載のポリペプチド；
- (b) 請求項 1 に記載のポリヌクレオチド；および 10
- (c) 請求項 2 に記載のポリペプチドを発現する抗原提示細胞、
の群より選択される少なくとも 1 つの成分と、T 細胞を、T 細胞の刺激および／または拡大を可能にするのに十分な条件下および時間で接触させる工程、を包含する、方法。

【請求項 10】

請求項 9 に記載の方法に従って調製された T 細胞を含む、単離された T 細胞集団。

【請求項 11】

第 1 の成分および第 2 の成分を含む組成物であって、該第 1 の成分は、生理学的に受容可能なキャリアおよび免疫刺激剤からなる群より選択され、そして該第 2 の成分は、以下：

- (a) 請求項 2 に記載のポリペプチド；
- (b) 請求項 1 に記載のポリヌクレオチド； 20
- (c) 請求項 5 に記載の抗体；
- (d) 請求項 7 に記載の融合タンパク質；
- (e) 請求項 10 に記載の T 細胞集団；および
- (f) 請求項 2 に記載のポリペプチドを発現する抗原提示細胞、
からなる群より選択される、組成物。

【請求項 12】

請求項 11 に記載の組成物を患者に投与する工程を包含する、患者における免疫応答を刺激するための方法。

【請求項 13】

請求項 11 に記載の組成物を患者に投与する工程を包含する、患者における癌の処置のための方法。 30

【請求項 14】

患者における癌の存在を決定するための方法であって、該方法は、以下の工程：

- (a) 該患者から生物学的サンプルを得る工程；
- (b) 請求項 8 に記載のオリゴヌクレオチドと該生物学的サンプルを接触させる工程；
- (c) 該オリゴヌクレオチドにハイブリダイズするポリヌクレオチドの量を該サンプル中で検出する工程；
- (d) 該オリゴヌクレオチドにハイブリダイズするポリヌクレオチドの量を、所定のカットオフ値と比較し、そして、それから、該患者における癌の存在を決定する工程、
を包含する、方法。 40

【請求項 15】

請求項 8 に記載の少なくとも 1 つのオリゴヌクレオチドを備える、診断キット。

【請求項 16】

請求項 5 に記載の少なくとも 1 つの抗体および検出試薬を備える診断キットであって、該検出試薬がレポーター基を含む、診断キット。

【請求項 17】

患者における癌の発生を阻害するための方法であって、該方法は、以下の工程：

- (a) T 細胞が増殖するように、患者から単離された CD4+ および／または CD8+ T 細胞を、以下：
- (i) 請求項 2 に記載のポリペプチド； 50

(i i) 請求項 1 に記載のポリヌクレオチド；および
(i i i) 請求項 2 に記載のポリペプチドを発現する抗原提示細胞、
からなる群より選択される少なくとも 1 つの成分と共にインキュベートする工程；
(b) 有効量の該増殖した T 細胞を該患者に投与する工程、
を包含し、
それにより、該患者における癌の発生を阻害する、方法。

【発明の詳細な説明】

【 0 0 0 1 】

(発明の技術分野)

本発明は、一般に、癌（例えば、乳癌）の治療および診断に関する。本発明は、より詳細には、乳癌タンパク質を少なくとも一部含むポリペプチド、およびこのようなポリペプチドをコードするポリヌクレオチド分子に関する。このようなポリペプチドおよびポリヌクレオチドは、乳癌の診断および処置のための、薬学的組成物（例えば、ワクチン）および他の組成物において有用である。

【 0 0 0 2 】

(発明の背景)

乳癌は、米国および世界中で、女性についての重大な健康問題である。その疾患の検出および処置において、進歩が成されてきたが、乳癌は、女性の癌関係の第二番目の死因のままであり、毎年、米国で 1 8 0 , 0 0 0 人を超える女性が罹患している。北米の女性に関して、生存中の乳癌に罹る可能性は、現在 8 人に 1 人である。

【 0 0 0 3 】

現在のところ利用可能である、乳癌の予防または処置のためのワクチンまたは他の一般に優れた方法は存在しない。その疾患の対応は、現在のところ、早期の診断（慣用的な乳房のスクリーニング手順を通じて）および積極的な処置の組み合わせに依存し、その処置は、手術、放射線治療、化学療法、およびホルモン療法のような、1 以上の種々の処置を含み得る。特定の乳癌の処置の過程は、しばしば、特異的な腫瘍マーカーの分析を含む、種々の予後のパラメーターに基づいて選択される。例えば、Porter-Jordan および Lippman、Breast Cancer 8 : 7 3 - 1 0 0 (1 9 9 4) を参照のこと。しかし、確立されたマーカーの使用は、しばしば、解釈が難しい結果を導き、そして乳癌患者において観察される高い死亡率は、その疾患の処置、診断、および予防において、改善が必要とされることを示している。

【 0 0 0 4 】

従って、当該分野において、乳癌の治療および診断の改善された方法の必要性が存在する。本発明は、これらの必要性を満たし、さらに他の関係する利点を提供する。

【 0 0 0 5 】

(発明の要旨)

1 つの局面において、本発明はポリヌクレオチド組成物を提供し、このポリヌクレオチドは、以下：

(a) 配列番号 1、3 ~ 8 6、1 4 2 ~ 2 9 8、3 0 1 ~ 3 0 3、3 0 7、3 1 3、3 1 4、3 1 6、3 1 7、3 2 5 および 3 2 7 ~ 3 3 0 に提供される配列；

(b) 配列番号 1、3 ~ 8 6、1 4 2 ~ 2 9 8、3 0 1 ~ 3 0 3、3 0 7、3 1 3、3 1 4、3 1 6、3 1 7、3 2 5 および 3 2 7 ~ 3 3 0 に提供される配列の相補体；

(c) 配列番号 1、3 ~ 8 6、1 4 2 ~ 2 9 8、3 0 1 ~ 3 0 3、3 0 7、3 1 3、3 1 4、3 1 6、3 1 7、3 2 5 および 3 2 7 ~ 3 3 0 に提供される配列の少なくとも 2 0 個連続する残基からなる配列；

(d) 配列番号 1、3 ~ 8 6、1 4 2 ~ 2 9 8、3 0 1 ~ 3 0 3、3 0 7、3 1 3、3 1 4、3 1 6、3 1 7、3 2 5 および 3 2 7 ~ 3 3 0 に提供される配列に中程度にストリンジェントな条件下でハイブリダイズする配列；

(e) 配列番号 1、3 ~ 8 6、1 4 2 ~ 2 9 8、3 0 1 ~ 3 0 3、3 0 7、3 1 3、3 1 4、3 1 6、3 1 7、3 2 5 および 3 2 7 ~ 3 3 0 の配列に少なくとも 7 5 % の同一性を

有する配列；

(f) 配列番号 1、3～86、142～298、301～303、307、313、314、316、317、325 および 327～330 の配列に少なくとも 90% の同一性を有する配列；ならびに

(g) 配列番号 1、3～86、142～298、301～303、307、313、314、316、317、325 および 327～330 に提供される配列の縮重改変体、からなる群より選択される配列を含む。

【0006】

1 つの好ましい実施形態において、本発明のポリヌクレオチド組成物は、試験された乳癌サンプルの少なくとも約 20% において、より好ましくは少なくとも約 30% において、そして最も好ましくは少なくとも約 50% において、正常組織のレベルの少なくとも約 2 倍、好ましくは少なくとも約 5 倍、そして最も好ましくは少なくとも約 10 倍のレベルで、発現される。

10

【0007】

本発明は、別の局面において、上記のポリヌクレオチド配列によってコードされるアミノ酸配列を含むポリペプチド組成物を提供する。

【0008】

本発明はさらに、配列番号 299、300、304～306、308～312、314、326 および 331～334 に示される配列からなる群より選択されるアミノ酸配列を含むポリペプチド組成物を提供する。

20

【0009】

特定の好ましい実施形態において、本発明のポリペプチドおよび／またはポリヌクレオチドは、本明細書中にさらに記載されるように、免疫原性である。すなわち、これらは、免疫応答（特に、体液性および／または細胞性免疫応答）を誘発し得る。

【0010】

本発明はさらに、開示されたポリペプチドおよび／またはポリヌクレオチド配列のフラグメント、改変体および／または誘導体を提供し、ここで、フラグメント、改変体および／または誘導体は、好ましくは、配列番号 299、300、304～306、308～312、314、326 および 331～334 に示されるポリペプチド配列、または配列番号 1、3～86、142～298、301～303、307、313、314、316、317、325 および 327～330 に示されるポリヌクレオチド配列によってコードされるポリペプチド配列の少なくとも約 50% 免疫原性活性のレベル、好ましくは少なくとも約 70% そしてより好ましくは少なくとも約 90% の免疫原性活性のレベルを有する。

30

【0011】

本発明はさらに、上記ポリペプチドをコードするポリヌクレオチド、このようなポリヌクレオチドを含む発現ベクターおよびこのような発現ベクターで形質転換されたかまたはトランスフェクトされた宿主細胞を提供する。

【0012】

他の局面において、本発明は、上記のようなポリペプチドまたはポリヌクレオチドを含む薬学的組成物および生理学的に受容可能なキャリアを提供する。

40

【0013】

本発明の関連した局面において、薬学的組成物（例えば、ワクチン組成物）が、予防適用または治療適用のために提供される。このような組成物は、一般に、本発明の免疫原性ポリペプチドまたはポリヌクレオチドおよび免疫刺激剤（例えば、アジュバント）を含む。

【0014】

本発明はさらに、以下を含む薬学的組成物を提供する：(a) 本発明のポリペプチドに特異的に結合する抗体もしくは抗原結合フラグメント、またはそれらのフラグメント；および (b) 生理学的に受容可能なキャリア。

【0015】

さらなる局面において、本発明は、以下を含む薬学的組成物を提供する：(a) 上記のよ

50

うなポリペプチドを発現する抗原提示細胞および (b) 薬学的に受容可能なキャリアまたは賦形剤。例示的な抗原提示細胞としては、樹状細胞、マクロファージ、単球、線維芽細胞および B 細胞が挙げられる。

【0016】

関連した局面において、以下を含む薬学的組成物が提供される： (a) 上記のようなポリペプチドを発現する抗原提示細胞および (b) 免疫刺激剤。

【0017】

本発明はさらに、他の局面において、代表的に、生理学的に受容可能なキャリアおよび／または免疫刺激剤を含む薬学的組成物（例えば、ワクチン組成物）の形態で、上記のような少なくとも 1 つのポリペプチドを含む融合タンパク質、ならびにこのような融合タンパク質をコードするポリヌクレオチドを提供する。この融合タンパク質は、上記のように、複数の免疫原性ポリペプチドまたはその部分／改変体を含み得、そしてポリペプチドの発現、精製および／または免疫原性を容易にする 1 つ以上のポリペプチドセグメントをさらに含み得る。

10

【0018】

さらなる局面において、本発明は、患者における免疫応答（好ましくは、ヒト患者における T 細胞応答）を刺激するための方法を提供し、この方法は、本明細書中に記載された薬学的組成物を投与する工程を包含する。患者は、乳癌に冒され得るか（この場合、この方法が疾患の処置を提供する）、またはこのような疾患についての危険性が考えられる患者は、予防的に処置され得る。

20

【0019】

さらなる局面において、本発明は、患者における癌の発生を阻害するための方法を提供し、この方法は、上記のような薬学的組成物を患者に投与する工程を包含する。患者は、乳癌に冒され得るか（この場合、この方法が疾患の処置を提供する）、またはこのような疾患についての危険性が考えられる患者は、予防的に処置され得る。

【0020】

本発明は、他の局面において、生物学的サンプルから腫瘍細胞を除去する方法をさらに提供する。この方法は、生物学的サンプルと、本発明のポリペプチドと特異的に反応する T 細胞とを接触させる工程を包含し、ここで、この接触させる工程は、このサンプルからのこのタンパク質を発現する細胞の除去を可能にするための条件下および十分な時間で実施される。

30

【0021】

関連する局面において、患者における癌の発達を阻害するための方法が提供される。この方法は、上記のように処理された生物学的サンプルを患者に投与する工程を包含する。

【0022】

他の局面において、本発明のポリペプチドに特異的な T 細胞を刺激するか、そして／または、これらを増大する方法がさらに提供される。これらの方法は、T 細胞の刺激および／または増大を可能にするための条件下および十分な時間で、以下の 1 つ以上と T 細胞とを接触させる工程を包含する： (i) 上記のようなポリペプチド； (ii) このようなポリペプチドをコードするポリヌクレオチド；および／または (iii) このようなポリペプチドを発現する抗原提示細胞。上記のように調製された T 細胞を含む単離された T 細胞集団がまた、提供される。

40

【0023】

さらなる局面において、本発明は、患者における癌の発達を阻害する方法を提供する。この方法は、上記のように治療有効量の T 細胞集団を患者に投与する工程を包含する。

【0024】

本発明はさらに、患者における癌の発達を阻害する方法を提供する。この方法は、以下の工程を包含する： (a) 以下の 1 つ以上を有する患者から単離した CD4⁺ T 細胞および／または CD8⁺ T 細胞をインキュベートする工程： (i) 本明細書中で開示されたポリペプチドの少なくとも 1 つの免疫原性部位を含むポリペプチド； (ii) このようなポリ

50

ペプチドをコードするポリヌクレオチド；および（i i i）このようなポリペプチドを発現した抗原提示細胞；ならびに（b）有効量の増殖したT細胞を患者に投与し、それによって、患者における癌の発達を阻害する工程。増殖した細胞は、患者への投与前にクローン化され得るが、必ずしもクローン化するに及ばない。

【0025】

さらなる局面において、本発明は、患者において癌（好ましくは乳癌）の存在または非存在を決定するための方法を提供する。この方法は、以下を包含する：（a）患者から得た生物学的サンプルと、上に列挙したポリペプチドに結合する結合因子とを接触させる工程；（b）サンプル中で、この結合因子に結合するポリペプチドの量を検出する工程；および（c）このポリペプチドの量と事前に決定していたカットオフ値とを比較し、それによって患者における癌の存在または非存在を決定する工程。好ましい実施形態において、この結合因子は、抗体であり、より好ましくはモノクローナル抗体である。

10

【0026】

他の局面において、本発明はまた、患者における癌の進行をモニタリングする方法を提供する。このような方法は、以下の工程を包含する：（a）最初の時点で患者から得た生物学的サンプルと上に列挙したポリペプチドに結合する結合因子とを接触させる工程；（b）サンプルにおいて、この結合因子に結合するポリペプチドの量を検出する工程；（c）次の時点において、患者から得られた生物学的サンプルを使用し、工程（a）および工程（b）を反復する工程；ならびに（d）工程（c）において検出されたポリペプチドの量を工程（b）において検出された量と比較し、それらによって、患者における癌の進行をモニタリングする工程。

20

【0027】

他の局面において、本発明は、患者の癌の存在または非存在を決定するための方法をさらに提供する。この方法は以下の工程を包含する：（a）患者から得られた生物学的サンプルと、本発明のポリペプチドをコードするポリヌクレオチドにハイブリダイズするオリゴヌクレオチドとを接触させる工程；（b）そのサンプルにおいて、そのオリゴヌクレオチドとハイブリダイズするポリヌクレオチド、好ましくはmRNAのレベルを検出する工程；ならびに（c）そのオリゴヌクレオチドとハイブリダイズするポリヌクレオチドのレベルと事前に決定されたカットオフ値とを比較して、それによって患者の癌の存在または非存在を検出する工程。特定の実施形態において、mRNAの量は、例えば、上記に列挙されるようなポリペプチドをコードするポリヌクレオチドまたはそのようなポリヌクレオチドの相補体とハイブリダイズする、少なくとも1つのオリゴヌクレオチドプライマーを使用するポリメラーゼ連鎖反応によって検出される。他の実施形態において、mRNAの量は、上記に列挙されるようなポリペプチドをコードするポリヌクレオチドまたはそのようなポリヌクレオチドの相補体とハイブリダイズするオリゴヌクレオチドプローブを使用する、ハイブリダイゼーション技術を使用して検出される。

30

【0028】

関連する局面において、患者の癌の進行をモニタリングする方法が、提供され、その方法は以下の工程を包含する：（a）患者から得られた生物学的サンプルと、本発明のポリペプチドをコードするポリヌクレオチドにハイブリダイズするオリゴヌクレオチドとを接触させる工程；（b）そのサンプルにおいて、そのオリゴヌクレオチドとハイブリダイズするポリヌクレオチドの量を検出する工程；ならびに（c）次の時点において、患者から得られた生物学的サンプルを使用し、工程（a）および工程（b）を反復する工程；ならびに（d）（c）において検出されたポリヌクレオチドの量を工程（b）において検出された量と比較し、それらによって、患者における癌の進行をモニタリングする工程。

40

【0029】

さらなる局面において、本発明は、上記のようなポリペプチドに結合する抗体（例えば、モノクローナル抗体）ならびにこのような抗体を含む診断キットを提供する。上記のような1つ以上のオリゴヌクレオチドプローブまたはオリゴヌクレオチドプライマーを含む診断キットがまた、提供される。

50

【0030】

本発明のこれらの局面および他の局面は、以下の詳細な説明および添付の図面を参照することによって明らかになる。本明細書中で開示される全ての参考文献は、各々が個々に援用されるがごとく参考としてその全体が本明細書によって援用される。

【0031】

(発明の詳細な説明)

本発明は、一般に、組成物、ならびに癌（特に、乳癌）の治療および診断におけるその使用に関する。以下でさらに記載するように、本発明の例示的な組成物としては、以下を含むがこれらに限定されない：ポリペプチド、特に免疫原性ポリペプチド、このようなポリペプチドをコードするポリヌクレオチド、抗体および他の結合因子、抗原提示細胞（APC）ならびに免疫系細胞（例えば、T細胞）。 10

【0032】

本発明の実施は、特にそれと逆であることを示さなければ、ウイルス学、免疫学、細菌学、分子生物学、および当業者の範囲である組換えDNA技術の従来の方法（これらの多くは、例示目的で以下に記載される）を使用する。このような技術は、文献によって完全に説明されている。例えば、Sambrookら、Molecular Cloning: A Laboratory Manual（第2版）1989；Maniatisら、Molecular Cloning: A Laboratory Manual（1982）；DNA Cloning: A Practical Approach、第I巻および第II巻（D. Glover編）；Oligonucleotide Synthesis（N. Gait編、1984）；Nucleic Acid Hybridization（B. HamesおよびS. Higgins編、1985）；Transcription and Translation（B. HamesおよびS. Higgins編、1984）；Animal Cell Culture（R. Freshney編、1986）；Perbal、A Practical Guide to Molecular Cloning（1984）を参照のこと。 20

【0033】

本明細書中で引用される、全ての刊行物、特許および特許出願は、上記または下記にかかわらず、全体が参考して本明細書によって援用される。

【0034】

本明細書中および添付の特許請求の範囲で使用される場合、単数形「a」、「an」、および「the」は、文脈によって明らかに他の方法で示されていない限り、複数の参照を含む。 30

【0035】

(ポリペプチド組成物)

本明細書中で使用される場合、用語「ポリペプチド」は、その従来の意味、すなわちアミノ酸の配列として使用される。これらのポリペプチドは、特定の長さの生成物に限定されず；従って、ペプチド、オリゴペプチド、およびタンパク質は、このポリペプチドの定義に含まれ、そしてこのような用語は、別段示さない限り、本明細書中で交換可能に使用され得る。この用語はまた、このポリペプチドの発現後修飾（例えば、グリコシル化、アセチル化、リン酸化など）ならびに天然および非天然の両方の、当該分野で公知の他の改変を称さない。すなわち、この用語は、これらを排除する。ポリペプチドは、タンパク質全体でもあり、または部分配列でもあり得る。本発明の状況において特定の目的のポリペプチドは、エピトープ、すなわち、ポリペプチドの免疫原特性の実質的に原因となり、かつ免疫応答を誘起し得る抗原決定基を含むアミノ酸部分配列である。 40

【0036】

本発明の特に例示的なポリペプチドとしては、配列番号1、3～86、142～298、301～303、307、313、314、316、317、325、および327～330のいずれか1つに示されたポリヌクレオチド配列、あるいは中度のストリンジェントな条件下または高度にストリンジェントな条件下で、配列番号1、3～86、142～2 50

98、301～303、307、313、314、316、317、325、および327～330のいずれか1つに示されるポリペプチド配列に対してハイブリダイズする配列によってコードされるポリペプチドを含む。特定の他の例示的な本発明のポリヌクレオチドとしては、配列番号の299、300、304～306、308～312、314、326および331～334いずれか1つに示されるアミノ酸配列が挙げられる。

【0037】

本発明のポリペプチドは、これらの同定が少なくとも部分的に乳房腫瘍サンプルにおけるこれらの発現の増大したレベルに基づくことの現れとして、本明細書中で時々、乳房腫瘍タンパク質または乳房腫瘍ポリペプチドとして参照される。従って、「乳房腫瘍ポリペプチド」または「乳房腫瘍タンパク質」は、一般に本発明のポリペプチド配列、またはこの
10
ようなポリペプチドをコードするポリヌクレオチド配列を示し、これらは、例えば、本明細書中で提供された代表的なアッセイを使用して決定されるように、正常組織での発現レベルより、少なくとも2倍のレベル、および好ましくは5倍のレベルで、試験された乳房腫瘍サンプルの、好ましくは約20%を超える、より好ましくは約30%を超える、そして最も好ましくは約50%以上を超える、乳房腫瘍サンプルに実質的な部分において発現される。本発明の乳房腫瘍ポリペプチド配列は、その腫瘍細胞での増大した発現レベルに基づき、以下にさらに記載されるように、診断マーカーならびに治療標的の両方として特定の有用性を有する。

【0038】

特定の好ましい実施形態において、本発明のポリペプチドは免疫原性であり、すなわちこれらは、乳癌に有する患者からの抗血清および／またはT細胞を用いる免疫アッセイ（例えば、ELISAまたはT細胞刺激アッセイ）で検出可能に反応する。免疫原性活性についてのスクリーニングは、当業者に周知の技術を使用して実施され得る。例えば、この
20
ようなスクリーニングは、例えば、HarlowおよびLane、Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory, 1988に記載されるスクリーニングのような方法を使用して実施され得る。1つの例示的な例において、ポリペプチドは固体支持体に固定され得、そして患者の血清と接触され得、血清中の抗体の固定されたポリペプチドへの結合を可能にする。次いで、非結合血清は除去され、そして結合した抗体は、例えば、¹²⁵I 標識プロテインAに用いて検出され得る。
30

【0039】

当業者に認識されるように、本明細書中で開示されるポリペプチドの免疫原性部分もまた、本発明により包含される。本明細書中で用いられる場合、「免疫原性部分」は、本発明の免疫原性ポリペプチドを認識するB細胞および／またはT細胞表面抗原レセプターと免疫学的に反応性である（すなわち、特異的に結合する）本発明の免疫原性ポリペプチドのフラグメントである。免疫原性部分は、一般に、周知の技術（例えば、Paul、Fundamental Immunology, 第3版、243-247 (Raven Press, 1993) およびこの文献で引用される参考文献に要約されている）を用いて同定され得る。このような技術は、抗原特異的抗体、抗血清、および／またはT細胞株もしくはクローンと反応する能力について、ポリペプチドをスクリーニングすることを包含する。
40
本明細書中で用いられる場合、抗血清および抗体は、これらが、抗原に特異的に結合する（すなわち、これらが、ELISAまたは他の免疫アッセイにおいてこのタンパク質と反応し、そして無関係のタンパク質とは検出可能に反応しない）場合に、「抗原特異的」である。このような抗血清および抗体は、本明細書中に記載されるように、そして周知技術を用いて、調製され得る。

【0040】

1つの好ましい実施形態において、本発明のポリペプチドの免疫原性部分は、（例えば、ELISAおよび／またはT細胞反応性アッセイにおいて）全長ポリペプチドの反応性よりも実質的に低くないレベルで、抗血清および／またはT細胞と反応する部分である。好ましくは、免疫原性部分の免疫原性活性のレベルは、全長ポリペプチドの免疫原性の少な
50

くとも約50%、好ましくは少なくとも約70%、そしてもっとも好ましくは90%以上である。いくつかの例において、対応する全長ポリペプチドよりも大きい免疫原性活性のレベルを有する（例えば、約100%または150%もしくはそれより大きい免疫原性活性を有する）好ましい免疫原性部分が同定される。

【0041】

特定の他の実施形態において、例示的な免疫原性部分は、N末端リーダー配列および／または膜貫通ドメインが欠失されたペプチドを含み得る。他の例示的な免疫原性部分は、成熟タンパク質に対して小さなN末端および／またはC末端の欠失（例えば、1～30アミノ酸、好ましくは5～15アミノ酸）を含む。

【0042】

別の実施形態において、本発明のポリペプチド組成物また、T細胞および／または本発明のポリペプチド（特に、本明細書中に開示されたアミノ酸配列を有するポリペプチドまたはそれらのフラグメントもしくは改変体）に対して生成された抗体と免疫学的に反応性である1つ以上のポリペプチドを含み得る。

【0043】

本発明の別の実施形態において、本明細書中に記載される1つ以上のポリペプチドによりコードされる1つ以上のポリペプチドと免疫学的に反応性であるT細胞および／または抗体を惹起し得る、1つ以上のポリペプチドを含むポリペプチド、または本明細書中に開示されるポリヌクレオチド配列に含まれる連続的な核酸配列、またはそれらの免疫原性フラグメントもしくは改変体、あるいは中程度から高度のストリンジェンシーの条件下で1つ以上のこれらの配列にハイブリダイズする1つ以上に核酸配列が提供される。

【0044】

別の局面において、本発明は、本明細書中に示されるポリペプチド組成物（例えば、配列番号299、300、304～306、308～312、314、326、および331～334に示されるポリペプチド、または配列番号1、3～86、142～298、301～303、307、313、314、316、317、325、および327～330の配列に示されるポリヌクレオチド配列によりコードされるポリペプチド）の少なくとも約5、10、15、20、25、50、または100連続するアミノ酸またはそれ以上（全ての中間の長さを含む）を含むポリペプチドフラグメントを提供する。

【0045】

別の局面において、本発明は、本明細書中に記載されるポリペプチド組成物の改変体を提供する。一般的に本発明により包含されるポリペプチド改変体は、代表的に、本明細書中に示されるポリペプチド配列に対して、その長さに沿って、少なくとも約70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、もしくは99%またはそれ以上の同一性（以下に記載するように決定された）を示す。

【0046】

1つの好ましい実施形態において、本発明により提供されるポリペプチドフラグメントおよび改変体は、本明細書中に詳細に示される全長ポリペプチドと反応する抗体および／またはT細胞と免疫学的に反応性である。

【0047】

別の好ましい実施形態において、本発明により提供されるポリペプチドフラグメントおよび改変体は、本明細書中に詳細に示される全長ポリペプチド配列により示される免疫原性活性のレベルの少なくとも約50%、好ましくは少なくとも約70%、そして最も好ましくは少なくとも約90%またはそれ以上の免疫原性活性のレベルを示す。

【0048】

ポリペプチド「改変体」は、本明細書中でこの用語が使用される場合、代表的に、本明細書中で詳細に開示されるポリペプチドと1つ以上の置換、欠失、付加、および／または挿入において異なるポリペプチドである。このような改変体は、天然に生じるか、または例えば、本発明の1つ以上の上記のポリペプチド配列を改変し、そして本明細書中に記載さ

10

20

30

40

50

れるようなそれらの免疫学的活性を評価することにより、ならびに／または当該分野で周知の任意の多くの技術を用いて、合成的に作製され得る。

【0049】

例えば、本発明のポリペプチドの特定の例示的な改変体として、1つ以上の部分（例えば、N末端リーダー配列または膜貫通ドメイン）が除去された改変体が挙げられる。他の例示的な改変体として、小さな部分（例えば、1～30アミノ酸、好ましくは5～15アミノ酸）が成熟タンパク質のN末端および／またはC末端から除去された改変体が挙げられる。

【0050】

多くの例において、改変体は、保存的置換を含む。「保存的置換」は、アミノ酸が類似の特性を有する別のアミノ酸で置換され、その結果、ペプチド化学の当業者がポリペプチドの二次構造および疎水的性質が実質的に変化していないことを期待するような置換である。上記のように、改変は、本発明のポリヌクレオチドおよびポリペプチドの構造においてなされ得、そしてなお所望の特性（例えば、免疫原性特性）を有する改変体または誘導体ポリペプチドをコードする機能的分子を獲得し得る。本発明のポリペプチドの等価物を（または本発明のポリペプチドの改善された免疫原性の改変体もしくは部分さえも）作製するようポリペプチドのアミノ酸配列が変更されることが所望される場合、当業者は、代表的に、表1に基づくコードDNA配列の1つ以上のコドンを変更する。

【0051】

例えば、特定のアミノ酸は、例えば、抗体の抗原結合領域または基質分子上の結合部位のような構造を有するタンパク質構造中の他のアミノ酸に、感知可能な相互作用的結合能の損失なしで、置換され得る。タンパク質の相互作用的な能力および性質がタンパク質の生物学的機能的活性を規定するので、特定のアミノ酸配列置換は、タンパク質配列、および当然ながら、その根底にあるDNAコード配列においてなされ得、そしてそれにも関わらず、同様の特性を有するタンパク質が入手される。従って、種々の変化が、ペプチドの生物学的有用性または活性の感知可能な損失を伴わずに、開示された組成物のペプチド配列またはそのペプチドをコードする対応するDNA配列においてなされ得ることが意図される。

【0052】

【表1】

表 I

アミノ酸			コドン						
アラニン	Ala	A	GCA	GCC	GCG	GCU			
システイン	Cys	C	UGC	UGU					
アスパラギン酸	Asp	D	GAC	GAU					
グルタミン酸	Glu	E	GAA	GAG					
フェニルアラニン	Phe	F	UUC	UUU					
グリシン	Gly	G	GGA	GGC	GGG	GGU			
ヒスチジン	His	H	CAC	CAU					
イソロイシン	Ile	I	AUA	AUC	AUU				
リジン	Lys	K	AAA	AAG					
ロイシン	Leu	L	UUA	UUG	CUA	CUC	CUG	CUU	
メチオニン	Met	M	AUG						
アスパラギン	Asn	N	AAC	AAU					
プロリン	Pro	P	CCA	CCC	CCG	CCU			
グルタミン	Gln	Q	CAA	CAG					
アルギニン	Arg	R	AGA	AGG	CGA	CGC	CGG	CGU	
セリン	Ser	S	AGC	AGU	UCA	UCC	UCG	UCU	
スレオニン	Thr	T	ACA	ACC	ACG	ACU			
バリン	Val	V	GUA	GUC	GUG	GUU			
トリプトファン	Trp	W	UGG						
チロシン	Tyr	Y	UAC	UAU					

10

20

30

40

50

このような変化を作製する際に、アミノ酸の疎水性親水性指標が考慮され得る。タンパク質に相互作用的な生物学的機能を付与する際の疎水性親水性アミノ酸指標の重要性は、一般的に当該分野において理解されている（K y t e および D o o l i t t l e、1982、本明細書中に参考として援用される）。アミノ酸の相対的な疎水性親水性の性質は、得られるタンパク質の二次構造に寄与し、これが次に、他の分子（例えば、酵素、基質、レセプター、DNA、抗体、抗原など）とのそのタンパク質の相互作用を規定することが受け入れられている。各アミノ酸は、その疎水性および電荷特性に基づいて疎水性親水性指標を割り当てられている（K y t e および D o o l i t t l e、1982）。これらの値は以下である：イソロイシン（+4.5）；バリン（+4.2）；ロイシン（+3.8）；フェニルアラニン（+2.8）；システイン／シスチン（+2.5）；メチオニン（+1.9）；アラニン（+1.8）；グリシン（-0.4）；スレオニン（-0.7）；セリン（-0.8）；トリプトファン（-0.9）；チロシン（-1.3）；プロリン（-1.6）；ヒスチジン（-3.2）；グルタミン酸（-3.5）；グルタミン（-3.5）；アスパラギン酸（-3.5）；アスパラギン（-3.5）；リジン（-3.9）；およびアルギニン（-4.5）。

【0053】

特定のアミノ酸が類似の疎水性親水性指標またはスコアを有する他のアミノ酸によって置換され得、そしてなお類似の生物学的活性を有するタンパク質を生じる（すなわち、生物学的に機能的に等価なタンパク質をなお入手する）ことが、当該分野において公知である。このような変化を作製する際に、その疎水性親水性指標が±2以内であるアミノ酸の置換が好ましく、その疎水性親水性指標が±1以内であるアミノ酸の置換が特に好ましく、そしてその疎水性親水性指標が±0.5以内であるアミノ酸の置換がなおより特に好ましい。同様なアミノ酸の置換が、親水性に基づいて有効になされ得ることもまた、当該分野で理解されている。米国特許第4,554,101号（その全体が本明細書中に参考として詳細に援用される）は、そのタンパク質の最大局所的平均親水性が、その隣接するアミノ酸の親水性によって支配されるので、そのタンパク質の生物学的特性と相関することを言及している。

【0054】

米国特許第 4, 554, 101 号に詳述されるように、以下の親水性値がアミノ酸残基に割り当てられた：アルギニン（+3.0）；リジン（+3.0）；アスパラギン酸（+3.0±1）；グルタミン酸（+3.0±1）；セリン（+0.3）；アスパラギン（+0.2）；グルタミン（+0.2）；グリシン（0）；スレオニン（-0.4）；プロリン（-0.5±1）；アラニン（-0.5）；ヒスチジン（-0.5）；システイン（-1.0）；メチオニン（-1.3）；バリン（-1.5）；ロイシン（-1.8）；イソロイシン（-1.8）；チロシン（-2.3）；フェニルアラニン（-2.5）；トリプトファン（-3.4）。アミノ酸は、類似の親水性の値を有する別のアミノ酸に置換され得、そしてなお、生物学的な等価物（特に、免疫学的に等価なタンパク質）を得ることが理解される。このような変化において、その親水性値が±2 以内であるアミノ酸の置換が好ましく、その親水性値が±1 以内であるアミノ酸の置換が特に好ましく、そしてその親水性値が±0.5 以内であるアミノ酸の置換がなおより特に好ましい。

10

【0055】

上記で概説したように、従って、アミノ酸置換は、一般的にアミノ酸側鎖置換基の相対的な類似性（例えば、その疎水性、親水性、電荷、大きさ、など）に基づく。前述の種々の特徴を考慮する典型的な置換は当業者に周知であり、そして以下を含む：アルギニンおよびリジン；グルタミン酸およびアスパラギン酸；セリンおよびスレオニン；グルタミンおよびアスパラギン；ならびにバリン、ロイシン、およびイソロイシン。

【0056】

さらに、任意のポリヌクレオチドが、インビボでの安定性を増加させるためにさらに改変され得る。可能性のある改変として以下が挙げられるがこれらに限定されない：5' 末端および／または 3' 末端での隣接配列の付加；骨格におけるホスホジエステラーゼ結合ではなくホスホロチオエートまたは 2' O-メチルの使用；ならびに／あるいは従来とは異なる塩基（例えば、イノシン、キューオシン、およびワイブトシン、ならびにアデニン、シチジン、グアニン、チミン、およびウリジンのアセチル、メチル、チオ、および他の修飾形態）の含有。

20

【0057】

アミノ酸置換はさらに、残基の極性、電荷、溶解性、疎水性、親水性および／または両親媒性特性の類似性に基づいて作製され得る。例えば、負に荷電したアミノ酸としては、アスパラギン酸およびグルタミン酸；正に荷電したアミノ酸としては、リジンおよびアルギニン；そして類似の親水性値を有する非荷電性の極性頭部を有するアミノ酸としては、ロイシン、イソロイシンおよびバリン；グリシンおよびアラニン；アスパラギンおよびグルタミン；ならびにセリン、スレオニン、フェニルアラニンおよびチロシンが挙げられる。保存的变化を示し得るアミノ酸の他のグループとしては、(1) ala、pro、gly、glu、asp、gln、asn、ser、thr；(2) cys、ser、tyr、thr；(3) val、ile、leu、met、ala、phe；(4) lys、arg、his；および(5) phe、tyr、trp、his が挙げられる。改変体はまた、またはあるいは、非保存的变化を含み得る。好ましい実施形態において、改変体ポリペプチドは、5 アミノ酸またはそれより少ないアミノ酸の置換、欠失または付加によって、ネイティブの配列とは異なる。改変体はまた（またはあるいは）、例えば、ポリペプチドの免疫原性、二次構造および疎水性親水性特性に対する最小の影響を有するアミノ酸の欠失または付加によって、改変され得る。

30

40

【0058】

上記のように、ポリペプチドは、タンパク質の N 末端にシグナル（または、リーダー）配列を含み得、これは、翻訳と同時に、または翻訳後に、そのタンパク質の転移を指向する。このポリペプチドはまた、このポリペプチドの合成、精製または同定を容易にするために、またはこのポリペプチドの固体支持体への結合を増強するために、リンカー配列または他の配列（例えば、ポリ His）に結合体化され得る。例えば、ポリペプチドは、免疫グロブリン Fc 領域に結合体化され得る。

【0059】

50

ポリペプチド配列が比較される場合、以下に記載されるように最大の一致で整列するとき
に2つの配列中のアミノ酸の配列が同じである場合、2つの配列は、「同一」であるとい
われる。2つの配列間の比較は、代表的には、比較ウィンドウによって配列を比較して、
配列類似性の局所的領域を同定および比較することによって行われる。本明細書中で使用
される「比較ウィンドウ」とは、少なくとも約20、通常は30～約75、40～約50
の連続する位置のセグメントをいう。ここで、配列は、2つの配列が必要に応じて整列さ
れた後、同じ数の連続する位置の参照配列と比較され得る。

【0060】

比較のための配列の最適な整列は、生命情報科学ソフトウェアのLasergene s
uiteにおけるMegalignプログラム(DNA STAR, Inc., Madis 10
on, WI)を使用して、デフォルトパラメータを用いて行われ得る。このプログラムは
、以下の参考文献に記載のいくつかの整列スキームを統合する: Dayhoff, M. O
.(編) Atlas of Protein Sequence and Struct
ure, National Biomedical Research Foundat
ion, Washington DC 第5巻, 補遺3, 345～358頁におけるDa
yhoff, M. O. (1978) A model of evolutionar
y change in proteins—Matrices for detectin
g distant relationships.; Hein J. (1990) Un
ified Approach to Alignment and Phylogen
es 626～645頁 Methods in Enzymology 20
第183巻, Academic Press, Inc., San Diego, CA; Higgins
 , D. G. および Sharp, P. M. (1989) CABIOS 5:151—153
 ; Myers, E. W. および Muller W. (1988) CABIOS 4:11
 —17; Robinson, E. D. (1971) Comb. Theor 11:105
 ; Santou, N. Nes, M. (1987) Mol. Biol. Evol. 4:40
 6—425; Sneath, P. H. A. および Sokal, R. R. (1973) Nu
merical Taxonomy—the Principles and Prac
tice of Numerical Taxonomy, Freeman Press
 , San Francisco, CA; Wilbur, W. J. および Lipman, D
 . J. (1983) Proc. Natl. Acad., Sci. USA 80:726— 30
730。

【0061】

あるいは、比較のための配列の最適な整列は、SmithおよびWaterman (19
81) Add. APL. Math 2:482の部分的同一性アルゴリズムによって、N
eedlemanおよびWunsch (1970) J. Mol. Biol. 48:443
の同一性整列アルゴリズムによって、PearsonおよびLipman (1988) P
roc. Natl. Acad. Sci. USA 85:2444の類似性検索方法によっ
て、これらのアルゴリズムのコンピュータ化した実行(Wisconsin Genet
ics Software Package, Genetics Computer G
roup (GCG), 575 Science Dr., Madison, WIにおける 40
GAP、BESTFIT、BLAST、FASTA、およびTFASTA)によって、ま
たは検査によって行われ得る。

【0062】

配列同一性および配列類似性の割合を決定するために適切なアルゴリズムの1つの好まし
い例は、BLASTおよびBLAST 2.0アルゴリズムであり、これらはそれぞれAl
tschulら (1977) Nucl. Acids Res. 25:3389—3402
およびAltschulら (1990) J. Mol. Biol. 215:403—410
に記載される。BLASTおよびBLAST 2.0は、本発明のポリヌクレオチドおよび
ポリペプチドについての配列同一性の割合を決定するために、例えば、本明細書中に記載
のパラメータを使用して使用され得る。BLAST分析を行うためのソフトウェアは、N 50

ational Center for Biotechnology Informationを通して公に利用可能である。アミノ酸配列について、スコアリングマトリクスは、累積スコアを算出するために使用され得る。各指示におけるワードヒットの拡大は、以下の場合に停止する：累積整列スコアが、その最大到達値から量Xだけ低下した場合；累積スコアが、1以上の負のスコアリングの残基整列の累積に起因してゼロ以下になった場合；またはいずれかの配列の末端に到達した場合。このBLASTアルゴリズムパラメーターW、TおよびXは、整列の感度および速度を決定する。

【0063】

好ましい1つのアプローチにおいて、「配列同一性のパーセンテージ」は、少なくとも20位置の比較ウィンドウにわたる2つの最適に整列した配列を比較することによって決定され、ここで比較ウィンドウ中のポリペプチド配列の部分は、参照配列（これは、付加または欠失を含まない）と比較して、2つの配列の最適な整列について20パーセント以下、通常は5～15パーセント、または10～12パーセントの付加または欠失（すなわち、ギャップ）を含み得る。このパーセンテージは、両方の配列で同一のアミノ酸残基が生じる位置の数を決定して一致する位置の数を得、この一致する位置の数を参照配列における位置の総数（すなわち、ウィンドウサイズ）で除算し、そしてこの結果に100を掛けて配列同一性のパーセンテージを得ることによって算出される。

【0064】

他の例示的实施形態において、ポリペプチドは、本明細書中に記載の複数のポリペプチドを含むか、または本明細書に記載の少なくとも1つのポリペプチドおよび関連しない配列（例えば、公知の腫瘍タンパク質）を含む融合ポリペプチドであり得る。例えば、融合パートナーは、Tヘルパーエピトープ、好ましくはヒトによって認識されるTヘルパーエピトープを提供する際に補助し得るか（免疫学的融合パートナー）、またはネイティブの組換えタンパク質より高い収量でタンパク質を発現する際に補助し得る（発現エンハンサー）。特定の好ましい融合パートナーは、免疫学的融合パートナーおよび発現増強融合パートナーの両方である。他の融合パートナーは、ポリペプチドの溶解性を増加するように、またはポリペプチドが所望の細胞内コンパートメントに標的化されることを可能にするように選択され得る。なおさらなる融合パートナーとしては、親和性タグ（これは、タンパク質の精製を容易にする）が挙げられる。

【0065】

融合ポリペプチドは、一般に、標準的な技術（化学的結合体化を含む）を使用して調製され得る。好ましくは、融合ポリペプチドは、発現系において、組換えポリペプチドとして発現され、非融合ポリペプチドと比較して、増加したレベルの産生を可能にする。手短に言うと、このポリペプチド成分をコードするDNA配列を、別々に組立て得、そして適切な発現ベクターに連結し得る。1つのポリペプチド成分をコードするDNA配列の3'末端は、ペプチドリinkerを用いてまたは用いずに、第2のポリペプチド成分をコードするDNA配列の5'末端に、これらの配列のリーディングフレームが同位相にあるように連結される。このことが、両方の成分ポリペプチドの生物学的活性を保持する単一の融合ポリペプチドへの翻訳を可能にする。

【0066】

ペプチドリinker配列は、各ポリペプチドがその二次構造および三次構造へと折り畳まれるのを保証するために十分な距離で第一および第二のポリペプチド構成要素を隔てるために用いられ得る。このようなペプチドリinker配列は、当該分野で周知の標準的な技術を用いて融合タンパク質中に組み込まれる。適切なペプチドリinker配列は、以下の因子に基づいて選択され得る：（1）伸長した可変コンホメーションを取る能力；（2）第一および第二のポリペプチド上の機能的なエピトープと相互作用し得る二次構造を取ることができないこと；ならびに（3）ポリペプチドの機能的なエピトープと反応し得る疎水性または荷電した残基の無いこと。好ましいペプチドリinker配列は、Gly、AsnおよびSer残基を含む。ThrおよびAlaのような中性に近い他のアミノ酸もまた、リンカー配列に用いられ得る。リンカーとして有用に用いられ得るアミノ酸配列としては、Ma

rateaら、Gene 40:39-46、1985; Murphyら、Proc. Natl. Acad. Sci. USA 83:8258-8262、1986; 米国特許第4,935,233号および米国特許第4,751,180号に開示されるアミノ酸配列が挙げられる。リンカー配列は、一般的に1から約50アミノ酸長であり得る。リンカー配列は、第一および第二のポリペプチドが、機能的ドメインを分離するため、および立体的な干渉を防ぐために用いられ得る非必須N末端アミノ酸領域を有する場合、必要とされない。

【0067】

連結されたDNA配列は、適切な転写または翻訳調節エレメントに作動可能に連結される。DNAの発現を担う調節エレメントは、第一のポリペプチドをコードするDNA配列の5'側にのみ位置する。同様に、翻訳および転写終結シグナルを終了するために必要とされる終止コドンは、第二のポリペプチドをコードするDNA配列の3'側にのみ存在する。

10

【0068】

融合ポリペプチドは、関連しない免疫原性ポリペプチドと共に、本明細書中に記載されるようなポリペプチド（例えば、リコール（recall）応答を惹起し得る免疫原性ポリペプチド）を含み得る。このようなタンパク質の例としては、破傷風タンパク質、結核タンパク質および肝炎タンパク質が挙げられる（例えば、Stoutera、New Engl. J. Med.、336:86-91（1997）を参照のこと）。

【0069】

1つの好ましい実施形態において、免疫学的融合パートナーは、Mycobacterium tuberculosis由来のRa12フラグメントのようにMycobacterium sp. 由来である。Ra12組成物ならびに異種ポリヌクレオチド／ポリペプチド配列の発現および／または免疫原性を増強する際に、Ra12を使用するための方法は、米国特許出願60/158,585（この開示は、本明細書中でその全体が参考として援用される）に記載されている。簡単には、Ra12は、Mycobacterium tuberculosis MTB32A核酸のサブ配列であるポリヌクレオチド領域をいう。MTB32Aは、M. tuberculosisの毒性菌株および無毒性菌株の遺伝子によってコードされる分子量32KDのセリンプロテアーゼである。MTB32Aのヌクレオチド配列およびアミノ酸配列が、記載されている（例えば、米国特許出願60/158,585; また、Skeikyら、Infection and Immun.（1999）67:3998-4007も参照のこと。これらは、本明細書中で参考として援用される）。MTB32Aコード配列のC末端フラグメントは高いレベルで発現し、精製プロセス全体にわたって可溶性ポリペプチドとして残存する。さらに、Ra12は、これが融合される、異種免疫原性ポリペプチドの免疫原性を増強し得る。1つの好ましいRa12融合ポリペプチドは、MTB32Aのアミノ酸残基192-323に対応する14KDのC末端フラグメントを含む。他の好ましいRa12ポリヌクレオチドは、一般的に、Ra12ポリペプチドの一部をコードする少なくとも約15個の連続するヌクレオチド、少なくとも約30個のヌクレオチド、少なくとも60個のヌクレオチド、少なくとも約100個のヌクレオチド、少なくとも約200個のヌクレオチド、または少なくとも約300個のヌクレオチドを含む。Ra12ポリヌクレオチドは、ネイティブな配列（すなわち、Ra12ポリペプチドまたはその一部をコードする内因性配列）を含んでもよいし、このような配列の改変体を含んでもよい。Ra12ポリヌクレオチド改変体は、1以上の置換、付加、欠失および／または挿入を含み得、その結果、コードされた融合ポリペプチドの生物学的活性が、ネイティブなRa12ポリペプチドを含む融合ポリペプチドと比較して、実質的に減少していない。好ましくは、改変体は、ネイティブなRa12ポリペプチドまたはその一部をコードするポリヌクレオチド配列に、少なくとも約70%の同一性、より好ましくは少なくとも約80%の同一性、そして最も好ましくは少なくとも約90%の同一性を示す。

20

30

40

【0070】

50

好ましい他の実施形態において、免疫学的融合パートナーは、グラム陰性の細菌 *Haemophilus influenza B* の表面タンパク質である、プロテイン D (WO 91/18926) に由来する。好ましくは、プロテイン D 誘導体は、ほぼ 3 分の 1 のタンパク質 (例えば、最初の N 末端 100~110 アミノ酸) を含み、そしてプロテイン D 誘導体は、脂質化 (lipidated) され得る。特定の好ましい実施形態において、リボプロテイン D 融合パートナーの最初の 109 残基は、さらなる外因性 T 細胞エピートープを有するポリペプチドを提供するように、そして *E. coli* 中の発現レベルを増加するように、N 末端に含まれる (従って、発現エンハンサーとして機能する)。脂質テールは、抗原提示細胞への抗原の最適な提示を保証する。他の融合パートナーは、インフルエンザウイルス由来の非構造的タンパク質、NS1 (血球凝集素) を含む。代表的に、N 末端の 81 アミノ酸が用いられるが、T ヘルパーエピートープを含む異なるフラグメントが使用され得る。

10

【0071】

別の実施形態において、免疫学的融合パートナーは、LYTA として公知のタンパク質、またはその部分 (好ましくは C 末端部分) である。LYTA は、アミダーゼ LYTA (*Lyta* 遺伝子によりコードされる; *Gene* 43:265~292, 1986) として公知の N-アセチル-L-アラニンアミダーゼを合成する *Streptococcus pneumoniae* 由来である。LYTA は、ペプチドグリカン骨格中の特定の結合を特異的に分解する自己溶解素である。LYTA タンパク質の C 末端ドメインは、コリンまたはいくつかのコリンアナログ (例えば、DEAE) に対する親和性を担う。この性質は、融合タンパク質の発現のために有用な *E. coli* C-LYTA 発現プラスミドの開発のために利用されてきた。アミノ末端で C-LYTA フラグメントを含むハイブリッドタンパク質の精製が、記載されている (*Biotechnology* 10:795~798, 1992 を参照のこと)。好ましい実施形態において、LYTA の反復部分は、融合タンパク質に組み込まれ得る。反復部分は、残基 178 で開始する C 末端領域中に見出される。特に好ましい反復部分は、残基 188~305 を組み込む。

20

【0072】

なお別の例示的な実施形態は、融合ポリヌクレオチド、およびそれらをコードするポリヌクレオチドを含み、ここでこの融合パートナーは、米国特許第 5,633,234 号に記載のようにポリペプチドをエンドソーム/リソソームコンパートメントへ指向させ得る標的シグナルを含む。本発明の免疫原性ポリペプチドは、この標的シグナルと融合する場合、より効率的に MHC クラス II 分子と結合し、これによってこのポリペプチドに特異的な CD4⁺ T 細胞のインビボでの増強された刺激が提供される。

30

【0073】

本発明のポリペプチドは、周知の種々の合成技術/組換え技術のいずれかを使用して調製される。これらの技術の後者を以下でさらに記載する。約 150 アミノ酸よりも少ないポリペプチド、部分および他の改変体は、一般的に当業者に周知の技術を使用して、合成手段により作製され得る。例示的な 1 つの実施例において、このようなポリペプチドは、商業的に入手可能な固相技術のいずれか (例えば、*Merrifield* 固相合成方法) を使用して合成され、ここでアミノ酸は、成長しているアミノ酸鎖に連続的に付加される。*Merrifield*, J. Am. Chem. Soc. 85:2149-2146, 1963 を参照のこと。ポリペプチドの自動化された合成のための機器が、*Perkin Elmer/Applied BioSystems Division* (*Foster City, CA*) などの供給業者から商業的に入手可能であり、そして製造者の指示に従って操作し得る。

40

【0074】

一般に、本発明のポリペプチド組成物 (融合ポリペプチドを含む) が単離される。「単離された」ポリペプチドは、その元来の環境から取り出されたものである。例えば、天然に存在するタンパク質またはポリペプチドは、それが天然の系中で共存する物質のいくつかまたは全てから分離されている場合、単離されている。好ましくは、このようなポリペ

50

チドは、精製され、例えば、少なくとも約 90% 純粋、より好ましくは少なくとも約 95% 純粋、そして最も好ましくは少なくとも約 99% 純粋である。

【0075】

(ポリヌクレオチド組成物)

他の局面において、本発明は、ポリヌクレオチド組成物を提供する。用語「DNA」および「ポリヌクレオチド」は、本明細書中で実質的に交換可能に使用され、特定の種の全ゲノム DNA を含まない単離された DNA 分子をいう。本明細書中で使用される場合、「単離された」は、ポリヌクレオチドが、他のコード配列から実質的に分離されており、そして DNA 分子が、無関係のコード DNA の大部分（例えば、大きな染色体フラグメントまたは他の機能的遺伝子またはポリペプチドコード領域）を含まないことを意味する。当然のことながら、これは、もともと単離された DNA 分子をいい、後で人工的にセグメントを付加された遺伝子またはコード領域を除外しない。

10

【0076】

当業者に理解されるように、本発明のポリヌクレオチド組成物は、タンパク質、ポリペプチド、ペプチドなどを発現するか、または発現し得るように適応されたゲノム配列、ゲノム外配列およびプラスミドにコードされる配列、ならびにより小さな操作された遺伝子セグメントを含み得る。このようなセグメントは、自然に単離され得るか、または人の手によって合成的に改変され得る。

【0077】

当業者に認識されるように、本発明のポリヌクレオチドは、一本鎖（コードもしくはアンチセンス）または二本鎖であり得、そして DNA 分子（ゲノム、cDNA もしくは合成）または RNA 分子であり得る。RNA 分子は、HnRNA 分子（これはイントロンを含み、そして DNA 分子に 1 対 1 の様式で対応する）、および mRNA 分子（これは、イントロンを含まない）を含み得る。さらなるコード配列または非コード配列は、本発明のポリヌクレオチド内に存在し得るがその必要はなく、そしてポリヌクレオチドは、他の分子および／または支持材料に連結され得るがその必要はない。

20

【0078】

ポリヌクレオチドは、ネイティブの配列（すなわち、本発明のポリペプチド／タンパク質、またはその部分をコードする内因性配列）を含み得るか、あるいはこのような配列の改変体または誘導体、そして好ましくは免疫原性改変体または誘導体をコードする配列を含み得る。

30

【0079】

従って、本発明の別の局面に従い、配列番号 1、3～86、142～298、301～303、307、313、314、316、317、325 および 327～330 のいずれか 1 つに示されるポリヌクレオチド配列、配列番号 1、3～86、142～298、301～303、307、313、314、316、317、325 および 327～330 のいずれか 1 つに示されるポリヌクレオチド配列の相補体、ならびに配列番号 1、3～86、142～298、301～303、307、313、314、316、317、325 および 327～330 のいずれか 1 つに示されるポリヌクレオチド配列の縮重改変体を含む本発明のポリヌクレオチド組成物が提供される。特定の好ましい実施形態において、本明細書中に示されるポリヌクレオチド配列は、上記のような免疫原性ポリペプチドをコードする。

40

【0080】

他の関連する実施形態において、本発明は、本明細書中で配列番号 1、3～86、142～298、301～303、307、313、314、316、317、325 および 327～330 に開示される配列に実質的な同一性を有するポリヌクレオチド改変体を提供する。例えば、これらは、少なくとも 70% の配列同一性、好ましくは少なくとも 75%、80%、85%、90%、95%、96%、97%、98%、または 99% より高い配列同一性を含む。配列同一性は、本明細書中で記載の方法（例えば、以下に記載のような標準的なパラメーターを使用する BLAST 分析）を使用して本明細書のポリヌクレオチ

50

ド配列を比較した。当業者は、これらの値が、コドンの縮重、アミノ酸類似性、リーディングフレームの位置付けなどを考慮することによって、2つのヌクレオチド配列にコードされるタンパク質の対応する同一性を決定するように、適切に調整され得ることを理解する。

【0081】

代表的には、ポリヌクレオチド改変体は、1つ以上の置換、付加、欠失および／または挿入を含み、好ましくは、その結果、改変体ポリヌクレオチドによってコードされるポリペプチドの免疫原性が、本明細書中に具体的に示されるポリヌクレオチド配列によってコードされるポリペプチドに対して実質的に減少されない。用語「改変体」はまた、異種起源の相同遺伝子を包含することが理解されるべきである。

10

【0082】

さらなる実施形態において、本発明は、本明細書中に記載の配列の1つ以上に同一であるかまたは相補的な配列の種々の長さの連続したストレッチを含むポリヌクレオチドフラグメントを提供する。例えば、本発明に開示される配列の1つ以上の、少なくとも約10、15、20、30、40、50、75、100、150、200、300、400、500または1000以上連続したヌクレオチド、ならびにその間の中間の長さの連続したヌクレオチドを含むポリヌクレオチドが、本発明によって提供される。この状況において、「中間の長さ」が、示された値の間の任意の長さ（例えば、16、17、18、19など；21、22、23など；30、31、32など；50、51、52、53など；100、101、102、103など；150、151、152、153など；200～500；500～1,000などの間の全ての整数を含む）を意味することが容易に理解される。

20

【0083】

本発明の別の実施形態において、本明細書において提供されるポリヌクレオチド配列もしくはそのフラグメントまたはその相補的な配列に対して、中程度～高いストリンジェンシー条件下でハイブリダイズし得るポリヌクレオチド組成物が提供される。ハイブリダイゼーション技術は、分子生物学の当該分野において周知である。例示の目的であるが、他のポリヌクレオチドと本発明のポリヌクレオチドとのハイブリダイゼーションを試験するために適切な中程度にストリンジェントな条件は、5×SSC、0.5% SDS、1.0 mM EDTA (pH 8.0) の溶液中の事前洗浄；50℃～65℃、5×SSCでの、一晩のハイブリダイゼーション；続いて、0.1% SDSを含有する2×、0.5×および0.2×SSCのそれぞれを用いた、65℃で20分間の2回の洗浄を含む。当業者は、ハイブリダイゼーションのストリンジェンシーが、例えば、ハイブリダイゼーション溶液の塩含量および／またはハイブリダイゼーションが実施される温度を変更することによって、容易に操作され得ることを理解する。例えば、別の実施形態において、適切な高いストリンジェントのハイブリダイゼーション条件としては、ハイブリダイゼーションの温度が、例えば、60～65℃または65～70℃に増大される点を除いて、上記の条件が挙げられる。

30

【0084】

特定の好ましい実施形態において、上記のポリヌクレオチド（例えば、ポリヌクレオチド改変体、フラグメント、およびハイブリダイズする配列）は、本明細書中に具体的に示されるポリペプチド配列と免疫学的に交差反応するポリペプチドをコードする。他の好ましい実施形態において、このようなポリヌクレオチドは、本明細書中に具体的に示されるポリペプチド配列の免疫原性活性の、少なくとも約50%、好ましくは少なくとも約70%、そしてより好ましくは少なくとも約90%の免疫原性活性レベルを有するポリペプチドをコードする。

40

【0085】

本発明のポリヌクレオチド、またはそのフラグメントは、そのコード配列の長さに関わらず、他のDNA配列（例えば、プロモーター、ポリアデニル化シグナル、さらなる制限酵素部位、マルチクローニング部位、他のコードセグメントなど）と組合わされ得、その結

50

果、その全体の長さは、相当変化し得る。従って、ほとんどいずれの長さの核酸フラグメントをも使用し得ることが意図され、その全長は、好ましくは意図した組換えDNAプロトコルにおける調製および使用の容易さによって制限される。例えば、約10,000、約5000、約3000、約2,000、約1,000、約500、約200、約100、約50塩基対長など（全ての中間の長さを含む）の全長を有する例示的なポリヌクレオチドセグメントが、本発明の多くの実行において有用であることが意図される。

【0086】

ポリヌクレオチド配列を比較する場合、2つの配列におけるヌクレオチドの配列が、以下に記載されるように最大一致について整列されるのと同じ場合に、2つの配列が「同一」であるといわれる。2つの配列の間の比較は、代表的には、配列類似性の局部領域を同定および比較するために、比較ウィンドウにわたって配列を比較することによって行われる。本明細書中で使用される場合、「比較ウィンドウ」は、少なくとも約20、通常は30～約75、40～約50連続した位置のセグメントをいい、ここで、2つの配列が最適に整列された後に、配列は、連続した位置の同じ番号の参照配列と比較され得る。

【0087】

比較のための配列の最適な整列は、バイオインフォマティクスソフトウェアの *Laser gene suite* における *Megalign* プログラム (DNASTAR, Inc., Madison, WI) を使用して、デフォルトパラメーターを用いて行われ得る。このプログラムは、以下の参考文献に記載のいくつかの整列スキームを統合する：Dayhoff, M. O. (編) *Atlas of Protein Sequence and Structure*, National Biomedical Research Foundation, Washington DC Vol. 5, 補遺3, 345-358頁におけるDayhoff, M. O. (1978) *A model of evolutionary change in proteins—Matrices for detecting distant relationships.*; Hein J. (1990) *Unified Approach to Alignment and Phylogenies* 626-645頁 *Methods in Enzymology* vol. 183, Academic Press, Inc., San Diego, CA; Higgins, D. G. および Sharp, P. M. (1989) *CABIOS* 5:151-153; Myers, E. W. および Muller W. (1988) *CABIOS* 4:11-17; Robinson, E. D. (1971) *Comb. Theor* 11:105; Santou, N. Nes, M. (1987) *Mol. Biol. Evol.* 4:406-425; Sneath, P. H. A. および Sokal, R. R. (1973) *Numerical Taxonomy—the Principles and Practice of Numerical Taxonomy*, Freeman Press, San Francisco, CA; Wilbur, W. J. および Lipman, D. J. (1983) *Proc. Natl. Acad. Sci. USA* 80:726-730。

【0088】

あるいは、比較のための配列の最適な整列は、Smith および Waterman (1981) *Add. Appl. Math* 2:482 の局部同定アルゴリズムによって、Needleman および Wunsch (1970) *J. Mol. Biol.* 48:443 の同定整列アルゴリズムによって、Pearson および Lipman (1988) *Proc. Natl. Acad. Sci. USA* 85:2444 の類似性検索方法によって、これらのアルゴリズムのコンピュータ化した実行 (*Wisconsin Genetics Software Package*, Genetics Computer Group (GCG), 575 Science Dr., Madison, WI における GAP、BESTFIT、BLAST、FASTA、および TFASTA) によって、または検査によって行われ得る。

【0089】

配列同一性および配列類似性の割合を決定するために適切なアルゴリズムの1つの好ましい例は、BLASTおよびBLAST 2.0アルゴリズムであり、これらはそれぞれ、Altschulら(1977) Nucl. Acids Res. 25:3389-3402およびAltschulら(1990) J. Mol. Biol. 215:403-410に記載される。BLASTおよびBLAST 2.0は、本発明のポリヌクレオチドについての配列同一性の割合を決定するために、例えば、本明細書中に記載のパラメータを用いて使用され得る。BLAST分析を行うためのソフトウェアは、National Center for Biotechnology Informationを通して公に利用可能である。1つの例示的な例において、累積スコアは、ヌクレオチド配列について、パラメーターM（一致する残基の対についての報酬スコア（reward score）；常に >0 ）およびN（ミスマッチ残基についてのペナルティスコア；常に <0 ）を使用して算出され得る。各指示におけるワードヒットの拡大は、以下の場合に停止する：累積整列スコアが、その最大到達値から量Xだけ低下した場合；累積スコアが、1以上の負のスコアの残基整列の累積に起因してゼロ以下になった場合；またはいずれかの配列の末端に到達した場合。このBLASTアルゴリズムパラメーターW、TおよびXは、整列の感度および速度を決定する。BLASTNプログラム（ヌクレオチド配列について）は、ワード長さ（W）11、および期待値（E）10をデフォルトとして、そしてBLOSUM62スコアリングマトリックス（HenikoffおよびHenikoff(1989) Proc. Natl. Acad. Sci. USA 89:10915を参照のこと）アラインメントは、（B）50、期待値（E）10、M=5、N=-4、および両方の鎖の比較を使用する。

【0090】

好ましくは、「配列同一性のパーセンテージ」は、少なくとも20の位置の比較ウィンドウにわたる2つの最適に整列した配列を比較することによって決定され、ここで比較ウィンドウ中のポリヌクレオチド配列の一部は、参照配列（これは、付加または欠失を含まない）と比較して、2つの配列の最適な整列について20パーセント以下、通常は5～15パーセント、または10～12パーセントの付加または欠失（すなわち、ギャップ）を含み得る。このパーセンテージは、両方の配列で同一の核酸塩基が生じる位置の数を決定して一致する位置の数を得、この一致する位置の数を参照配列中の位置の総数（すなわち、ウィンドウサイズ）で除算し、そしてこの結果に100を掛けて配列同一性のパーセンテージを得ることによって算出される。

【0091】

遺伝コードの縮重の結果として、本明細書中に記載されるようなポリペプチドをコードする多くのヌクレオチド配列が存在することが、当業者に理解される。これらのポリヌクレオチドのうちいくつかは、任意のネイティブな遺伝子のヌクレオチド配列に対して最小の相同性を有する。それにもかかわらず、コドンの用法における差異に起因して変化するポリヌクレオチドは、本発明によって具体的に意図される。さらに、本明細書中に提供されるポリヌクレオチド配列を含む遺伝子の対立遺伝子は、本発明の範囲内である。対立遺伝子は、1以上の変異（例えば、ヌクレオチドの欠失、付加および／または置換）の結果として変化する内因性遺伝子である。得られたmRNAおよびタンパク質は、変化した構造または機能を有し得るが、有さなくてもよい。対立遺伝子は、標準的な技術（例えば、ハイブリダイゼーション、増幅および／またはデータベース配列比較）を用いて、同定され得る。

【0092】

従って、本発明の別の実施形態において、変異誘発アプローチ（例えば、部位特異的変異誘発）は、本明細書中に記載のポリペプチドの免疫原性改変体および／または誘導体の調製のために使用される。このアプローチによって、ポリペプチド配列における特定の改変が、これらをコードする、基礎となるポリヌクレオチドの変異誘発を介してなされ得る。これらの技術は、例えば、ポリヌクレオチド中に1つ以上のヌクレオチド配列変化を導入することによって先の考慮の1つ以上を具体化する、配列改変体を調製および試験するた

めの直接的なアプローチを提供する。

【0093】

部位特異的変異誘発は、所望の変異のDNA配列をコードする特定のオリゴヌクレオチド配列ならびに十分な数の隣接ヌクレオチドの使用を介して変異体の生成を可能にし、相対する欠失連結部の両側に安定な二重鎖を形成するために、十分なサイズのプライマー配列および配列複雑性を提供する。変異は、ポリヌクレオチド自体の特性を改善するか、変更するか、減少させるか、改変するか、さもなければ変化させるため、そして／またはコードされたポリペプチドの特性、活性、組成、安定性または一次配列を変更するために、選択されたポリヌクレオチド配列において使用され得る。

【0094】

本発明の特定の実施形態において、本発明者らは、コードされたポリペプチドの1つ以上の特性（例えば、ポリペプチドワクチンの免疫原性）を変更するために、開示されたポリヌクレオチド配列の変異誘発を意図する。部位特異的変異誘発の技術は、当該分野に周知であり、ポリペプチドおよびポリヌクレオチドの両方の改変体を作製するために広範に使用される。例えば、部位特異的変異誘発は、しばしば、DNA分子の特定部分を変更するために使用される。このような実施形態において、代表的に約14ヌクレオチド長～約25ヌクレオチド長程度の長さを含むプライマーが使用され、配列の連結部の両側にある、約5残基～約10残基が変更される。

【0095】

当業者によって理解されるように、部位特異的変異誘発技術は、しばしば、一本鎖形態および二本鎖形態の両方で存在するファージベクターを使用してきた。部位特異的変異誘発において有用である代表的なベクターとしては、例えば、M13ファージのようなベクターが挙げられる。これらのファージは、容易に商業的に入手可能であり、そしてそれらの使用は、一般的に当業者に周知である。二本鎖プラスミドもまた、目的の遺伝子をプラスミドからファージに転移する工程を除去する部位特異的変異誘発において慣用的に使用される。

【0096】

一般的に、本明細書に従う部位特異的変異誘発は、所望のペプチドをコードするDNA配列をその配列中に含む一本鎖ベクターを最初に入手する工程、またはこの配列を含む二本鎖ベクターの2本の鎖を融解して分ける工程によって実行される。所望の変異配列を保有するオリゴヌクレオチドプライマーは、一般的に、合成的に調製される。次いで、このプライマーは、一本鎖ベクターとともにアニーリングされ、そして変異を保有する鎖の合成を完了するために、DNA重合酵素（例えば、E. coli ポリメラーゼ I Klenowフラグメント）に供される。このようにヘテロ二重鎖が形成され、ここで一方の鎖が変異を有さないもともとの鎖をコードし、そして2つめの鎖は所望の変異を保有する。次いで、このヘテロ二重鎖ベクターは、適切な細胞（例えば、E. coli 細胞）を形質転換するために使用され、そして、変異した配列配置を保有する組換えベクターを含むクローンが選択される。

【0097】

部位特異的変異誘発を使用する、選択されたペプチドコードDNAセグメントの配列改変体の調製は、潜在的に有用な種を産生する手段を提供し、そしてこれは、限定を意味するものではない。なぜなら、ペプチドの配列改変体およびそれらをコードするDNA配列を入手し得る他の方法が存在するからである。例えば、所望のペプチド配列をコードする組換えベクターは、変異誘発剤（例えば、ヒドロキシルアミン）で処理されて、配列改変体を入手し得る。これらの方法およびプロトコルに関する具体的な詳細は、Maloyら、1994；Segal、1976；ProkopおよびBajpai、1991；Kuby、1994；およびManiatisら、1982（各々がその目的のために本明細書中に参考として援用される）の教示において見出され得る。

【0098】

本明細書中で使用される場合、用語「オリゴヌクレオチド特異的変異誘発手順」とは、鋳

10

20

30

40

50

型依存的プロセスおよびベクター媒介性増殖をいい、これは、その初期の濃度と比較して、特定の核酸分子の濃度の増加を生じるか、または検出可能なシグナルの濃度の増加（例えば、増幅）を生じる。本明細書中で使用される場合、用語「オリゴヌクレオチド特異的変異誘発手順」は、プライマー分子の鋳型依存的な伸長を含むプロセスをいうことが意図される。用語、鋳型依存的プロセスとは、RNA分子またはDNA分子の核酸合成をいい、ここで新規に合成された核酸の鎖の配列は、相補的塩基対形成の周知の規則によって指示される（例えば、Watson、1987を参照のこと）。代表的には、ベクター媒介性の方法論は、DNAまたはRNAベクターへの核酸フラグメントの導入、ベクターのクローン性増幅、および増幅した核酸フラグメントの回収を包含する。これらの方法論の例は、その全体が具体的に参考として本明細書中に援用される、米国特許第4,237,224号によって提供される。

【0099】

本発明のポリペプチド改変体の生成のための別のアプローチにおいて、米国特許第5,837,458号に記載のような、再帰的な配列組換えが使用され得る。このアプローチにおいて、組換えおよびスクリーニングまたは選択の反復性のサイクルが実施されて、例えば、増強された免疫原性活性を有する本発明の個々のポリヌクレオチド改変体を「展開させる」。

【0100】

本発明の他の実施形態において、本明細書中に提供されるポリヌクレオチド配列は、核酸ハイブリダイゼーションのためのプローブまたはプライマーとして、有利に使用され得る。このように、本明細書中に開示される15ヌクレオチド長の連続配列と同じか、またはこれに相補的な配列を有する、少なくとも約15ヌクレオチド長の連続配列の配列領域を含む核酸セグメントが特定の有用性を見出すことが、意図される。例えば、約20、30、40、50、100、200、500、1000（全ての中間的な長さのものを含む）およびまさに全長配列までの、より長い連続した同一配列または相補的な配列がまた、特定の実施形態において使用される。

【0101】

このような核酸プローブが目的の配列に特異的にハイブリダイズする能力は、所定のサンプル中の相補的配列の存在を検出する際にこれらが使用されることを可能にする。しかし、他の用途（例えば、変異種プライマーまたは他の遺伝的構築物の調製の際の使用のためのプライマーを調製するための配列情報の使用）がまた意図される。

【0102】

10～14、15～20、30、50、または100～200ヌクレオチド程度（同様に中間的な長さを含む）の連続したヌクレオチドストレッチからなる配列領域を有し、本明細書中に開示されるポリヌクレオチド配列に同一または相補的なポリヌクレオチド分子は、例えばサザンブロットティングまたはノーザンブロットティングにおける使用のためのハイブリダイゼーションプローブとして特に意図される。これは、遺伝子産物またはそのフラグメントの、多様な細胞型およびまた種々の細菌細胞の両方における分析を可能とする。フラグメントの全サイズ、ならびに相補的ストレッチのサイズは、究極的には、特定の核酸セグメントの意図される用途または適用に依存する。より小さなフラグメントは、一般にハイブリダイゼーション実施形態における用途を見出し、ここで連続した相補領域の長さが変化され得る（例えば、約15ヌクレオチドと約100ヌクレオチドとの間）が、検出を望む相補配列の長さに従って、より長い連続した相補ストレッチが使用され得る。

【0103】

約15～25ヌクレオチド長のハイブリダイゼーションプローブの使用は、安定で選択的な二重鎖分子の形成を可能にする。しかし、15塩基長より長いストレッチの連続した相補配列を有する分子が、ハイブリッドの安定性および選択性を増加するために一般に好ましく、それによって、得られる特異的ハイブリッド分子の質および程度が改善される。15～25の連続したヌクレオチド、または望まれる場合にはより長い遺伝子相補的ストレッチを有する核酸分子を設計するのが一般には好ましい。

10

20

30

40

50

【0104】

ハイブリダイゼーションプローブは、本明細書中に記載される配列のいずれかの、任意の部分から選択され得る。必要とされることの全ては、本明細書中に記載の配列、あるいはプローブまたはプライマーとしての利用を望む場合は、約15～25ヌクレオチド長から全長配列まで、および全長配列を含む配列の任意の連続した部分までを再検討することである。プローブおよびプライマー配列の選択は、種々の要因によって支配され得る。例えば、全配列の末端に向かってプライマーを使用することが望まれ得る。

【0105】

小さなポリヌクレオチドセグメントまたはフラグメントは、通常は自動オリゴヌクレオチド合成機を使用して行われるので、例えば化学的手段によってフラグメントを直接合成することによって、容易に調製され得る。また、フラグメントは、核酸複製技術（例えば、米国特許第4,683,202号（本明細書中に参考として援用される）のPCRTM技術）の適用によって、組換え産物のために選択配列を組換えベクターに導入することによって、および一般に分子生物学の分野の当業者に公知の他の組換えDNA技術によって、得られ得る。

10

【0106】

本発明のヌクレオチド配列は、目的の完全遺伝子または遺伝子フラグメントのいずれかの相補的ストレッチと、二重鎖分子を選択的に形成するその能力のために使用され得る。想定される適用に依存して、代表的には、標的配列に対するプローブの選択性の種々の程度に達するために、ハイブリダイゼーションの種々の条件を使用することが望ましい。高い選択性を必要とする適用について、代表的には、ハイブリッドを形成するための比較的ストリンジェントな条件（例えば、比較的低濃度の塩、および／または高温の条件（例えば、約50℃～約70℃の温度で、約0.02M～約0.15Mの塩の塩濃度によって提供されるような）を使用することが望ましい。このような選択条件は、存在する場合、プローブとテンプレートまたは標的鎖との間のミスマッチにわずかに許容性であり、そして関連する配列の単離のために特に適切である。

20

【0107】

当然のことながら、いくつかの適用（例えば、潜在的なテンプレートにハイブリダイズする変異体プライマー鎖を使用する、変異体の調製が望ましい場合）について、低ストリンジェント（減少したストリンジェンシー）のハイブリダイゼーション条件は、ヘテロ二重鎖を形成するために代表的に必要とされる。これらの状況において、約20℃～約55℃の温度範囲で約0.15M～約0.9Mの塩の条件のような塩条件を使用することが望ましくあり得る。これによって、交差ハイブリダイズ種は、コントロールハイブリダイゼーションに関して、陽性ハイブリダイズシグナルとして容易に同定され得る。任意の場合において、ホルムアミド（これは、増加した温度と同じ様式で、ハイブリッド二重鎖を不安定化するように作用する）の増加した量の添加によって、条件がよりストリンジェントになり得ることが一般に理解される。従って、ハイブリダイゼーション条件は、容易に操作され得、従って、一般に、所望の結果に依存する最良の方法である。

30

【0108】

本発明の別の実施形態に従って、アンチセンスオリゴヌクレオチドを含むポリヌクレオチド組成物が、提供される。アンチセンスオリゴヌクレオチドは、タンパク質合成の効果的かつ標識化インヒビターであることが実証され、疾患が、疾患に起因するタンパク質の合成を阻害することによって処置され得ることによる治療的なアプローチを結果的に提供する。タンパク質合成を阻害するためのアンチセンスオリゴヌクレオチドの効力は、よく確立されている。例えば、ポリガラクタウロナーゼ（polygalacturonase）の合成およびムスカリン2型アセチルコリンレセプターは、そのそれぞれのmRNA配列に特異的なアンチセンスオリゴヌクレオチドによって阻害される（米国特許第5,739,119号および米国特許第5,759,829号）。さらに、アンチセンス阻害の例が、核タンパク質サイクリン、多剤耐性遺伝子（MDG1）、ICAM-1、E-セレクチン、STK-1、線条体GABAAレセプターおよびヒトEGFを用いて示されてい

40

50

る (Jaskulskiら、Science 1988 Jun 10; 240 (4858): 1544-6; VasanthakumarおよびAhmed、Cancer Commun. 1989; 1 (4): 225-32; Perisら、Brain Res Mol Brain Res. 1998 Jun 15; 57 (2) 310-20; 米国特許第5, 801, 154号; 同第5, 789, 573号; 同第5, 718, 709号および同第5, 610, 288号)。種々の異常な細胞増殖 (例えば、癌) を阻害しそして処置するために使用され得る、アンチセンス構築物もまた、記載されている (米国特許第5, 747, 470号; 同第5, 591, 317号および同第5, 783, 683号)。

【0109】

従って、特定の実施形態において、本発明は、本明細書中に記載されるポリヌクレオチド配列に特異的に結合し得る任意の配列もしくはその相補体のすべてまたは一部を含む、オリゴヌクレオチド配列を提供する。1つの実施形態において、そのアンチセンスオリゴヌクレオチドは、DNAまたはその誘導体を含む。別の実施形態において、そのオリゴヌクレオチドは、RNAまたはその誘導体を含む。第3の実施形態において、そのオリゴヌクレオチドは、ホスホロチオエート改変骨格を含む、改変DNAである。第4の実施形態において、そのオリゴヌクレオチド配列は、ペプチド核酸またはその誘導体を含む。各場合において、好ましい組成物は、本明細書中に開示されるポリヌクレオチドのうちの1つ以上の部分に相補的な、より好ましくは実質的に相補的な、そしてさらにより好ましくは完全に相補的な、配列領域を含む。所定の遺伝子配列に特異的なアンチセンス組成物の選択は、選択された標的配列の分析に基き、そして二次構造、 T_m 、結合エネルギー、および相対的安定性の決定に基く。アンチセンス組成物は、二量体、ヘアピン、または宿主細胞において標的mRNAへの特異的結合を減少または妨げる他の二次構造をそれらが形成できない相対的能力に基づいて選択され得る。mRNAの非常に好ましい標的領域は、AUG翻訳開始コドンの領域またはその付近の領域、およびmRNAの5'領域と実質的に相補的な配列である。これらの二次構造分析および標的部位選択の考慮は、例えば、OLIGOプライマー分析ソフトウェアのv. 4およびBLASTN 2.0.5アルゴリズムソフトウェア (Altschulら、Nucleic Acids Res. 1997, 25 (17): 3389-402) を使用して実施され得た。

【0110】

短いペプチドベクター (MPG (27残基) と称する) を使用するアンチセンス送達法の使用もまた意図される。このMPGペプチドは、HIV gp41の融合配列由来の疎水性ドメインと、SV40 T抗原の核局在化配列由来の親水性ドメインとを含む (Morrisら、Nucleic Acids Res. 1997 Jul 15; 25 (14): 2730-6)。このMPGペプチドのいくつかの分子がアンチセンスオリゴヌクレオチドをコートし、そして比較的高効率 (90%) で1時間未満で培養哺乳動物細胞へと送達され得ることが示された。さらに、MPGとの相互作用は、ヌクレアーゼに対するそのオリゴヌクレオチドの安定性および形質膜を横切る能力の両方を強力に増加させる。

【0111】

本発明の別の実施形態に従って、本明細書中に記載されるポリヌクレオチド組成物は、腫瘍ポリヌクレオチドおよび腫瘍細胞における本発明のタンパク質の発現を阻害するためのリボンザイム分子の設計および調製に使用される。リボザイムは、部位特異的様式で核酸を切断するRNA-タンパク質複合体である。リボザイムは、エンドヌクレアーゼ活性を保有する特異的触媒ドメインを有する (KimおよびCech、Proc Natl Acad Sci U S A. 1987 Dec; 84 (24): 8788-92; ForsterおよびSymons、Cell. 1987 Apr 24; 49 (2) 211-20)。例えば、多数のリボザイムが、高い程度の特異性でホスホエステル転移反応を促進し、しばしば、オリゴヌクレオチド基質中の数個のホスホエステルのうちの1つだけを切断する (Cechら、Cell. 1981 Dec; 27 (3 Pt 2): 487-96; MichelおよびWesthof、J Mol Biol. 1990 Dec 5; 216 (3): 585-610; Reinhold-HurekおよびShub、

Nature, 1992 May 14; 357 (6374): 173-6)。この特異性は、この基質が、化学反応の前にリボザイムの内部ガイド配列（「IGS」）との特異的塩基対形成相互作用を介して結合するという要件に帰せられている。

【0112】

天然に存在する酵素的RNAの6つの基本的変種が、現在公知である。各々が、生理学的条件下で、トランスで、RNAホスホジエステル結合の加水分解を触媒し得る（従って、他のRNA分子を切断し得る）。一般に、酵素的核酸は、まず、標的RNAへの結合によって作用する。このような結合は、標的RNAを切断するように作用する分子の酵素的部分に近接して保持される、酵素的核酸の標的結合部分を介して生じる。従って、その酵素的核酸はまず、標的RNAを認識し、次いで相補的塩基対形成を介してその標的RNAに結合し、そして一旦正確な部位に結合すると、その標的RNAを酵素的に切断するように作用する。このような標的RNAの戦略的切断は、コードされるタンパク質を直接合成する標的RNAの能力を破壊する。酵素的核酸がそのRNA標的に結合しそして切断した後、その酵素的核酸は、別の標的を探索するためにそのRNAから離れ、そして繰り返し、新しい標的に結合しそして切断し得る。

10

【0113】

リボザイムの酵素特性は、多くの技術（例えば、アンチセンス技術（ここで、核酸分子は、翻訳を阻害するために核酸標的に簡単に結合する））に対して有利である。なぜなら、治療的処置に影響を与えるのに必要であるリボザイムの濃度は、アンチセンスオリゴヌクレオチドの濃度よりも低いからである。この利点は、酵素学的に作用するリボザイムの能力に影響を与える。従って、単一のリボザイム分子は、標的RNAの多くの分子を切断し得る。さらに、このリボザイムは、高度に選択的なインヒビターであり、この阻害の特異性は、標的RNAに結合する塩基対形成の機構に依存するだけでなく、標的RNAの切断の機構にも依存する。切断部位付近の1個のミスマッチまたは塩基置換により、リボザイムの触媒活性が完全になくなり得る。アンチセンス分子内の同様のミスマッチでは、それらの作用は妨げられない（Woolfら、Proc Natl Acad Sci U S A, 1992 Aug 15; 89 (16): 7305-9）。従って、リボザイムの作用の特異性は、同じRNA部位に結合するアンチセンスオリゴヌクレオチドの作用の特異性よりも高い。

20

【0114】

この酵素学的核酸分子は、ハンマーヘッド（hammerhead）モチーフ、ヘアピンモチーフ、 δ 型肝炎ウイルスモチーフ、I群イントロンモチーフまたはRNaseP RNAモチーフ（RNA誘導配列に関連する）あるいはNeurospora VS RNAモチーフに形成され得る。ハンマーヘッドモチーフの例は、Rossiら、Nucleic Acids Res. 1992 Sep 11; 20 (17): 4559-65に記載される。ヘアピンモチーフの例は、Hampelら（欧州特許出願公開番号EP 0360257）、HampelおよびTritz（Biochemistry 1989 Jun 13; 28 (12): 4929-33）、Hampelら、Nucleic Acids Res. 1990 Jan 25; 18 (2): 299-304および米国特許第5,631,359号に記載される。 δ 型肝炎ウイルスモチーフの例は、PerrottaおよびBeen, Biochemistry. 1992 Dec 1; 31 (47): 11843-52に記載される；RNasePモチーフの例は、Guerrier-Takadara, Cell. 1983 Dec; 35 (3 Pt 2): 849-57に記載される；Neurospora VS RNAリボザイムモチーフは、Collins（SavilleおよびCollins, Cell. 1990 May 18; 61 (4): 685-96；SavilleおよびCollins, Proc Natl Acad Sci U S A. 1991 Oct 1; 88 (19): 8826-30；CollinsおよびOlive, Biochemistry. 1993 Mar 23; 32 (11): 2795-9）に記載される；そしてI群イントロンの例は、米国特許第4,987,071号に記載される。本発明の酵素学的核酸分子において重要であるこ

30

40

50

とは、1以上の標的遺伝子RNA領域に相補的である特定の基質結合部位を有すること、およびこの分子にRNA切断活性を付与する、基質結合部位内のヌクレオチド配列またはこの基質結合部位の周囲のヌクレオチド配列を有することだけである。従って、リボザイム構築物は、本明細書中に記載される特定のモチーフに限定する必要はない。

【0115】

リボザイムは、国際特許出願公開番号WO 93/23569および国際特許出願公開番号WO 94/02595（各々は、本明細書中で参考として詳細に援用される）に記載されるように設計され得、そして記載されるようにインビトロおよびインビボにおいて試験するために合成され得る。このようなリボザイムはまた、送達するために最適化され得る。特定の例が提供されるが、当業者は、他の種において等価なRNA標的が、必要な場合に利用され得ることを理解する。

10

【0116】

リボザイムの活性は、リボザイムの結合アームの長さを変更することにより、または血清リボヌクレアーゼによる分解を防ぐ改変（例えば、国際特許出願番号WO 92/07065；国際特許公開番号WO 93/15187；国際特許出願公開番号WO 91/03162；欧州特許出願公開番号92110298.4；米国特許第5,334,711号、および国際特許出願公開番号WO 94/13688（これには、酵素学的RNA分子の糖部分に対して成され得る、種々の化学的修飾が記載される）を参照のこと）、細胞中でのリボザイムの有効性を増強する改変、ならびにRNAの合成時間を短縮および化学的必要性を低下させるための幹（stem）II塩基を除去して、リボザイムを化学的に合成することによって最適化され得る。

20

【0117】

Sullivanら（国際特許出願公開番号WO 94/02595）は、酵素学的RNA分子の送達のための一般的な方法を記載する。リボザイムは、当業者に公知の種々の方法によって細胞に投与され得、これらには、リボソームへのカプセル化、イオン泳動法、あるいは、例えば、ヒドロゲル、シクロデキストリン、生分解性ナノカプセルおよび生体接着性ミクロスフィアのような他のビヒクルへの取り込みによるもの、が挙げられるが、これらに限定されない。いくつかの指針のために、リボザイムは、エキソビボで細胞または組織に、上記のビヒクルを使用してか、または使用せずに直接送達され得る。あるいは、RNA/ビヒクルの組合せは、直接的な吸入、直接的な注射、またはカテーテル、注入ポンプまたはステントの使用により、局所的に送達され得る。他の送達経路としては、血管内注射、筋内注射、皮下注射または関節注射、エアロゾル吸入、経口送達（錠形態または丸薬形態）、局所送達、全身送達、眼送達、腹腔内送達および／またはくも膜下腔内送達（錠形態または丸薬形態）が挙げられるが、これらに限定されない。リボザイム送達および投与のより詳細な記載は、国際特許出願公開番号WO 94/02595および国際特許出願公開番号WO 93/23569（各々は、本明細書中で参考として詳細に援用される）に提供される。

30

【0118】

高濃度のリボザイムを細胞内に蓄積する別の手段は、リボザイムコード配列をDNA発現ベクターに組み込むことである。リボザイム配列の転写は、真核生物RNAポリメラーゼI（pol I）、RNAポリメラーゼII（pol II）、またはRNAポリメラーゼIII（pol III）のためのプロモーターにより駆動される。pol IIプロモーターまたはpol IIIプロモーターからの転写物は、全ての細胞内で高レベルで発現される；所定の細胞型における所定のpol IIプロモーターのレベルは、近くに存在する遺伝子調節配列（エンハンサー、サイレンサーなど）の性質に依存する。原核生物のRNAポリメラーゼプロモーターもまた使用され得るが、但し、原核生物RNAポリメラーゼ酵素は、適切な細胞において発現される。このようなプロモーターから発現するリボザイムは、哺乳動物細胞において機能し得る。このような転写ユニットは、哺乳動物細胞内への導入のための種々のベクター（プラスミドDNAベクター、ウイルスDNAベクター（例えば、アデノウイルスベクターまたはアデノ随伴ベクター）、またはウイルスRNAベクター（例えば、レトロウイルスベクター、セムリキ森林ウイルスベクター、シ

40

50

ンドビスウイルスベクター)が挙げられるがこれらに限定されない)に組み込まれ得る。

【0119】

本発明の別の実施形態において、ペプチド核酸(PNA)組成物が提供される。PNAは、DNA模倣物であり、核酸塩基は、偽ペプチド骨格に結合される(GoodおよびNielsen、Antisense Nucleic Acid Drug Dev. 1997 7(4)431~37)。PNAは、RNAまたはDNAを伝統的に使用した多くの方法において利用され得る。しばしば、PNA配列は、対応するRNA配列またはDNA配列よりも技術的により良く実施され、そしてRNAにもDNAにも固有でない有用性を有する。産生方法、特徴、および使用方法を含むPNAの総説は、Corey(Trends Biotechnol 1997 6月;15(6):224~9)によって提供される。このように、特定の実施形態において、ACE mRNA配列の1以上の部分に相補的であるPNA配列を調製し得、そしてこのようなPNA組成物は、ACE-特異的mRNAの翻訳を調節、変更、減少、または低下させるために使用され得、これにより、このようなPNA組成物が投与された宿主細胞におけるACE活性のレベルを変更する。

10

【0120】

PNAは、DNAの正常なホスホジエステル骨格を置換する2-アミノエチル-グリシン連結を有する(Nielsenら、Science 1991 12月6;254(5037):1497~500;Hanveyら、Science. 1992 11月27;258(5087):1481~5;HyrupおよびNielsen, Bioorg Med Chem. 1996 1月;4(1):5~23)。この化学は、3つの重要な結果を有する。第1に、DNAまたはホスホロチオエートオリゴヌクレオチドと比較して、PNAは、中性の分子であり;第2に、PNAは、アキラル(立体選択的な合成を開発する必要を避ける)であり;そして第3に、PNA合成は、標準的なBocプロトコルまたはFmocプロトコルを、固相ペプチド合成に使用する。一方、改良型Merrifield法を含む他の方法が使用されている。

20

【0121】

PNAモノマーまたは既製のオリゴマーは、PerSeptive Biosystems(Framingham, MA)から市販されている。BocプロトコルまたはFmocプロトコルのいずれかによるPNAの合成は、手動プロトコルまたは自動プロトコルを使用して簡単に実施される(Nortonら、Bioorg Med Chem. 1995 4月;3(4):437~45)。このマニュアルプロトコルは、それ自体で、化学的に修飾されたPNAの産生または密接に関係するPNAのファミリーの同時合成を導く。

30

【0122】

ペプチドの合成に関して、特定のPNA合成の成功は、選択した配列の特性に依存する。例えば、理論上、PNAは、ヌクレオチド塩基の任意の組合せを取り込み得るが、隣接するプリンの存在は、産物中に1以上の残基の欠失を導き得る。この困難性を見込みにおいて、隣接するプリンを有するPNAを産生する際に、非効率的に加えられているような残基の連結を繰り返すべきであることを提案する。この後、ペプチド合成の間に観測されるのと類似する産物の収率および純度を提供する逆相高圧液体クロマトグラフィーによってPNAを精製すべきである。

40

【0123】

所与の適用のためのPNAの修飾が、固相合成の間にアミノ酸を連結することによって、または露出されたN末端アミンにカルボン酸基を含む化合物を付加することによって達成され得る。あるいは、PNAは、導入されたリジンまたはシステインに連結することによって合成した後に修飾され得る。PNAが修飾され得る簡便さにより、より良い溶解度または特定の機能的必要条件の最適化を容易にする。一旦合成されると、PNAおよびそれらの誘導体の同定は、質量分析法によって確認され得る。いくつかの研究が、PNAの修飾物を作製および利用した(例えば、Nortonら、Bioorg Med Chem

50

. 1995 4月; 3 (4) : 437~45; Petersenら、J Pept Sci. 1995 5~6月; 1 (3) 175~83; Orumら、Biotechniques. 1995 9月; 19 (3) : 472~80; Footerら、Biochemistry. 1996 8月 20; 35 (33) : 10673~9; Griffithsら、Nucleic Acid Res. 1995 8月 11; 23 (15) : 3003~8; Pardridgeら、Proc Natl Acad Sci USA 1995 6月 6; 92 (12) : 5592~6; Boffaら、Proc Natl Acad Sci USA 1995 3月 14; 92 (6) : 1901~5; Gambacorti-Passeriniら、Blood、1996 8月 15; 88 (4) : 1411~7; Armitageら、Proc Natl Acad Sci USA、1997 11月 11; 94 (23) : 12320~5; Seegerら、Biotechniques、1997 9月; 23 (3) : 512~7)。米国特許第5, 700, 922号は、PNA-DNA-PNAキメラ分子および診断におけるその分子の使用、生体におけるタンパク質の調節、ならびに治療に対して影響を受け易い状態の処置について議論する。

【0124】

PNAのアンチセンス結合特性を特徴づける方法は、Rose (Anal Chem、1993 12月 15; 65 (24) : 3545~9) およびJensenら (Biochemistry、1997 4月 22; 36 (16) : 5072~7) において議論される。Roseは、キャピラリーゲル電気泳動を使用して、相対的な結合速度論および化学量論を測定することで、PNAのそれらの相補的オリゴヌクレオチドに対する結合を決定する。同様のタイプの測定が、BIAcoreTM 技術を使用してJensenらによって成された。

【0125】

記載され、そして当業者に明らかであるPNAの他の適用には、DNA鎖の侵入、アンチセンス阻害、変異分析、転写のエンハンサー、核酸の精製、転写活性遺伝子の単離、転写因子結合の阻害、ゲノムの切断、バイオセンサー、インサイチュハイブリダイゼーションなどにおける使用が挙げられる。

【0126】

(ポリヌクレオチドの同定、特徴づけおよび発現)

本発明のポリヌクレオチド組成物は、種々の十分に確立された技術のいずれかを使用して、同定、調製および／または操作され得る（一般に、Sambrookら、Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratories, Cold Spring Harbor, NY, 1989、および他の類似の参考文献を参照のこと）。例えば、ポリヌクレオチドは、cDNAのマイクロアレイを腫瘍に関連する発現（すなわち、本明細書中で提供される代表的なアッセイを使用して決定する場合、正常な組織における発現よりも、腫瘍での少なくとも2倍高い発現）についてスクリーニングすることによって、同定され得る。このようなスクリーニングは、例えば、製造者の指示に従って、Affymetrix, Inc. (Santa Clara, CA) のマイクロアレイ技術を使用して、（そして本質的にSchenaら、Proc. Natl. Acad. Sci. USA 93: 10614~10619, 1996およびHellerら、Proc. Natl. Acad. Sci. USA 94: 2150~2155, 1997に記載されるように）、行われ得る。あるいは、ポリヌクレオチドは、本明細書中に記載されるタンパク質を発現する細胞（例えば、腫瘍細胞）から調製されるcDNAから増幅され得る。

【0127】

多数の鋳型依存性プロセスが、サンプル中に存在する目的の標的配列を増幅するために利用可能である。最もよく知られた増幅方法の1つは、ポリメラーゼ連鎖反応（PCRTM）であり、これは、米国特許第4, 683, 195号、同第4, 683, 202号、および同第4, 800, 159号に詳細に記載され、これらの各々はその全体が本明細書中に

参考として援用される。手短に言えば、PCR^{T M}においては、標的配列の向かい合った相補鎖上の領域に対して相補的な2つのプライマー配列が調製される。過剰のデオキシヌクレオシド三リン酸を、DNAポリメラーゼ（例えば、Taqポリメラーゼ）とともに反応混合液に添加する。標的配列がサンプル中に存在する場合、プライマーはその標的に結合し、そしてポリメラーゼは、ヌクレオチドを付加することによって標的配列に沿ってプライマーを伸長させる。反応混合液の温度を上昇および下降させることによって、伸長したプライマーは、標的から解離して反応生成物を形成し、過剰のプライマーは標的および反応生成物に結合し、そしてこのプロセスが反復される。好ましくは、逆転写およびPCR^{T M}増幅手順が、増幅されたmRNAの量を定量するために実行され得る。ポリメラーゼ連鎖反応方法論は、当該分野で周知である。

10

【0128】

多くの他の鋳型依存性プロセスのいずれか（この多くは、PCR^{T M}増幅技術のバリエーションである）は、当該分野で容易に知られており、利用可能である。例示として、いくつかのこのような方法としては、リガーゼ連鎖反応（LCRともいわれる）（欧州特許出願公開番号320,308号、米国特許第4,883,750号において記載される）；Q β レプリカーゼ（PCT国際特許出願公開番号PCT/US87/00880に記載される）；鎖置換増幅（Strand Displacement Amplification）（SDA）および修復鎖反応（Repair Chain Reaction）（RCR）が挙げられる。なお他の増幅法は、英国特許出願番号2202328号およびPCT国際特許出願公開番号PCT/US89/01025において記載される。他の核酸増幅手順には、転写に基づく増幅系（TAS）（PCT国際特許出願公開番号WO88/10315）が挙げられ、これには、核酸配列に基づく増幅（NASBA）および3SRが含まれる。欧州特許出願公開番号329,822は、一本鎖RNA（「ssRNA」）、ssDNA、および二本鎖DNA（dsDNA）をサイクル的に合成する工程を包含する核酸増殖プロセスを記載する。PCT国際特許出願公開番号WO89/06700は、プロモーター／プライマー配列の標的一本鎖DNA（「ssDNA」）へのハイブリダイゼーションに基づく核酸配列増幅スキーム、引き続くこの配列の多くのRNAコピーの転写を記載する。「RACE」（Frohman、1990）および「片側（one-sided）PCR」（Ohara、1989）のような他の増幅方法はまた、当業者に周知である。

20

30

【0129】

本発明のポリヌクレオチドの増幅した部分を使用して、適切なライブラリー（例えば、腫瘍cDNAライブラリー）から周知の技術を使用して全長遺伝子を単離し得る。このような技術において、増幅に適した1以上のポリヌクレオチドプローブまたはプライマーを使用して、ライブラリー（cDNAライブラリーまたはゲノムライブラリー）をスクリーニングする。好ましくは、ライブラリーはより大きな分子を含むようにサイズが選択される。無作為に準備されたライブラリー（random primed library）もまた、遺伝子の5'領域および上流領域の同定ために好ましくあり得る。ゲノムライブラリーは、イントロンを入手することおよび5'配列を伸長させることについて好ましい。

【0130】

ハイブリダイゼーション技術に関して、部分配列は、周知の技術を使用して標識（例えば、ニックトランスレーションまたは³²Pでの末端標識）され得る。次いで、細菌ライブラリーまたはバクテリオファージライブラリーは、一般に、標識プローブと、変性した細菌コロニー（またはファージプラークを含む菌叢）を含むフィルターとをハイブリダイズすることによってスクリーニングされる（Sambrookら、Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratories, Cold Spring Harbor, NY, 1989を参照のこと）。ハイブリダイズするコロニーまたはプラークを選択し、そして増殖させる。そしてそのDNAをさらなる分析のために単離する。cDNAクローンを、付加配列の量を決定するために、例えば、部分配列由来のプライマーおよびそのベクタ

40

50

ー由来のプライマーを使用するPCRによって分析し得る。制限地図および部分配列を作成して、1以上の重複クローンを同定し得る。次いで、標準的な技術（これは、一連の欠失クローンを作製することを包含し得る）を使用して完全配列を決定し得る。次いで、得られた重複配列を1つの連続配列中に構築し得る。周知の技術を使用して、適切なフラグメントに連結することにより全長cDNA分子を生成し得る。

【0131】

あるいは、上記のような増幅技術は、部分的なcDNA配列から全長コード配列を得るために有用であり得る。このような増幅技術の1つは、インバースPCRである（Trigliaら、Nuc l. A c i d s R e s. 16:8186, 1988を参照のこと）。この技術は、制限酵素を使用して、その遺伝子の既知の領域内のフラグメントを生成する。次いで、このフラグメントを分子内連結により環状化し、そして既知の領域に由来する多岐したプライマーを用いたPCRのための鋳型として使用する。代替のアプローチにおいて、部分配列に隣接した配列を、リンカー配列に対するプライマーおよび既知の領域に特異的なプライマーを使用する増幅により取り出し得る。この増幅した配列を、代表的に、同じリンカープライマーおよび既知の領域に特異的な第2のプライマーを使用する2回目の増幅に供する。既知の配列から反対方向に伸長を開始する2つのプライマーを使用するこの手順についての改変は、WO 96/38591に記載される。そのような別の技術は、「迅速なcDNA末端の増幅」またはRACEとして公知である。この技術は、公知配列の5'および3'である配列を同定するために、内側プライマーおよび外側プライマー（これらは、ポリA領域またはベクター配列にハイブリダイズする）の使用を含む。さらなる技術としては、キャプチャーPCR（Langerstromら、PCR Methods Applic. 1:111-19, 1991）およびウォーキングPCR（Parkerら、Nuc l. A c i d s R e s. 19:3055-60, 1991）が挙げられる。増幅を利用する他の方法もまた使用して、全長cDNA配列を入手し得る。

【0132】

特定の場合において、発現配列タグ（EST）データベース（例えば、GenBankより入手可能なもの）に提供される配列の分析により、全長cDNA配列を入手することが可能である。重複ESTの検索は、一般に、周知のプログラム（例えば、NCBI BLAST検索）を使用して行われ得、そしてこのようなESTを使用して連続した全長配列を生成し得る。全長DNA配列はまた、ゲノムフラグメントの分析により入手され得る。

【0133】

本発明の他の実施形態において、本発明のポリペプチドまたはその融合タンパク質もしくは機能的等価物をコードするポリヌクレオチド配列またはそのフラグメントは、適切な宿主細胞におけるポリペプチドの発現に方向付けるように組換えDNA分子において使用され得る。遺伝コードの固有の縮重に起因して、実質的に同じか、または機能的に等価なアミノ酸配列をコードする他のDNA配列が産生され得、そしてこれらの配列を使用して所定のポリペプチドをクローン化し、そして発現させ得る。

【0134】

当業者によって理解されるように、いくつかの場合において、天然に存在しないコドンをも所有するポリペプチドコードヌクレオチド配列を作製することが有利であり得る。例えば、特定の原核生物宿主または真核生物宿主に好まれるコドンは、タンパク質発現の速度を増加させるためか、または所望の特性（例えば、天然に存在する配列から生成される転写物の半減期よりも長い半減期）を有する組換えRNA転写物を産生するように選択され得る。

【0135】

さらに、本発明のポリヌクレオチド配列は、種々の理由（遺伝子産物のクローニング、プロセッシング、および／または発現を変更する改変が挙げられるが、それらに限定されない）のためにポリペプチドコード配列を変更するために、当該分野において一般的に公知の方法を使用して操作され得る。例えば、無作為切断によるDNAシャッフリングならびに遺伝子フラグメントおよび合成オリゴヌクレオチドのPCR再アセンブリを使用して、ヌ

クレオチド配列を操作し得る。さらに、部位指向性変異誘発 (site-directed mutagenesis) を使用して、新たな制限部位を挿入し得るか、グリコシル化パターンを改変し得るか、コドン優先 (preference) を変化させ得るか、スプライス改変体を作製し得るか、または変異を導入したりなどし得る。

【0136】

本発明の別の実施形態において、天然の核酸配列、改変された核酸配列、または組換え核酸配列は、融合タンパク質をコードする異種配列に連結され得る。例えば、ポリペプチド活性のインヒビターについてペプチドライブラリーをスクリーニングするために、市販の抗体により認識され得るキメラタンパク質をコードすることが有用であり得る。融合タンパク質はまた、ポリペプチドコード配列と異種タンパク質配列との間に位置する切断部位を含むように操作され得、その結果そのポリペプチドは、切断されて、そして異種部分から精製されて取り出される。

10

【0137】

所望のポリペプチドをコードする配列が、当該分野において周知の化学的方法を使用して、全体または部分的に合成され得る (Caruthers, M. H. ら、(1980) Nucl. Acids Res. Symp. Ser. 215~223、Horn, T. ら、(1980) Nucl. Acids Res. Symp. Ser. 225~232を参照のこと)。あるいは、タンパク質自体は、ポリペプチドまたはその一部のアミノ酸配列を合成するための化学的方法を使用して産生され得る。例えば、ペプチド合成は、種々の固相技術を使用して実施され得 (Roberge, J. Y. ら、(1995) Science 269:202~204)、そして自動化合成は、例えば、ABI 431A Peptide Synthesizer (Perkin Elmer, Palo Alto, CA) を使用して達成され得る。

20

【0138】

新たに合成されたペプチドは、分取用高速液体クロマトグラフィー (例えば、Creighton, T. (1983) Proteins, Structures and Molecular Principles, WH Freeman and Co., New York, N. Y.) または当該分野において利用可能な他の同等技術により実質的に精製され得る。合成ペプチドの組成は、アミノ酸分析または配列決定 (例えば、エドマン分解手順) により確認され得る。さらに、ポリペプチドまたはその任意の部分のアミノ酸配列は、直接合成の間改変されて、そして／または化学的方法を使用して、他のタンパク質もしくはその任意の部分に由来する配列と組み合わせられて、改変体ポリペプチドを産生し得る。

30

【0139】

所望のポリペプチドを発現させるために、ポリペプチドをコードするヌクレオチド配列または機能的等価物は、適切な発現ベクター (すなわち、挿入されたコード配列の転写および翻訳のために必要なエレメントを含むベクター) 内に挿入され得る。当業者に周知である方法を使用して、目的のポリペプチドをコードする配列および適切な転写制御エレメントおよび翻訳制御エレメントを含む発現ベクターを構築し得る。これらの方法としては、インビトロ組換えDNA技術、合成技術、およびインビボ遺伝子組換えが挙げられる。そのような技術は、例えば、Sambrook, J. ら、(1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview, N. Y., および Ausubel, F. M. ら、(1989) Current Protocols in Molecular Biology, John Wiley & Sons, New York, N. Y. に記載される。

40

【0140】

種々の発現ベクター／宿主系が、ポリヌクレオチド配列を含み、そして発現させるように利用され得る。これらとしては、以下が挙げられるが、これらに限定されない：微生物 (例えば、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

50

形質転換した細菌)；酵母発現ベクターで形質転換した酵母；ウイルス発現ベクター（例えば、バキュロウイルス）で感染させた昆虫細胞系；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；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もしくはpBR322プラスミド）で形質転換した植物細胞系；あるいは動物細胞系。

【0141】

発現ベクター内に存在する「制御エレメント」または「調節配列」は、転写および翻訳を実行するように宿主細胞タンパク質と相互作用する、それらのベクターの非翻訳領域（エンハンサー、プロモーター、5'および3'非翻訳領域）である。そのようなエレメントは、その長さおよび特異性を変更し得る。利用されるベクター系および宿主に依存して、多数の適切な転写エレメントおよび翻訳エレメント（構成的プロモーターおよび誘導性プロモーターを含む）が使用され得る。例えば、細菌系においてクローニングする場合、誘導性プロモーター（例えば、PBLUESCRIPTファージミド（Stratagene, La Jolla, Calif.）またはPSPORT1プラスミド（Gibco BRL, Gaithersburg, MD）などのハイブリッドlacZプロモーターが使用され得る。哺乳動物細胞系において、哺乳動物遺伝子または哺乳動物ウイルスに由来するプロモーターが、一般的に好ましい。ポリペプチドをコードする配列の複数のコピーを含む細胞株を生成することが必要とされる場合、SV40またはEBVベースのベクターが、適切な選択マーカーと共に有利に使用され得る。

10

【0142】

細菌系において、多数の発現ベクターのいずれかが、発現されるポリペプチドに対して意図される使用に依存して選択され得る。例えば、大量に必要とされる場合（例えば、抗体の誘導のために）、容易に精製される融合タンパク質の高レベルの発現を指向するベクターが使用され得る。そのようなベクターとしては、以下が挙げられるがそれらに限定されない：多機能性E. coliクローニングベクターおよび発現ベクター（例えば、BLUESCRIPT（Stratagene）、ここでは目的のポリペプチドをコードする配列が、 β -ガラクトシダーゼのアミノ末端Metおよびそれに引続く7残基についての配列とインフレームでベクター内に連結され得、その結果、ハイブリッドタンパク質が産生される）；pINベクター（Van Heeke, G. およびS. M. Schuster (1989) J. Biol. Chem. 264: 5503-5509) など。pGEXベクター（Promega, Madison, Wis.）もまた、グルタチオンSトランスフェラーゼ（GST）との融合タンパク質として外来ポリペプチドを発現させるために使用され得る。一般的に、そのような融合タンパク質は、可溶性であり、そしてグルタチオン-アガロースビーズに吸着させ、次に遊離のグルタチオンの存在下において溶出させることによって、溶解した細胞から容易に精製され得る。そのような系において作製されたタンパク質は、ヘパリン、トロンビン、または第XA因子プロテアーゼ切断部位を含むように設計され得、その結果、クローニングされた目的のポリペプチドが、随意にGST部分から放出され得る。

20

30

【0143】

酵母（*Saccharomyces cerevisiae*）において、構成的プロモーターまたは誘導性プロモーター（例えば、 α 因子、アルコールオキシダーゼおよびPGH）を含む多数のベクターが使用され得る。概説については、Ausubelら、（前出）およびGrantら、（1987）Methods Enzymol. 153: 516-544を参照のこと。

40

【0144】

植物発現ベクターを使用する場合において、ポリペプチドをコードする配列の発現は、任意の多数のプロモーターにより駆動され得る。例えば、ウイルスプロモーター（例えば、CaMVの35Sプロモーターおよび19Sプロモーター）は、単独でか、またはTMVに由来する ω リーダー配列と組み合わせて使用され得る（Takamatsu, N. (1987) EMBO J. 6: 307-311)。あるいは、植物プロモーター（例えば、

50

RUBISCOの小サブユニットまたは熱ショックプロモーター)を使用し得る(Coruzzi, G. ら、(1984) EMBO J. 3:1671-1680; Broglie, R. ら、(1984) Science 224:838-843; および Winter, J ら、(1991) Results Probl. Cell Differ. 17:85-105)。これらの構築物は、直接的DNA形質転換または病原体媒介性トランスフェクションによって植物細胞内に導入され得る。そのような技術は、多数の一般的に入手可能な概説に記載されている(例えば、Hobbs, S. または Murry, L. E., McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York, N. Y.; 191-196 頁を参照のこと)。

10

【0145】

昆虫系はまた、目的のポリペプチドを発現させるために使用され得る。例えば、そのような1つの系において、オートグラフィカリフォルニア核発汗病ウイルス(Autographa californica nuclear polyhedrosis virus) (AcNPV)は、Spodoptera frugiperda細胞またはTrichoplusia larvaeにおける外来遺伝子を発現させるベクターとして使用される。ポリペプチドをコードする配列は、ウイルスの非必須領域内(例えば、ポリヘドリン(polyhedrin)遺伝子)にクローニングされ得て、ポリヘドリンプロモーターの制御下に置かれ得る。ポリペプチドコード配列の首尾良い挿入は、ポリヘドリン遺伝子を不活性化し、そしてコートタンパク質を欠損している組換えウイルスを産生する。次いで、この組換えウイルスを使用して、例えば、目的のポリペプチドが発現され得る、S. frugiperda細胞またはTrichoplusia larvaeに感染させ得る(Engelhardt, E. K. ら、(1994) Proc. Natl. Acad. Sci. 91:3224-3227)。

20

【0146】

哺乳動物宿主細胞において、多数のウイルスベースの発現系が一般的に利用可能である。例えば、アデノウイルスが、発現ベクターとして使用される場合において、目的のポリペプチドをコードする配列は、後期プロモーターおよび3部からなるリーダー配列から構成されるアデノウイルス転写/翻訳複合体内に連結され得る。ウイルスゲノムの非必須E1またはE3領域における挿入を使用して、感染した宿主細胞においてポリペプチドを発現し得る、生存可能ウイルスを入手し得る(Logan, J. および Shenk, T. (1984) Proc. Natl. Acad. Sci. 81:3655-3659)。さらに、転写エンハンサー(例えば、ラウス肉腫ウイルス(RSV)エンハンサー)を使用して、哺乳動物宿主細胞における発現を増加させ得る。

30

【0147】

特定の開始シグナルはまた、目的のポリペプチドをコードする配列のより効率的な翻訳を達成するために使用され得る。そのようなシグナルとしては、ATG開始コドンおよび隣接配列が挙げられる。ポリペプチドをコードする配列、その開始コドンおよび上流配列が適切な発現ベクター内に挿入される場合において、さらなる転写制御シグナルまたは翻訳制御シグナルは必要とされなくとも良い。しかし、コード配列のみ、またはその一部のみが挿入される場合、ATG開始コドンを含む外因性翻訳制御シグナルが提供されるべきである。さらに、開始コドンは、インサート全体の翻訳を確実にするために正確なリーディングフレーム内にあるべきである。外因性翻訳エレメントおよび開始コドンは、種々の起源(天然および合成の両方)に由来し得る。発現の効率は、使用される特定の細胞系に適切なエンハンサー(例えば、文献(Scharf, D. ら、(1994) Results Probl. Cell Differ. 20:125-162)に記載されるエンハンサー)の封入によって増大され得る。

40

【0148】

さらに、宿主細胞株は、挿入された配列の発現を調節するか、または所望の様式において発現されたタンパク質をプロセッシングするその能力について選択され得る。ポリペプチド

50

のそのような改変としては、アセチル化、カルボキシル化、グリコシル化、リン酸化、リビデーション (lipidation) およびアシル化が挙げられるが、これらに限定されない。タンパク質の「プレプロ」形態を切断する翻訳後プロセッシングはまた、正確な挿入、折り畳みおよび／または機能を促進させるために使用され得る。異なる宿主細胞 (例えば、CHO、COS、HeLa、MDCK、HEK293 および WI38) (これらは、そのような翻訳後活性についての特定の細胞機構 (machinery) および特徴的機構 (mechanisms) を有する) は、正確な改変および外来タンパク質のプロセッシングを確実にするように選択され得る。

【0149】

長期間の組換えタンパク質の高収率の産生のために、安定な発現が一般的に好ましい。例えば、目的のポリヌクレオチドを安定に発現する細胞株が、ウイルス複製起点および／または内因性発現エレメントならびに同じかもしくは別個のベクター上の選択マーカー遺伝子を含み得る、発現ベクターを使用して形質転換され得る。そのベクターの導入後、細胞は、それらが選択培地に切換えられる前に、富化 (enriched) 培地において1～2日間の増殖が可能とされ得る。選択マーカーの目的は、選択に対する耐性を与えることであり、そしてその存在は、導入された配列を首尾良く発現する細胞の増殖および回収を可能とする。安定に形質転換された細胞の耐性クローンは、その細胞型に適切な組織培養技術を使用して、増殖され得る。

【0150】

多数の選択系を使用して、形質転換細胞株を回収し得る。これらの選択系としては、それぞれ、tk. sup. 一細胞または aprt. sup. 一細胞において使用され得る、単純疱疹ウイルスチミジンキナーゼ (Wigler, M. ら、(1977) Cell 11: 223-32) およびアデニンホスホリボシルトランスフェラーゼ (Lowy, I. ら、(1990) Cell 22: 817-23) 遺伝子が挙げられるが、これらに限定されない。また、代謝拮抗剤耐性、抗生物質耐性または除草剤耐性は、選択のための基礎として使用され得る；例えば、dhfr (これは、メトトレキセートに対する耐性を与える (Wigler, M. ら、(1980) Proc. Natl. Acad. Sci. 77: 3567-70))；npt (これは、アミノグリコシド、ネオマイシンおよび G-418 に対する耐性を与える (Colbere-Garapin, F. ら、(1981) J. Mol. Biol. 150: 1-14))；ならびに als または pat (これらは、それぞれ、クロルスルフロンおよびホスフィノスリシンアセチルトランスフェラーゼに対する耐性を与える (Murry, 前出))。さらなる選択遺伝子が記載されている。例えば、trpB (これは、細胞がトリプトファンの代わりにインドールを利用することを可能にする) または hisD (これは、細胞がヒスチジンの代わりにヒスチノールを利用することを可能にする) (Hartman, S. C. および R. C. Mulligan (1988) Proc. Natl. Acad. Sci. 85: 8047-51)。アントシアニン、β-グルクロニダーゼおよびその基質の GUS、ならびにルシフェラーゼおよびその基質のルシフェリンのようなマーカーを用いる、可視マーカーの使用の人気の高くなっており、形質転換体を同定するためのみならず、特定のベクター系に起因し得る一過性のタンパク質発現または安定なタンパク質発現の量を定量するためにもまた広範に使用されている (Rhodes, C. A. ら、(1995) Methods Mol. Biol. 55: 121-131)。

【0151】

マーカー遺伝子発現の存在／非存在は、目的の遺伝子もまた存在するということを示唆しているが、その存在および発現は、確認されることを必要とし得る。例えば、ポリペプチドをコードする配列が、マーカー遺伝子配列内に挿入される場合、配列を含む組換え細胞は、マーカー遺伝子機能の非存在により同定され得る。あるいは、マーカー遺伝子は、1つのプロモーターの制御下にポリペプチドコード配列と直列に配置され得る。誘導または選択に応じてのマーカー遺伝子の発現は、通常、その直列遺伝子の発現も同様に示す。

【0152】

10

20

30

40

50

あるいは、所望のポリヌクレオチド配列を含み、かつ発現する宿主細胞は、当業者に公知の種々の手順によって同定され得る。これらの手順としては、以下が挙げられるが、これらに限定されない：核酸またはタンパク質の検出および／または定量化のための、例えば、膜ベースの技術、溶液ベースの技術、またはチップベースの技術を含む、DNA-DNAハイブリダイゼーション技術もしくはDNA-RNAハイブリダイゼーション技術およびタンパク質バイオアッセイ技術またはイムノアッセイ技術。

【0153】

ポリヌクレオチドコード化産物に特異的なポリクローナル抗体もしくはモノクローナル抗体のいずれかを使用して、ポリヌクレオチドコード化産物の発現を検出および測定するための種々のプロトコルは、当該分野において公知である。例としては、酵素結合免疫測定法（ELISA）、ラジオイムノアッセイ（RIA）および蛍光活性化セルソーティング（FACS）が挙げられる。所定のポリペプチド上の2つの非干渉性エピトープに対して反応性であるモノクローナル抗体を利用する2つの部位の、モノクローナルベースのイムノアッセイが、いくつかの適用のために好まれ得るが、競合的結合アッセイもまた、利用され得る。これらのアッセイおよび他のアッセイは、とりわけHampton, R. ら、（1990; Serological Methods, a Laboratory Manual, APS Press, St Paul, Minn.）およびMaddox, D. E. ら、（1983; J. Exp. Med. 158: 1211-1216）に記載される。

10

【0154】

広範に種々の標識技術および結合技術は、当業者に公知であり、そして種々の核酸アッセイおよびアミノ酸アッセイにおいて使用され得る。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またはPCRプローブを作製するための手段としては、オリゴ標識（oligolabeling）、ニックトランスレーション、末端標識または標識されたヌクレオチドを使用するPCR増幅が挙げられる。あるいは、mRNAプローブの産生のために、配列またはその任意の部分が、ベクター内にクローニングされ得る。そのようなベクターは、当該分野において公知であり、市販されており、そして適切なRNAポリメラーゼ（例えば、T7、T3、またはSP6）および標識されたヌクレオチドを添加することにより、RNAプローブをインビトロで合成するために使用され得る。これらの手順は、種々の市販キットを使用して実施され得る。使用され得る適切なレポーター分子または標識としては、放射性核種、酵素、蛍光、化学発光、または色素形成剤ならびに基質、コファクター（補助因子）、インヒビター、磁気粒子などが挙げられる。

20

30

【0155】

目的のポリヌクレオチド配列で形質転換された宿主細胞は、発現および細胞培養物からのタンパク質の回収に適した条件下において培養され得る。組換え細胞により産生されたタンパク質は、使用された配列および／またはベクターに依存して、細胞内に分泌され得るかまたは含まれ得る。当業者によって理解されるように、本発明のポリヌクレオチドを含む発現ベクターは、原核細胞膜または真核細胞膜を介して、コードされたポリペプチドの分泌を指向するシグナル配列を含むように設計され得る。他の組換え構築物を使用して、目的のポリペプチドをコードする配列を、可溶性タンパク質の精製を促進するポリペプチドドメインをコードするヌクレオチド配列に結合させ得る。そのような精製促進ドメインとしては、以下が挙げられるが、これらに限定されない：固定された金属上での精製を可能にする金属キレート化ペプチド（例えば、ヒスチジントリプトファンモジュール）、固定された免疫グロブリン上での精製を可能にするプロテインAドメイン、およびFLAG伸長／アフィニティー精製システム（ImmuneX Corp., Seattle, Wash.）において利用されるドメイン。精製ドメインとコードされるポリペプチドとの間への、切断可能なリンカー配列（例えば、第XA因子またはエンテロキナーゼに対して特異的な配列（Invitrogen, San Diego, Calif.））の封入を使用して、精製を促進させ得る。そのような発現ベクターの1つは、目的のポリペ

40

50

チドを含む融合タンパク質およびチオレドキシンまたはエンテロキナーゼ切断部位の前に6つのヒスチジン残基をコードする核酸の発現を提供する。ヒスチジン残基は、P o r a t h, J. ら、(1992、Prot. Exp. Purif. 3:263-281)に記載されるように、IMIA C (固定化金属イオンアフィニティークロマトグラフィー)上での精製を促進する一方で、エンテロキナーゼ切断部位は、融合タンパク質から所望のポリペプチドを精製するための手段を提供する。融合タンパク質を含むベクターの考察は、K r o l l, D. J. ら、(1993; DNA Cell Boil. 12:441-453)において提供される。

【0156】

組換え産生方法に加えて、本発明のポリペプチドおよびそのフラグメントは、固相技術 (M e r r i f i e l d J. (1963) J. Am. Chem. Soc. 85:2149-2154) を使用する、直接的ペプチド合成によって産生され得る。タンパク質合成は、手動技術を使用してか、または自動化によって実施され得る。自動化合成は、例えば、A p p l i e d B i o s y s t e m s 431A ペプチド合成装置 (P e r k i n E l m e r) を使用して、達成され得る。あるいは、種々のフラグメントは、別々に化学的に合成されて、そして全長分子を産生するために化学的方法を使用して組み合わせられ得る。

【0157】

(抗体組成物、そのフラグメントおよび他の結合因子)

別の局面によると、本発明はさらに、本明細書中で開示される腫瘍ポリペプチドに対して、もしくは一部に対して免疫学的結合を示す結合因子 (例えば、抗体およびその抗原結合フラグメント)、その改変体または誘導体を提供する。抗体、またはその抗原結合フラグメントとは、本発明のポリペプチドに対して、「特異的に結合する」こと、「免疫学的に結合する」ことを言い、そして/または、それが (例えば、E L I S A アッセイにおいて) 検出可能なレベルでそのポリペプチドと反応し、かつ類似の条件下で非関連ポリペプチドと検出可能に反応しない場合、「免疫学的反応性」であると言う。

【0158】

この文脈で使用される場合、免疫学的結合は、一般に、免疫グロブリン分子と、この免疫グロブリンが特異的な抗原との間で起こるタイプの非共有結合的な相互作用をいう。免疫学的結合相互作用の強度または親和性は、相互作用の解離定数 (K_d) の観点から表現され得、ここで、より小さい K_d は、より大きな親和性を示す。選択されたポリペプチドの免疫学的結合特性は、当該分野で周知の方法を使用して、定量され得る。1つのこのような方法は、抗原結合部位/抗原複合体形成および解離の速度を測定する工程を必要とし、ここで、これらの速度は、複合体パートナーの濃度、相互作用の親和性、および両方の方向における速度に等しく影響を与える幾何学的パラメータに依存する。従って、両方の「会合速度定数 (on rate constant)」(K_{on}) および「解離速度定数 (off rate constant)」(K_{off}) の両方が、濃度ならびに会合および解離の実際の速度を計算することによって決定され得る。 K_{off}/K_{on} の比は、親和性に関係しない全てのパラメータの排除を可能にし、そして従って、解離定数 K_d に等しい。一般的には、D a v i e s ら (1990) A n n u a l R e v. B i o c h e m. 59:439-473 を参照のこと。

【0159】

抗体の「抗原結合部位」または「結合部分」は、抗原結合に関係する免疫グロブリン分子の一部分をいう。この抗原結合部位は、重鎖 (「H」) および軽鎖 (「L」) のN末端可変 (「V」) 領域のアミノ酸残基によって形成される。重鎖および軽鎖のV領域内の3つの高度に分岐したストレッチが、「超可変領域」といわれ、これは、「フレームワーク領域」または「FR」として公知のより保存された隣接ストレッチの間に挿入される。従って、用語「FR」は、免疫グロブリンの超可変領域の間で、それに隣接して天然に見出されるアミノ酸配列をいう。抗体分子において、軽鎖の3つの超可変領域および重鎖の3つの超可変領域が、3次元空間において互いに相対的に配置されて、抗原結合表面を形成す

る。この抗原結合表面は、結合した抗原の3次元表面に相補的であり、そして重鎖および軽鎖の各々の3つの超可変領域は、「相補性決定領域」または「CDR」といわれる。

【0160】

結合剤は、本明細書中に提供される代表的なアッセイを使用して、癌（例えば、乳癌）を有する患者と有さない患者との間をさらに区別し得る。例えば、腫瘍タンパク質に結合する抗体または他の結合剤は、好ましくは、この疾患を有する患者の少なくとも約20%、より好ましくは患者の少なくとも約30%において、癌の存在を示すシグナルを生成する。あるいは、または加えて、この抗体は、癌を有さない個体の少なくとも約90%において、この疾患の非存在を示す陰性シグナルを生成する。結合剤が、この要求を満たすか否かを決定するために、癌（標準的な臨床試験によって決定されるような）を有する患者および有さない患者からの生物学的サンプル（例えば、血液、血清、痰、尿および／または腫瘍生検）が、この結合剤に結合するポリペプチドの存在について、本明細書に記載されるようにアッセイされ得る。好ましくは、疾患を有するサンプルおよび有さないサンプルの統計的に有意な数が、アッセイされる。各結合剤は、上記の基準を満足すべきである；しかし、当業者は、結合剤が、組合せて使用されて、感度を改善し得ることを認識する。

10

【0161】

上記の要求を満足する任意の薬剤は、結合剤であり得る。例えば、結合剤は、ペプチド成分を含むか含まないリボソーム、RNA分子またはポリペプチドであり得る。好ましい実施形態において、結合剤は、抗体またはその抗原結合フラグメントである。抗体は、当業者に公知の種々の技術のいずれかによって、調製され得る。例えば、HarlowおよびLane、Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory, 1988を参照のこと。一般に、抗体は、細胞培養技術（本明細書に記載されるようなモノクローナル抗体の生成を含む）によって、または組換え抗体の生成を可能にするために、適切な細菌細胞宿主または哺乳動物細胞宿主への抗体遺伝子のトランスフェクションを介して、生成され得る。1つの技術において、このポリペプチドを含む免疫原が、最初に、広範な種々の哺乳動物（例えば、マウス、ラット、ウサギ、ヒツジまたはヤギ）のいずれかに注射される。この工程において、本発明のポリペプチドは、改変を有さない免疫原として機能し得る。あるいは、特に、比較的短いポリペプチドについて、優れた免疫応答が、このポリペプチドがキャリアタンパク質（例えば、ウシ血清アルブミンまたはキーホールリンペットヘモシアニン）に結合される場合に、惹起され得る。この免疫原は、好ましくは、1回以上のブースト免疫を組み込む所定のスケジュールに従って、動物宿主に注射され、そしてこれらの動物が、定期的に採血される。このポリペプチドに特異的なポリクローナル抗体は、次いで、例えば、適切な固体支持体に結合されたポリペプチドを使用するアフィニティークロマトグラフィーによって、このような抗血清から精製され得る。

20

30

【0162】

目的の抗原性ポリペプチドに特定なモノクローナル抗体が、例えば、KohlerおよびMilstein, Eur. J. Immunol. 6: 511-519, 1976の技術およびその改善物を使用して、調製され得る。簡潔には、これらの方法は、所望の特異性（すなわち、目的のポリペプチドとの反応性）を有する抗体を生成し得る不死細胞株の調製を含む。このような細胞株は、例えば、上記のように免疫された動物から得られる脾臓細胞から生成され得る。次いで、この脾臓細胞が、例えば、骨髓腫細胞融合パートナー、好ましくは、免疫された動物と同系であるものとの融合によって不死化される。種々の融合技術が使用され得る。例えば、脾臓細胞および骨髓腫細胞は、数分間、非イオン性界面活性剤と組合せられ、次いで、ハイブリッド細胞の増殖は支持するが、骨髓腫細胞の増殖は支持しない選択培地に低密度でプレートされ得る。好ましい選択技術は、HAT（ヒポキサンチン、アミノプテリン、チミジン）選択を使用する。十分な時間の後、通常、約1～2週間の後、ハイブリッドのコロニーが観察される。単一コロニーを選択し、そしてそれらの培養上清を、このポリペプチドに対する結合活性について試験する。高い反応性および特異性を有するハイブリドーマが好ましい。

40

50

【0163】

モノクローナル抗体は、増殖ハイブリドーマコロニーの上清から単離され得る。さらに、種々の技術が、収量を増加させるために使用され得、この技術は、例えば、適切な脊椎動物宿主（例えば、マウス）の腹腔へのハイブリドーマ細胞株の注入である。次いで、モノクローナル抗体が、腹水または血液から回収され得る。混入物が、従来の技術（例えば、クロマトグラフィー、ゲル濾過、沈澱、および抽出）によって抗体から除去され得る。本発明のポリペプチドは、例えば、アフィニティークロマトグラフィー工程における精製プロセスにおいて使用され得る。

【0164】

抗体分子の免疫学的結合特性を示し得る抗原結合部位を含む多くの治療的に有効な分子が、当該分野で公知である。このタンパク質分解性酵素パパインは、優先的に、IgG分子を切断して、いくつかのフラグメントを生じさせ、これら（「F(ab)フラグメント」）のうちの2つの各々は、インタクトな抗原結合部位を含む共有結合ヘテロダイマーを含む。酵素ペプシンは、IgG分子を切断して、いくつかのフラグメント（両方の抗原結合部位を含む「F(ab')₂」フラグメントを含む）を提供し得る。「Fv」フラグメントは、IgM、および稀な場合にはIgGまたはIgA免疫グロブリン分子の優先的なタンパク質分解的切断によって生成され得る。しかし、Fvフラグメントは、当該分野で公知の組換え技術を使用して、より一般的に誘導される。Fvフラグメントは、非共有結合的なV_H：：V_Lヘテロダイマーを含み、このヘテロダイマーは、ネイティブな抗体分子の抗原認識および結合能力のほとんどを保持する抗原結合部位を含む。Inbarら（1972）Proc. Nat. Acad. Sci. USA 69：2659-2662；Hochmanら（1976）Biochem 15：2706-2710；およびEhrlichら（1980）Biochem 19：4091-4096。

【0165】

単鎖Fv（「sFv」）ポリペプチドは、共有結合されたV_H：：V_Lヘテロダイマーであり、このヘテロダイマーは、ペプチドをコードするリンカーによって連結されたV_Hをコードする遺伝子およびV_Lをコードする遺伝子を含む遺伝子融合物から発現される。Hustonら（1988）Proc. Nat. Acad. Sci. USA 85（16）：5879-5883。抗体V領域からの天然に凝集する（化学的には分離された）ポリペプチド軽鎖および重鎖を、sFv分子に転換するための化学構造を識別するための多くの方法が記述され、このsFvは、抗原結合部位の構造に実質的に類似する3次元構造の折り畳まれる。例えば、米国特許第5,091,513号および同第5,132,405号（Hustonら）；ならびに米国特許第4,946,778号（Ladnerら）を参照のこと。

【0166】

上記分子の各々は、重鎖および軽鎖CDRセットを含み、各々は、CDRSへの支持を提供し、そして互い対してCDRの空間的關係を規定する重鎖および軽鎖FRセットの間に挿入される。本明細書中で使用される場合、用語「CDRセット」は、重鎖または軽鎖V領域の3つの超可変領域をいう。重鎖または軽鎖のN末端から進んで、これらの領域は、各々、「CDR1」、「CDR2」および「CDR3」と表記される。従って、抗原結合部位は、6つのCDRを含み、この6つのCDRは、重鎖および軽鎖V領域の各々からのCDRセットを含む。単一CDR（例えば、CDR1、CDR2、またはCDR3）を含むポリペプチドは、本明細書中で「分子認識ユニット」といわれる。多くの抗原-抗体複合体の結晶学的分析により、CDRのアミノ酸残基は、結合した抗原との広範な接触を形成することが立証され、ここで、最も広範な接触は、重鎖CDR3とである。従って、分子認識ユニットは、主に、抗原結合部位の特異性の原因である。

【0167】

本明細書中で使用される場合、用語「FRセット」は、重鎖または軽鎖のV領域のCDRセットのCDRを構成する、4つの隣接アミノ酸配列をいう。いくつかのFR残基は、結合抗原と接触し得る；しかし、FRは、主に、V領域を抗原結合部位、特にCDRSに直

接隣接するF R残基に折り畳む原因である。F Rにおいて、特定のアミノ酸残基および特定の構造特徴が、非常に高度に保存される。この点において、全てのV領域配列は、約90アミノ酸残基の内部ジスルフィドループを含む。V領域が、結合部位に折り畳まれる場合、このCDRは、抗原結合表面を形成する突出したループモチーフとして表示される。正確なCDRアミノ酸配列に関わらず、特定の「正準」構造に折り畳まれたCDRループの形状に影響するF Rの保存された構造領域が存在することが一般的に認識される。さらに、特定のF R残基は、抗体重鎖および軽鎖の相互作用を安定化する非共有結合的なドメイン間接触に関係することが知られている。

【0168】

非ヒト免疫グロブリン由来の抗原結合部位を含む、多くの「ヒト化」抗体分子が記載され、この抗体分子には、げっ歯類V領域およびヒト定常領域に融合されたそれらの会合CDR (Winterら(1991) Nature 349:293-299; Lobuglioら(1989) Proc. Nat. Acad. Sci. USA 86:4220-4224; Shawら(1987) J Immunol. 138:4534-4538; およびBrownら(1987) Cancer Res. 47:3577-3583)、適切なヒト抗体定常ドメインとの融合の前に、ヒト支持F Rにグラフトされたげっ歯類CDR (Riechmannら(1988) Nature 332:323-327; Verhoevenら(1988) Science 239:1534-1536; およびJonesら(1986) Nature 321:522-525)、ならびに組換えベニアリングされた(recombinantly veneered)げっ歯類F Rによって支持されるげっ歯類CDR (欧州特許公開番号519, 596(1992年12月23日公開)を含むキメラ抗体が挙げられる。これらの「ヒト化」分子は、ヒトレシピエントにおけるこれらの部分の治療的適用の持続期間および有効性を制限するげっ歯類抗ヒト抗体分子に対する所望されない免疫学的応答を最小化するように設計される。

【0169】

本明細書中で使用される場合、用語「ベニアリングされたF R」および「組換えベニアリングされたF R」は、ネイティブのF Rポリペプチド折り畳み構造の実質的に全てを保持する抗原結合部位を含む異種分子を提供するために、例えば、げっ歯類重鎖または軽鎖のV領域からのF R残基のヒトF R残基での選択的置換をいう。ベニアリング技術は、抗原結合部位のリガンド結合特徴が、主に抗原結合表面内の重鎖および軽鎖のCDRセットの構造および相対的な配置によって決定されるという理解に基づく。Daviesら(1990) Ann. Rev. Biochem. 59:439-473。従って、抗原結合特異性は、CDR構造、互いとのそれらの相互作用、およびV領域ドメインの残りとのそれらの相互作用が注意深く維持されるヒト化抗体のみにおいて保存され得る。ベニアリング技術を使用することによって、外部(例えば、溶媒がアクセス可能な)F R残基(これは、免疫系に容易に遭遇する)は、ヒト残基と選択的に置換されて、弱い免疫原性ベニアリング表面または実質的に非免疫原性のベニアリング表面のいずれかを含むハイブリッド分子を提供する。

【0170】

ベニアリングのプロセスは、Sequences of Proteins of Immunological Interest, 第4版(U. S. Dept. of Health and Human Services, U. S. Government Printing Office, 1987)において、Kabatraによって編集されたヒト抗体可変ドメインについての利用可能な配列データを使用し、Kabataデータベースに更新され、そして他の米国および海外のアクセス可能なデータベース(核酸およびタンパク質の両方)を使用する。V領域アミノ酸の溶媒アクセス可能性は、ヒトおよびマウス抗体フラグメントについての既知の3次元構造から推定され得る。マウス抗原結合部位をベニアリングする際に2つの一般的な工程が存在する。最初に、目的の抗体分子の可変ドメインのF Rが、上記の供給源から得られたヒト可変ドメインの対応するF R配列と比較される。次いで、最も相同的なヒトV領域が、対応するマウスアミノ酸と、残基ごとに

比較される。ヒト対応物とは異なるマウス F R における残基は、当該分野において周知の組換え技術を使用して、ヒト部分に存在する残基によって置換される。残基スイッチングは、少なくとも部分的に露出した（溶媒アクセス可能である）部分を用いて実施されるのみであり、そして V 領域ドメインの三次構造に対する有意な効果を有し得るアミノ酸残基（例えば、プロリン、グリシンおよび荷電したアミノ酸）の置換において、注意がなされる。

【0171】

この様式において、得られた「ベニアリングされた」マウス抗原結合部位は、従って、マウス C D R 残基、C D R に実質的に隣接する残基、埋没したかほとんど埋没した（溶媒のアクセスが不可能）として同定された残基、重鎖ドメインおよび軽鎖ドメインとの間の非共有結合的な（例えば、静電的および疎水的）接触に関係すると考えられる残基、および C D R ループの「正準」三次構造に影響を与えると考えられる F R の保存された構造領域からの残基を保持するように設計される。次いで、これらの設計基準は、ヒト様 F R へのマウス抗原結合部位の軽鎖および重鎖の両方の C D R を組合せる組換えヌクレオチド配列を調製するために使用され、このヒト様 F R は、マウス抗体分子の抗原特異性を示す組換えヒト抗体の発現のために哺乳動物細胞をトランスフェクトするために使用され得る。

【0172】

本発明の別の実施形態において、本発明のモノクローナル抗体は、1つ以上の治療剤に結合され得る。この点に関して、適切な薬剤としては、放射性核種、分化インデューサー、薬物、毒素、およびその誘導体が挙げられる。好ましい放射性核種としては、 ^{90}Y 、 ^{125}I 、 ^{125}I 、 ^{131}I 、 ^{186}Re 、 ^{188}Re 、 ^{211}At および ^{212}Bi が挙げられる。好ましい薬物としては、メトトレキサート、ならびにピリミジンおよびプリンアナログが挙げられる。好ましい分化インデューサーとしては、ホルボールエステルおよび酪酸が挙げられる。好ましい毒素としては、リシン、アブリン、ジフテリア毒素、コレラ毒素、ゲロニン、Pseudomonas 体外毒素、Shigella 毒素、およびアメリカヤマゴボウ抗ウイルスタンパク質が挙げられる。

【0173】

治療剤は、適切なモノクローナル抗体に直接的または間接的に（例えば、リンカー基を介して）に結合（例えば、共有結合）され得る。薬剤と抗体との間の直接的な反応は、各々が他のものと反応し得る置換基を有する場合に可能である。例えば、一方の上の求核基（例えば、アミノ基またはスルフヒドリル基）は、他方のカルボニル含有基（例えば、酸無水物または酸ハロゲン化物）または良好な脱離基（例えば、ハロゲン化物）を含むアルキル基と反応し得る。

【0174】

あるいは、リンカー基を介して治療剤と抗体とを結合させることが所望され得る。リンカー基は、結合の可能性を妨げることを回避するために、抗体を薬剤から隔てるためのスペーサーとして機能し得る。リンカー基はまた、薬剤または抗体上の置換基の化学的反応性を増加させるためにはたらき得、従って結合効率を増大させる。化学的反応性の増大はまた、薬剤または薬剤上の官能基の使用を促進し得る（さもないと、可能ではない）。

【0175】

種々の二官能性または多官能性試薬、ホモ官能性とヘテロ官能性との両方（例えば、Pierce Chemical Co., Rockford, IL のカタログ中に記載されるもの）が、リンカー基として使用され得ることが当業者には明らかである。結合は、例えば、アミノ基、カルボキシル基、スルフヒドリル基、または酸化された炭水化物残基を介してもたらされ得る。このような方法論を記載する多数の参考文献（例えば、Rodwell らに対する米国特許第 4,671,958 号）が存在する。

【0176】

本発明の免疫結合体の抗体部分がないときに治療剤がより強力である場合、細胞中へのインターナリゼーションの間に、またはその際に切断可能なリンカー基を使用することが望ましくあり得る。多数の異なる切断可能なリンカー基が記載されている。これらのリンカ

10

20

30

40

50

一基からの薬剤の細胞内放出についての機構は、ジスルフィド結合の還元（例えば、S p i t t l e r への米国特許第 4, 4 8 9, 7 1 0 号）、感光性結合の照射（例えば、S e n t e r への米国特許第 4, 6 2 5, 0 1 4 号）、誘導体化されたアミノ酸側鎖の加水分解（例えば、K o h n への米国特許第 4, 6 3 8, 0 4 5 号）、血清補体媒介性加水分解（例えば、R o d w e l l への米国特許第 4, 6 7 1, 9 5 8 号）、および酸触媒加水分解（例えば、B l a t t l e r への米国特許第 4, 5 6 9, 7 8 9 号）による切断を含む。

【0177】

1 つより多い薬剤を抗体に結合させることが望ましくあり得る。1 つの実施形態において、薬剤の複数の分子が 1 つの抗体分子に結合される。別の実施形態において、1 つより多い型の薬剤が 1 つの抗体に結合され得る。特定の実施形態に関わらず、1 つより多い薬剤を有する免疫結合体は、種々の方法で調製され得る。例えば、1 つより多い薬剤が、抗体分子に直接的に結合され得るか、または付着のための複数の部位を提供するリンカーが使用され得る。あるいは、キャリアが使用され得る。

10

【0178】

キャリアは、種々の方法（直接的にかまたはリンカー基を介するかのいずれかの共有結合を含む）で薬剤を保有し得る。適切なキャリアとしては、アルブミンのようなタンパク質（例えば、K a t o への米国特許第 4, 5 0 7, 2 3 4 号）、ペプチド、およびアミノデキストランのような多糖類（例えば、S h i h への米国特許第 4, 6 9 9, 7 8 4 号）が挙げられる。キャリアはまた、例えばリポソーム小胞内に、非共有結合によってかまたはカプセル化によって、薬剤を保有し得る（例えば、米国特許第 4, 4 2 9, 0 0 8 号および同第 4, 8 7 3, 0 8 8 号）。放射性核種薬剤に特異的なキャリアは、放射性ハロゲン化低分子およびキレート化合物を含む。例えば、米国特許第 4, 7 3 5, 7 9 2 号は、代表的な放射性ハロゲン化低分子およびそれらの合成を開示する。放射性核種キレートは、金属、または金属酸化物、放射性核種を結合するためのドナー原子として窒素原子および硫黄原子を含むキレート化合物から形成され得る。例えば、D a v i s o n への米国特許第 4, 6 7 3, 5 6 2 号は、代表的なキレート化合物およびそれらの合成を開示する。

20

【0179】

（T 細胞組成物）

別の局面において、本発明は、本明細書中に開示される腫瘍ポリペプチドに特異的な T 細胞、または本明細書中に開示される腫瘍ポリペプチドの改変体もしくは誘導体に特異的な T 細胞を含み得る。このような細胞は、一般的に標準的手順を使用して、インビトロまたはエキソビボで調製され得る。例えば、T 細胞は、市販の細胞分離システム（例えば、I s o l e x ^{T M} システム（N e x e l l T h e r a p e u t i c s, I n c. (I r v i n e, C A ; 米国特許第 5, 2 4 0, 8 5 6 号 ; 米国特許第 5, 2 1 5, 9 2 6 号 ; W O 8 9 / 0 6 2 8 0 ; W O 9 1 / 1 6 1 1 6 および W O 9 2 / 0 7 2 4 3 もまた参照のこと）から入手可能）を使用して、患者の骨髄、末梢血あるいは骨髄または末梢血の画分中から単離され得る。あるいは、T 細胞は、関連または無関連のヒト、非ヒト動物、細胞株または培養物から誘導され得る。

30

40

【0180】

T 細胞は、ポリペプチド、ポリペプチドをコードするポリヌクレオチドおよび／またはそのようなポリペプチドを発現する抗原提示細胞（A P C）を用いて刺激され得る。このような刺激は、目的のポリペプチドに特異的である T 細胞の生成を可能にする条件下および十分な時間、行われる。好ましくは、本発明の腫瘍ポリペプチドまたはポリヌクレオチドは、送達ビヒクル（例えば、ミクロスフェア）中に存在して、特定の T 細胞の生成を容易にする。

【0181】

T 細胞は、この T 細胞が、特異的に増殖するか、サイトカインを分泌するか、または本発明のポリペプチドで被覆されるか、もしくはこのポリペプチドをコードする遺伝子を発現

50

する標的細胞を殺傷する場合に、本発明のポリペプチドに特異的であるとみなされる。T細胞特異性は、種々の標準的技術のいずれかを使用して評価され得る。例えば、クロム放出アッセイまたは増殖アッセイにおいて、ネガティブコントロールと比較して、溶解および／または増殖における2倍を超える増加の刺激指標は、T細胞特異性を示す。このようなアッセイは、例えば、Chenら、Cancer Res. 54:1065-1070, 1994に記載されるように、実行され得る。あるいは、T細胞の増殖の検出は、種々の公知の技術によって達成され得る。例えば、T細胞増殖は、DNA合成の速度の増加を測定することによって検出され得る（例えば、トリチウム化チミジンでT細胞の培養物をパルス標識し、そしてDNAに取り込まれたトリチウム化チミジンの量を測定することによって）。3～7日間の腫瘍ポリペプチド（100 ng/ml～100 μg/ml、好ましくは、200 ng/ml～25 μg/ml）との接触は、代表的に、T細胞の増殖において少なくとも2倍の増加を生じる。2～3時間の上記のような接触は、標準的なサイトカインアッセイを使用して測定されるように（ここで、サイトカイン（例えば、TNFまたはIFN-γ）放出のレベルの2倍の増加が、T細胞の活性化を示す）、T細胞の活性化を生じるべきである（Coliganら、Current Protocols in Immunology, 第1巻、Wiley Interscience (Greene 1998)を参照のこと）。腫瘍ポリペプチド、ポリヌクレオチドまたはポリペプチド発現APCに応答して活性化されたT細胞は、CD4⁺および／またはCD8⁺であり得る。腫瘍ポリペプチド特異的T細胞は、標準的な技術を使用して拡大され得る。好ましい実施形態において、T細胞は、患者、関連するドナーまたは無関連のドナーに由来し、そして刺激および拡大後にその患者に投与される。

10

20

30

40

50

【0182】

治療目的で、腫瘍ポリペプチド、ポリヌクレオチドまたはAPCに応答して増殖するCD4⁺T細胞またはCD8⁺T細胞は、インビトロまたはインビボのいずれかで大量に拡大され得る。このようなT細胞のインビトロでの増殖は、種々の方法で達成され得る。例えば、T細胞は、T細胞増殖因子（例えば、インターロイキン-2）の添加を伴うか、または伴わずに、腫瘍ポリペプチド、またはこのようなポリペプチドの免疫原性部分に対応する短いペプチドに対して再曝露され得、そして／または腫瘍ポリペプチドを合成する刺激細胞に対して再曝露され得る。あるいは、腫瘍ポリペプチドの存在下で増殖する1つ以上のT細胞は、クローニングによって数の上で拡大され得る。細胞をクローニングするための方法は、当該分野で周知であり、そしてこれらとしては、限界希釈が挙げられる。

【0183】

（薬学的組成物）

さらなる実施形態において、本発明は、細胞または動物への投与のための、単独でかまたは治療の1つ以上の他の様相と組合わせてかのいずれかでの、薬学的に受容可能なキャリア中の本明細書中に開示される1つ以上のポリヌクレオチド、ポリペプチド、T細胞および／または抗体組成物の処方物に関する。

【0184】

所望される場合、本明細書中に開示される組成物が、他の薬剤（例えば、他のタンパク質もしくはポリペプチドもしくは種々の薬学的に活性な薬剤など）と組合わせても同様に投与され得ることが、理解される。事実、さらなる薬剤が標的細胞または宿主細胞と接触した際に有意な有害な影響をもたらさないとすれば、さらに含まれ得る他の成分に実質的に制限はない。従って、この組成物は、特定の例において必要とされる種々の他の薬剤と共に送達され得る。このような組成物は、宿主細胞または他の生物学的供給源から精製され得るか、あるいは、本明細書中に記載されるように化学的合成され得る。同様に、このような組成物はさらに、置換または誘導体化されたRNA組成物またはDNA組成物を含み得る。

【0185】

従って、本発明の別の局面において、生理学的に受容可能なキャリアと組合わせた、本明細書中に記載される1つ以上のポリヌクレオチド、ポリペプチド、抗体、および／または

T細胞組成物を含む薬学的組成物が、提供される。特定の好ましい実施形態において、本発明の薬学的組成物は、予防ワクチン適用および治療ワクチン適用における使用のための、本発明の免疫原性ポリヌクレオチド組成物および／または免疫原性ポリペプチド組成物を含む。ワクチン調製物は、一般的に、例えば、M. F. PowellおよびM. J. Newman編、「Vaccine Design (the subunit and adjuvant approach)」、Plenum Press (NY, 1995)に記載される。一般的にこのような組成物は、1つ以上の免疫刺激剤と組合わせた、本発明のポリヌクレオチド組成物および／またはポリペプチド組成物の1つ以上を含む。

【0186】

本明細書中に記載される任意の薬学的組成物が、本発明のポリヌクレオチドおよびポリペプチドの薬学的に受容可能な塩を含み得ることは、明らかである。このような塩は、例えば、薬学的に受容可能な非毒性の塩基（有機塩基（例えば、一級アミン、二級アミンおよび三級アミンならびに塩基性アミノ酸の塩）および無機塩基（例えば、ナトリウム塩、カリウム塩、リチウム塩、アンモニウム塩、カルシウム塩およびマグネシウム塩）を含む）から調製され得る。

【0187】

別の実施形態において、本発明の例示的な免疫原性組成物（例えば、ワクチン組成物）は、上記のような1つ以上のポリペプチドをコードするDNAを含み、その結果このポリペプチドは、インサイチュで生成される。上記のように、ポリヌクレオチドは、当業者に公知の種々の送達系の内のいずれかで投与され得る。実際、多数の遺伝子送達技術（例えば、Rolland, Crit. Rev. Therap. Drug Carrier Systems 15:143-198, 1998およびこの中に引用される参考文献に記載される遺伝子送達技術）が、当該分野で周知である。当然、適切なポリヌクレオチド発現系は、患者における発現のための必要な調節性のDNA調節配列（例えば、適切なプロモーターおよび終結シグナル）を含む。あるいは、細菌送達系は、その細胞表面上にポリペプチドの免疫原性部分を発現するか、またはエピトープなどを分泌する細菌（例えば、Bacillus-Calmette-Guerrin）の投与を含み得る。

【0188】

従って、特定の実施形態において、本明細書中に記載される免疫原性ポリペプチドをコードするポリヌクレオチドは、多数の公知のウイルスに基づく系のうちのいずれかを用いて、発現のための適切な哺乳動物宿主細胞に導入される。1つの例示的な実施形態において、レトロウイルスが、遺伝子送達系のための簡便かつ有効な基盤を提供する。本発明のポリペプチドをコードする選択されたヌクレオチド配列は、当該分野で公知の技術を用いて、ベクターに挿入され得、そしてレトロウイルス粒子にパッケージングされ得る。次いで、組換えウイルスが、単離され、そして被験体に送達され得る。多数の例示的なレトロウイルス系が、記載されている（例えば、米国特許第5,219,740号；MillerおよびRosman (1989) BioTechniques 7:980-990；Miller, A. D. (1990) Human Gene Therapy 1:5-14；Scarpaら (1991) Virology 180:849-852；Burnsら (1993) Proc. Natl. Acad. Sci. USA 90:8033-8037；ならびにBoris-LawrieおよびTemin (1993) Cur. Opin. Genet. Develop. 3:102-109）。

【0189】

さらに、多数の例示的なアデノウイルスに基づく系もまた、記載されている。宿主ゲノムに組込むレトロウイルスとは異なり、アデノウイルスは、染色体外に残り、このようにして、挿入変異誘発に関連する危険性を最小化する（Haj-AhmadおよびGraham (1986) J. Virol. 57:267-274；Bettrら (1993) J. Virol. 67:5911-5921；Mitterederら (1994) Human Gene Therapy 5:717-729；Sethら (1994) J. Virol. 68:933-940；Barrら (1994) Gene Therapy 1:

51-58; Berkner, K. L. (1988) *BioTechniques* 6: 616-629; ならびに Richら (1993) *Human Gene Therapy* 4: 461-476)。

【0190】

種々のアデノ随伴ウイルス (AAV) ベクター系もまた、ポリヌクレオチド送達に関して開発されている。AAVベクターは、当該分野で周知の技術を用いて容易に構築され得る。例えば、米国特許第5, 173, 414号および同第5, 139, 941号; 国際公開番号WO 92/01070およびWO 93/03769; Lebkowskiら (1988) *Molec. Cell. Biol.* 8: 3988-3996; Vincentら (1990) *Vaccines* 90 (Cold Spring Harbor Laboratory Press); Carter, B. J. (1992) *Current Opinion in Biotechnology* 3: 533-539; Muzyczka, N. (1992) *Current Topics in Microbiol. and Immunol.* 158: 97-129; Kotin, R. M. (1994) *Human Gene Therapy* 5: 793-801; ShellingおよびSmith (1994) *Gene Therapy* 1: 165-169; ならびに Zhouら (1994) *J. Exp. Med.* 179: 1867-1875を参照のこと。

【0191】

遺伝子移入による、本発明のポリペプチドをコードするポリヌクレオチドを送達するのに有用なさらなるウイルスベクターとしては、ポックスウイルスのファミリー (例えば、ワクシニアウイルスおよび鳥類ポックスウイルス) から誘導されるものが挙げられる。例として、新規分子を発現するワクシニアウイルスの組換え体は、以下のように構築され得る。ポリペプチドをコードするDNAを最初に、ワクシニアプロモーターおよび隣接ワクシニアDNA配列 (例えば、チミジンキナーゼ (TK) をコードする配列) に隣接するように適切なベクターに挿入する。次いで、このベクターを使用して、ワクシニアで同時に感染させる細胞にトランスフェクトする。相同組換えによって、ワクシニアプロモーターおよび目的のポリペプチドをコードする遺伝子を、ウイルスゲノムに挿入する。生じるTK^{sup} (-) 組換え体を、5-ブロモデオキシウリジンの存在下でこの細胞を培養し、そして5-ブロモデオキシウリジンに耐性のウイルスプラークをピックアップすることによって、選択し得る。

【0192】

ワクシニアに基づく感染/トランスフェクション系は、生物の宿主細胞における、本明細書中に記載される1つ以上のポリペプチドの誘導可能な一過性発現または同時発現を提供するために、好都合に使用され得る。この特定の系において、細胞を最初に、インビトロで、バクテリオファージT7 RNAポリメラーゼをコードするワクシニアウイルスの組換え体で感染させる。このポリメラーゼは、それがT7プロモーターを保有する鋳型のみを転写する点で、鋭敏な特異性を示す。感染後、細胞を、T7プロモーターによって駆動される目的のポリヌクレオチドでトランスフェクトさせる。ワクシニアウイルスの組換え体から細胞質中で発現されるポリメラーゼは、トランスフェクトされたDNAをRNAに転写し、このRNAは次いで、宿主翻訳機構によってポリペプチドに翻訳される。この方法は、大量のRNAおよびその翻訳産物の、高レベルで一過性の細胞質産生を提供する。例えば、Elroy-SteinおよびMoss, *Proc. Natl. Acad. Sci. USA* (1990) 87: 6743-6747; Fuerstら、*Proc. Natl. Acad. Sci. USA* (1986) 83: 8122-8126を参照のこと。

【0193】

あるいは、アビポックスウイルス (avipoxvirus) (例えば、鶏痘ウイルスおよびカナリア痘ウイルス) もまた使用されて、目的のコード配列を送達し得る。哺乳動物病原体由来の免疫原を発現する組換えアビポックスウイルスは、非鳥類種に投与された場合に、防御免疫を与えることが公知である。アビポックスベクターの使用は、ヒトおよび他の哺乳動物種において特に望ましい。なぜなら、アビポックス属のメンバーは、感受性

の鳥類種においてのみ生産的に複製し得、故に、哺乳動物細胞において感染性ではないからである。組換えアビボックスウイルスを産生するための方法は、当該分野で公知であり、そしてワクシニアウイルスの産生に関して上記に記載されるように、遺伝子組換えを使用する。例えば、WO 91/12882；WO 89/03429；およびWO 92/03545を参照のこと。

【0194】

任意の多数のアルファウイルスベクターもまた、本発明のポリヌクレオチド組成物の送達のために使用され得、これらとしては、米国特許第5,843,723号；同第6,015,686号；同第6,008,035号および同第6,015,694号に記載されるベクターが挙げられる。ベネズエラウマ脳脊髄炎（VEE）に基づく特定のベクターもまた、使用され得、この例示的な例は、米国特許第5,505,947号および同第5,643,576号に見出され得る。

10

【0195】

さらに、分子結合体化ベクター（例えば、Michaelら、J. Biol. Chem. (1993) 268:6866~6869およびWagnerら、Proc. Natl. Acad. Sci. USA (1992) 89:6099~6103に記載のアデノウイルススキメラベクター）はまた、本発明における遺伝子送達に用いられ得る。

【0196】

これらおよび他の公知のウイルスベースの送達系に対するさらなる例示的情報は、例えば、以下に見出され得る：Fisher-Hochら、Proc. Natl. Acad. Sci. USA 86:317~321、1989；Flexnerら、Ann. N. Y. Acad. Sci. 569:86~103、1989；Flexnerら、Vaccine 8:17~21、1990；米国特許第4,603,112号、同第4,769,330号および同第5,017,487号；WO89/01973；米国特許第4,777,127号；GB 2,200,651:EP 0,345,242号；WO 91/02805；Berkner、Biotechniques 6:616~627、1988；Rosenfeldら、Science 252:431~434、1991；Kollsら、Proc. Natl. Acad. Sci. USA 91:215~219、1994；Kass-Eislerら、Proc. Natl. Acad. Sci. USA 90:11498~11502、1993；Guzmanら、Circulation 88:2838~2848、1993；ならびにGuzmanら、Cir. Res. 73:1202~1207、1993。

20

30

【0197】

特定の実施形態において、ポリヌクレオチドを標的細胞のゲノムに組み込み得る。この組み込みは、相同組み換え（遺伝子置換）を介して特定の位置および方向であり得るか、または無作為な非特異的位置に組み込まれ得る（遺伝子増大）。なおさらなる実施形態において、ポリヌクレオチドはDNAの別のエピソードセグメントとして細胞において安定に維持され得る。このようなポリヌクレオチドセグメントすなわち「エピソード」は、宿主細胞の周期に独立してまたは同調して維持および複製を可能にするのに十分な配列をコードする。発現構築物が細胞に送達され、そしてこの細胞中にこのポリヌクレオチドが残る様式は、使用される発現構築物のタイプに依存する。

40

【0198】

本発明の別の実施形態において、例えば、Ulmerら、Science 259:1745~1749、1993に記載され、そしてCohen、Science 259:1691~1692、1993によって概説されるように、ポリヌクレオチドは、「裸の」DNAとして投与／送達される。裸のDNAの取り込みは、生分解性ビーズ（これは、細胞に効率的に輸送される）上にそのDNAをコーティングすることによって増加され得る。

【0199】

なお別の実施形態において、本発明の組成物は、微粒子銃アプローチ（その多くが記載さ

50

れている)を介して送達され得る。1つの代表的な例において、ガス駆動粒子加速(gas-driven acceleration)は、Powderject Pharmaceuticals PLC(Oxford, UK)およびPoderject Vaccines Inc.(Madison, WI)によって製造されたデバイスのようなデバイスで達成され得る。そのいくつかの例は、米国特許第5,846,796号;同第6,010,478号;同第5,865,796号;同第5,584,807号;およびEP特許番号第0500799号に記載されている。このアプローチは、注射針のない(無注射針)(ニードルフリー)送達アプローチを提供する。ここでは、顕微鏡的な粒子(例えば、ポリヌクレオチド粒子、またはポリペプチド粒子)の乾燥粉末処方物を、手持ちデバイスによって生成されたヘリウムガスジェット内で高速に加速し、目的の標的組織内へ粒子を噴射する。 10

【0200】

関連の実施形態において、本発明の組成物のガス駆動注射針なし注入に有用であり得る他のデバイスおよび方法は、Bioject, Inc.(Portland, OR)によって提供されるものが挙げられ、そのいくつかの例は、米国特許第4,790,824号;同第5,064,413号;同第5,312,335号;同第5,383,851号;同第5,399,163号;同第5,520,639号;および同第5,993,412号に記載されている。

【0201】

別の実施形態に従って、本明細書に記載される薬学的組成物は、本発明の免疫原性ポリヌクレオチド、ポリペプチド、抗体、T細胞および/またはAPC組成物に加えて、1つ以上の免疫賦活剤(免疫刺激因子)を含む。免疫賦活薬とは、本質的に、外因性抗原に対する免疫応答(抗原および/または細胞媒介)を増大または増強する任意の物質をいう。免疫賦活薬の1つの好ましい型は、アジュバントを含む。多くのアジュバントは、迅速な異化作用から抗原を保護するように設計された物質(例えば、水酸化アルミニウムまたは鉱油)および免疫応答の刺激物質(例えば、脂質A、Bordetella pertussisまたはMycobacterium tuberculosis誘導化タンパク質)を含む。特定のアジュバント、例えば、フロイント不完全アジュバントおよびフロイント完全アジュバント(Difco Laboratories, Detroit, MI);Merck Adjuvant 65(Merck and Company, Inc., Rahway, NJ);AS-2(SmithKline Beecham, Philadelphia, PA);アルミニウム塩(例えば、水酸化アルミニウムゲル(ミョウバン)またはリン酸アルミニウム);カルシウム、鉄または亜鉛の塩;アシル化チロシンの不溶性懸濁液;アセチル化糖;カチオンの誘導化多糖またはアニオンの誘導化多糖;ポリフォスファゼン(polyphosphazene);生分解性ミクロスフェア:モノホスホリル脂質Aおよびモノホスホリルクイル(quil)Aが、市販されている。サイトカイン(例えば、GM-CSF、インターロイキン-2、インターロイキン-7、インターロイキン-12および成長因子のような他のもの)はまた、アジュバントとして使用され得る。 20 30

【0202】

本発明の特定の実施形態において、アジュバント組成物は、Th1型の免疫応答を優勢に誘導するものが好ましい。高レベルのTh1型サイトカイン(例えば、IFN- γ 、TNF α 、IL-2およびIL-12)は、投与される抗原に対する細胞媒介性応答の誘導を支持する傾向がある。対照的に、高レベルのTh2型サイトカイン(例えば、IL-4、IL-5、IL-6およびIL-10)は、体液性免疫応答の誘導を支持する傾向がある。本明細書中に提供されるようなワクチンの適用の後、患者は、Th1型応答およびTh2型応答を含む免疫応答を支持する。応答が優勢にTh1型である好ましい実施形態において、Th1型サイトカインのレベルは、Th2型サイトカインのレベルより高い程度まで増加する。これらのサイトカインのレベルは、標準的アッセイを使用して容易に評価され得る。サイトカインのファミリーの概説については、MosmannおよびCoffm 40 50

an、Ann. Rev. Immunol. 7:145~173、1989を参照のこと。

【0203】

Th1型優勢の応答を誘発するための特定の好ましいアジュバンドは、例えば、モノホスホリルリピドA、好ましくは3-de-O-アシル化モノホスホリルリピドAとアルミニウム塩との組み合わせを含む。MPL（登録商標）アジュバンドは、Corixa Corporation（Seattle, WA；例えば、米国特許第4,436,727号；同第4,877,611号；同第4,866,034号および同第4,912,094号を参照のこと）から入手可能である。CpG含有オリゴヌクレオチド（ここで、CpGジヌクレオチドはメチル化されていない）はまた、Th1優勢の応答を誘導する。このようなオリゴヌクレオチドは周知であり、そして例えばWO96/02555、WO99/33488ならびに米国特許第6,008,200号および同第5,856,462号に記載される。免疫賦活薬DNA配列がまた、例えば、Satoら、Science 273:352, 1996によって記載される。別の好ましいアジュバンドは、サポニン（例えば、Quil A）、またはその誘導體（QS21およびQS7（Aquila Biopharmaceuticals Inc., Framingham, MA）を含む）；Escin；Digitonin；またはGypsophilaまたはChenopodium quinoaサポニンを含む。他の好ましい処方物としては、本発明のアジュバントの組み合わせ、例えば、QS21、QS7、Quil A、 β -エスシン、またはジギトニンを含む以下の軍のうち少なくとも2つの組み合わせ、において1つより多いサポニンを含む。

10

20

【0204】

あるいは、このサポニン処方物は、キトサン、または他のポリカチオン性ポリマーからなるワクチンビヒクル、ポリラクチドおよびポリラクチド-コ-グリコチド粒子、ポリ-N-アセチルグルコサミンベースのポリマーマトリックス、ポリサッカライドまたは化学的に改変されたポリサッカライドからなる粒子、リポソームおよび脂質ベースの粒子、グリセロールモノエステルからなる粒子、などと組み合わせられ得る。サポニンはまた、コレステロールの存在下で処方されて、リポソームまたはISCOMのような粒子構造を形成し得る。さらに、サポニンは、非粒子の溶液もしくは懸濁液で、または小数層リポソームもしくはISCOMのような粒子状構造で、ポリオキシエチレンエーテルまたはエステルと一緒に処方され得る。サポニンはまた、Carbopol（登録商標）のような賦形剤と一緒に処方されて、粘度を増大され得るか、またはラクトースのような粉末賦形剤とともに乾燥粉末形態に処方され得る。

30

【0205】

1つの好ましい実施形態において、アジュバントシステムは、モノホスホリルリピドAとサポニン誘導體との組み合わせ（例えば、WO94/00153に記載されるような、QS21と3D-MPL（登録商標）アジュバントとの組み合わせ、またはWO96/33739に記載されるような、QS21がコレステロールでクエンチ（quench）される、あまり反応発生的（reactogenic）でない組成物）を含む。他の好ましい処方物は、水中油型エマルジョンおよびトコフェロールを含む。水中油型エマルジョン中にQS21、3D-MPL（登録商標）アジュバントおよびトコフェロールを含む、別の特に強力なアジュバンド処方物は、WO95/17210に記載されている。

40

【0206】

別の強化されたアジュバント系は、CpG含有オリゴヌクレオチドとサポニン誘導體の組み合わせを含み、特にCpGとQS21の組み合わせがWO00/09159に開示されている。好ましくは、この処方物は、さらに、水中油型エマルジョンおよびトコフェロールを含む。

【0207】

本発明の薬学的組成物における使用のためのさらなる代表的アジュバンドとしては、Montanide ISA 720（Seppic, France）、SAF（Chiron, California, United States）、ISCOMS（CSL）、

50

MF-59 (Chiron)、アジュバンドのSBASシリーズ (例えば、SBAS-2 またはSBAS-4 (SmithKline Beecham、Rixensart、Belgiumから入手可能)、Detox (Enhanzyn (登録商標)) (Corixa, Hamilton, MT)、RC-529 (Corixa, Hamilton, MT) および他のアミノアルキルグルコサミニド4-ホスフェート (AGP) (例えば、係続中の米国出願登録番号08/853, 826および09/074, 720 (これらの開示は、その全体が本明細書中に参考として援用される) に記載されるアジュバンド)、ならびにWO 99/52549 A1に記載のポリオキシエチレンエーテルアジュバンドのようなアジュバンドが挙げられる。

【0208】

他の好ましいアジュバンドは、一般式 (I) : $\text{HO}(\text{CH}_2\text{CH}_2\text{O})_n\text{--A--R}$ のアジュバンド分子を含み、

ここで、 n は、1~50であり、 A は、結合または --C(O)-- であり、 R は、 $C_{1\sim50}$ アルキルまたはフェニル $C_{1\sim50}$ アルキルである。

【0209】

本発明の1つの実施形態は、一般式 (I) のポリオキシエチレンエーテルを含むワクチン処方物からなり、ここで、 n は、1と50との間、好ましくは4~24、最も好ましくは9であり、この R 成分は、 $C_{1\sim50}$ 、好ましくは $C_4\sim C_{20}$ アルキル、そして最も好ましくは C_{12} アルキルであり、そして A は、結合である。ポリオキシエチレンエーテルの濃度は、0.1~20%の範囲、好ましくは0.1~10%、そして最も好ましくは0.1~1%の範囲にあるべきである。好ましいポリオキシエチレンエーテルは、以下の群から選択される：ポリオキシエチレン-9-ラウロイルエーテル、ポリオキシエチレン-9-ステアシル (stearyl) エーテル、ポリオキシエチレン-8-ステアシル (stearyl) エーテル、ポリオキシエチレン-4-ラウロイルエーテル、ポリオキシエチレン-35-ラウロイルエーテルおよびポリオキシエチレン-23-ラウロイルエーテル。ポリオキシエチレンラウロイルエーテルのようなポリオキシエチレンエーテルは、Merk Index (第12版、7717項目) に記載される。これらのアジュバンド分子は、WO 99/52549に記載される。

【0210】

上記の一般式 (I) に従うポリオキシエチレンエーテルは、所望の場合、別のアジュバンドと組み合わせられ得る。例えば、好ましいアジュバンドの組み合わせは、好ましくは、係属中の英国特許出願第GB 9820956.2号に記載されるようなCPGとの組み合わせである。

【0211】

本発明の別の実施形態に従って、本明細書に記載される免疫原性組成物は、抗原提示細胞 (APC) (例えば、樹状細胞、マクロファージ、B細胞、単球、および有効なAPCであるように操作され得る他の細胞) を介して宿主に送達される。このような細胞は、抗原を提示する能力を増大するように、T細胞応答の活性化および/または維持を改良するように、それ自体で抗腫瘍効果を有するように、そして/あるいは受け手と免疫学的に適合性 (すなわち、一致するHLAハプロタイプ) であるように遺伝学的に改変されてもよいが、改変される必要はない。APCは、一般に、種々の生物学的な流体および器官 (腫瘍および腫瘍周辺組織を含む) のいずれかから単離されてもよいし、そして自己細胞、同種異系細胞、同系細胞、または異種細胞であってもよい。

【0212】

本発明の特定の好ましい実施形態は、抗原提示細胞として、樹状細胞またはその前駆細胞を使用する。樹状細胞は、高度に強力なAPCであり (BanchereauおよびSteinman、Nature 392:245-251、1998)、そして予防的または治療的な抗腫瘍免疫性を誘発するための生理学的アジュバンドとして有効であることが示されてきた (TimmermanおよびLevy、Ann. Rev. Med. 50:507-529、1999を参照のこと)。一般に、樹状細胞は、それらの代表的な形状 (

10

20

30

40

50

インサイチュでは星状、インビトロでは目に見える顕著な細胞質プロセス（樹枝状結晶）を有する）、高い効率で抗原を取り込み、プロセッシングし、そして提示するそれらの能力、およびナイーブな（naive）T細胞応答を活性化するそれらの能力に基づいて同定され得る。もちろん樹状細胞は、インビボまたはエキソビボで樹状細胞上に通常見出されない特定の細胞表面レセプターまたはリガンドを発現するように操作され得、このような改変樹状細胞は本発明によって意図される。樹状細胞の代替として、分泌小胞抗原装荷樹状細胞（secreted vesicles antigen-loaded dendritic cell）（エキソソーム（exosome）と呼ばれる）がワクチン内で使用され得る（Zitvogelら、Nature Med. 4: 594-600, 1998を参照のこと）。

10

【0213】

樹状細胞および前駆細胞は、末梢血、骨髓、腫瘍浸潤細胞、腫瘍周辺組織浸潤細胞、リンパ節、脾臓、皮膚、臍帯血、または他の適切な組織もしくは流体から得られ得る。例えば、樹状細胞は、末梢血から収集された単球の培養物に、GM-CSF、IL-4、IL-13および／またはTNF α のようなサイトカインの組み合わせを添加することによってエキソビボで分化され得る。あるいは、末梢血、臍帯血または骨髓から収集されたCD34陽性細胞は、培養培地にGM-CSF、IL-3、TNF α 、CD40リガンド、LPS、flt3リガンドおよび／または樹状細胞の分化、成熟、および増殖を誘導する他の化合物の組み合わせを添加することによって、樹状細胞に分化され得る。

【0214】

樹状細胞は、「未熟」細胞および「成熟」細胞として都合良く分類され、このことは、単純な方法で2つの充分に特徴付けられた表現型の間を区別するのを可能にする。しかしこの学名は、あらゆる可能な分化の中間段階を排除すると解釈されるべきではない。未熟な樹状細胞は、抗原の取り込みおよびプロセッシングの高い能力を有するAPCとして特徴付けられ、この能力は、Fc γ レセプターおよびマンノースレセプターの高度な発現と相関する。成熟表現型は、代表的に、クラスIおよびクラスII MHC、接着分子（例えば、CD54およびCD11）ならびに同時刺激性分子（例えば、CD40、CD80、CD86および4-1BB）のようなT細胞活性化を担う細胞表面分子の高度な発現ではなく、これらのマーカーのより低い発現によって特徴付けられる。

20

【0215】

APCは、一般に、本発明のポリヌクレオチド（またはその部分もしくは他の改変体）を用いてトランスフェクトされ得、その結果、このコードされたポリペプチドまたはその免疫原性部分が細胞表面上に発現される。このようなトランスフェクションはエキソビボで生じ得、次いでこのようなトランスフェクトされた細胞を含む薬学的組成物は、本明細書中に記載されるように、治療目的のために使用され得る。あるいは、樹状細胞または他の抗原提示細胞を標的とする遺伝子送達ビヒクルが、患者に投与され得、インビボで起こるトランスフェクションを生じる。樹状細胞のインビボおよびエキソビボでのトランスフェクションは、例えば、WO 97/24447に記載される方法、またはMahviら、Immunology and Cell Biology 75: 456-460、1997によって記載される遺伝子銃アプローチのような当該分野で公知の任意の方法を使用して一般に実施され得る。樹状細胞の抗原ローディングは、樹状細胞または前駆細胞を、腫瘍ポリペプチド、DNA（裸のもしくはプラスミドベクター中の）またはRNA；あるいは抗原発現性組換え細菌またはウイルス（例えば、牛痘、鶏痘、アデノウイルスまたはレンチウイルスのベクター）とインキュベートすることによって達成され得る。ローディングの前に、ポリペプチドは、T細胞補助を提供する免疫学的パートナー（例えば、キャリア分子）に共有結合され得る。あるいは、樹状細胞は、別々にまたはポリペプチドの存在下で、結合していない免疫学的パートナーでパルスされ得る。

30

【0216】

当業者に公知の任意の適切なキャリアが、本発明の薬学的組成物において使用され得るが、このタイプのキャリアは、代表的に投与の様式に依存して変化する。本発明の組成物は

40

50

、投与の適切な様式、例えば、局所投与、経口投与、経鼻投与、粘膜投与、静脈投与、頭蓋内投与、腹腔内投与、皮下投与、および筋肉内投与のために処方され得る。

【0217】

このような処方物内での使用のためのキャリアは、生体適合性であり、そしてまた生分解性であり得る。特定の実施形態において、好ましくは、この処方物は比較的一定レベルの活性成分の放出を提供する。しかし、他の実施形態において、投与直後のより迅速な放出速度が所望され得る。このような組成物の処方には十分に、公知の技術を使用する当業者のレベル内である。この点に関して有用な例示的なキャリアとしては、ポリ（ラクチド-ε-ε-グルコリド）、ポリアクリレート、ラテックス、デンプン、セルロース、デキストランなどのマイクロマプセルが挙げられる。他の例示的な徐放性キャリアとしては、非液性（non-liquid）親水性コア（例えば、架橋ポリサッカリドまたはオリゴサッカリド）を含む超分子バイオベクター、ならびに必要に応じて、両親媒性化合物を含む外部層（例えば、リン脂質）（例えば、米国特許第5,151,254号およびPCT出願WO94/20078、WO/94/23701およびWO96/06638を参照のこと）を含む超分子バイオベクターが挙げられる。徐放性処方物内に含まれる活性な化合物の量は、移植の部位、放出の速度および予期される持続期間、ならびに処置または予防されるべき状態の性質に依存する。

10

【0218】

別の例示的な実施形態において、生分解性ミクロスフェア（例えば、ポリ乳酸ポリグリコレート）は、本発明の組成物のためのキャリアとして使用される。適切な生分解性ミクロスフェアは、例えば、米国特許第4,897,268号；同第5,075,109号；同第5,928,647号；同第5,811,128号；同第5,820,883号；同第5,853,763号；同第5,814,344号；同第5,407,609号および同第5,942,252号に開示される。改変B型肝炎コアタンパク質キャリア系（例えば、WO/99 40934およびそこで引用される参考文献に記載されるような系）もまた、多くの用途に有用である。別の例示的なキャリア送達システムは、キャリア含有粒子-タンパク質複合体（例えば、米国特許第5,928,647号に記載される複合体）を使用し、これは、宿主においてクラスI拘束細胞傷害性Tリンパ球応答を誘導し得る。

20

【0219】

本発明の薬学的組成物はしばしば、1つ以上の、緩衝剤（例えば、中性の緩衝化生理食塩水またはリン酸緩衝化生理食塩水）；糖質（例えば、グルコース、マンノース、スクロースまたはデキストラン）；マンニトール；タンパク質；ポリペプチドまたはアミノ酸（例えば、グリシン）；抗酸化剤；静菌剤；キレート剤（例えば、EDTAまたはグルタチオン）；アジュバント（例えば、水酸化アルミニウム）；処方物をレシピエントの血液に対して等張性、低張性または弱く高張性にする溶質；懸濁剤；濃化剤および／または保存剤を含み得る。あるいは、本発明の組成物は、凍結乾燥物として処方され得る。

30

【0220】

本明細書中に記載される薬学的組成物は、単回用量容器または多用量容器（例えば、密閉アンプルまたはバイアル）中に存在し得る。このような容器は、代表的に、使用まで処方物の無菌性および安定性を維持するために、密封される。一般に、処方物は、油性ビヒクルまたは水性ビヒクル中の、懸濁液、溶液またはエマルジョンとして保存され得る。あるいは、薬学的組成物は、使用の直前に無菌水性キャリアの添加のみを必要とする、凍結乾燥状態で保存され得る。

40

【0221】

様々な処置レジメンにおいて本明細書中で記載される特定の組成物を使用するための適切な投薬レジメンおよび処置レジメン（例えば、経口、非経口、静脈内、鼻腔内、および筋肉内の投与および処方物を含む）の開発は、当該分野で周知であり、これらのいくつかは、一般的な例示目的のために以下で簡単に考察される。

【0222】

特定の用途において、本明細書中で開示される薬学的組成物は、経口投与によって動物に

50

送達され得る。このように、これらの組成物は、不活性希釈剤と共にかまたは同化可能食用キャリアと共に処方され得るか、あるいはこれらの組成物は、硬質または軟質殻ゼラチンカプセルに入れられ得るか、あるいはこれらの組成物は錠剤に圧縮され得るか、あるいはこれらの組成物は食餌の食品に直接組み込まれ得る。

【0223】

活性化合物は、さらに賦形剤に組み込まれ得、そして経口摂取錠剤、経頬粘膜錠剤、トローチ、カプセル剤、エリキシル、懸濁液、シロップ、カシェ剤などの形態で使用される（例えば、Mathiowitzら、Nature 1997 Mar 27; 386 (6623): 410-4; Hwangら、Crit Rev Ther Drug Carrier Syst 1998; 15 (3): 243-84; 米国特許第5,641,515号; 米国特許第5,580,579号および米国特許第5,792,451号を参照のこと）。錠剤、トローチ、丸剤、カプセル剤などはまた、任意の種々のさらなる成分を含み得、例えば、結合剤（例えば、トラガカントゴム、アラビアゴム、コーンスターチまたはゼラチン）; 賦形剤（例えば、リン酸二カルシウム）; 崩壊剤（例えば、コーンスターチ、ポテトデンプン、アルギン酸など）; 潤沢剤（例えば、ステアリン酸マグネシウム）; および甘味料（例えば、スクロース、ラクトースまたはサッカリン）または香料（例えば、ペパーミント、ウインターグリーン油、またはチェリー香料）が添加され得る。投薬量単位形態がカプセル剤である場合、これは上記のタイプの材料に加えて液体キャリアを含み得る。様々な他の材料が、コーティングとして存在し得るか、またはそうでなければ物理的形態の投薬単位を改変し得る。例えば、錠剤、丸剤またはカプセル剤は、セラック、糖またはその両方でコーティングされ得る。もちろん、任意の投薬単位形態を調製する際に使用される任意の材料は、薬学的に純粋でありかつ用いられる量で実質的に非毒性でなければならない。さらに、活性化合物が、持続性放出調製物および処方物に組み込まれ得る。

【0224】

代表的に、これらの処方物は、少なくとも約0.1%またはそれよりも多い活性化合物を含み得るが、活性成分の割合は、もちろん、変化し得、そして好都合に、全処方物の重量または体積の約1または2%と、約60%または70%以上のとの間であり得る。当然、薬学的に有用な処方物の各々の中の活性化合物の量は、適切な投薬量が、化合物の任意の所定の単位用量において得られるような様式で、調製され得る。溶解度、バイオアベイラビリティ、生物学的半減期、投与の経路、製品の有効期限のような因子、および他の薬理学的考慮が、このような薬学的処方物を調製する分野の当業者によって意図され、そして、種々の投薬量および処置レジメンはそれ自体が、所望され得る。

【0225】

あるいは、経口投与について、本発明の処方物は、うがい薬、歯みがき剤、バッカル錠、経口スプレー、または舌下経口投与処方物の形態で、1つ以上の賦形剤と混合され得る。あるいは、活性成分は、経口溶液（例えば、ホウ酸ナトリウム、グリセリンおよび炭酸水素カリウムを含む溶液）に組み込まれ得るか、または歯みがき剤に分散され得るか、または治療的有效量で、水、結合剤、研磨剤、香料、発泡剤、および湿潤剤を含み得る組成物に添加され得る。あるいは、これらの組成物は、舌下に置かれ得るか、そうでなければ口の中で溶解され得る錠剤形態または溶液形態に成形され得る。

【0226】

特定の状況において、本明細書中で開示される薬学的組成物を、非経口送達、静脈内送達、筋肉内送達またはさらに腹腔内送達することが所望される。このようなアプローチは、当業者に周知であり、これらのいくつかは例えば、例えば、米国特許第5,543,158号; 米国特許第5,641,515号および米国特許第5,399,363号にさらに記載される。特定の実施形態において、遊離塩基または薬学的に受容可能な塩としての活性化合物の溶液は、界面活性剤（例えば、ヒドロキシプロピルセルロース）と適切に混合された水中で調製され得る。分散物もまた、グリセロール、液体ポリエチレングリコール、およびそれらの混合物中で、ならびにオイル中で調製され得る。保存および使用の通常

の条件下で、これらの調製物は一般に、微生物の増殖を防止するために、防腐剤を含む。

【0227】

注入用途のために適切な例示的な薬学的形態として、滅菌水性液剤または分散物および滅菌注射用液剤または分散物の即時調製のための滅菌散剤が挙げられる（例えば、米国特許第5,466,468号を参照のこと）。全ての場合において、その形態は、滅菌でなければならない、そして容易に注射することができる程度に、流動性でなければならない。それは、製造および保存の条件下で安定でなければならない、そして微生物（例えば、細菌および真菌）の汚染作用に対して保存されなければならない。キャリアは、例えば、水、エタノール、ポリオール（例えば、グリセロール、プロピレングリコール、および液体ポリエチレングリコールなど）、それらの適切な混合物、および／もしくは植物油を含む溶媒または分散媒体であり得る。適切な流動性は、例えば、コーティング（例えば、レシチン）の使用によって、分散物の場合には必要とされる粒子サイズの維持によって、そして／または界面活性剤の使用によって、維持され得る。微生物の作用の防止は、種々の抗菌剤および抗真菌剤（例えば、パラベン（paraben）、クロロブタノール、フェノール、ソルビン酸、チメロサルなど）によって促進され得る。多くの場合において、等張剤（例えば、糖または塩化ナトリウム）を含めることが、望ましい。注射可能な組成物の長期にわたる吸収は、組成物中での吸収を遅延させる薬剤（例えば、モノステアリン酸アルミニウムおよびゼラチン）の使用によって、もたらされ得る。

10

【0228】

一実施形態において、水溶液の非経口投与のために、必要ならばその溶液は、適切に緩衝化されるべきであり、そして液体希釈剤が、最初に、十分な生理食塩水またはグルコースで等張にされる。これらの特定の水溶液は、特に、静脈内投与、筋肉内投与、皮下投与、および腹腔内投与に適切である。これに関連して、使用され得る滅菌水性媒体は、本開示の観点から当業者に公知である。例えば、一投与量は、1mlの等張性NaCl溶液に溶解され得、そして1000mlの皮下注入流体に添加され得るか、または注入予定部位で注入され得るかのいずれかである（例えば、「Remington's Pharmaceutical Sciences」、第15版、1035～1038頁および1570～1580頁を参照のこと）。投薬量におけるいくらかの変化が、処置される被験体の状態に依存して、さらに、ヒト投与について、調製物は、もちろん、FDA Office of Biologics standardsによって要求されるような、無菌性、

20

30

【0229】

本発明の別の実施形態において、本明細書中で開示される組成物は、中性形態または塩形態で処方され得る。例示的な薬学的に受容可能な塩としては、酸付加塩（タンパク質の遊離アミノ基とともに形成される）が挙げられ、そしてこれらの塩は、無機酸（例えば、塩酸またはリン酸）、または酢酸、シュウ酸、酒石酸、マンデル酸などの有機酸とともに形成される。遊離カルボキシル基とともに形成される塩もまた、無機塩基（例えば、水酸化ナトリウム、水酸化カリウム、水酸化アンモニウム、水酸化カルシウム、または水酸化鉄）、およびイソプロピルアミン、トリメチルアミン、ヒスチジン、プロカインなどの有機塩基から誘導され得る。処方の際に、溶液は、投薬処方物と適合する様式でかつ治療的に有効な量で投与される。

40

【0230】

キャリアは、任意および全ての溶媒、分散媒体、ビヒクル、コーティング、希釈剤、抗菌剤および抗真菌剤、等張剤および吸収を遅延させる薬剤、緩衝液、キャリア溶液、懸濁液、コロイドなどをさらに含み得る。薬学的活性物質のためのこのような媒体および薬剤の使用は、当該分野で周知である。任意の従来の媒体または薬剤が、活性成分と不適合である場合を除いては、治療組成物におけるその使用が、企図される。補助的活性成分もまた、これらの組成物に組み込まれ得る。句「薬学的に受容可能な」とは、ヒトに投与される場合に、アレルギー反応または類似の厄介な反応を生成しない分子実体および組成物をいう。

50

【0231】

特定の実施形態において、薬学的組成物は、鼻腔内スプレー、吸入、および／または他のエアロゾル送達ビヒクルによって送達され得る。遺伝子、核酸およびペプチド組成物は、経鼻エアロゾルスプレーを介して肺に直接送達するための方法は、例えば、米国特許第5,756,353号および米国特許第5,804,212号に記載されている。同様に、鼻腔内微粒子樹脂 (Takenagaら、J Controlled Release 1998 Mar 2; 52(1-2): 81-7) およびリゾホスファチジルグリセロール化合物 (米国特許第5,725,871号) を使用する薬物の送達はまた、薬学分野で周知である。同様に、ポリテトラフルオロエチレン支持マトリックスの形態での経粘膜薬物送達は、米国特許第5,780,045号に記載される。

10

【0232】

特定の実施形態において、リポソーム、ナノカプセル、微粒子、脂質粒子、小胞などを、適切な宿主細胞／生物への本発明の組成物の導入のために使用される。特に、本発明の組成物は、脂質粒子、リポソーム、小胞、ナノスフィア、またはナノ粒子などのいずれかにカプセル化して送達するために処方され得る。あるいは、本発明の組成物は、このようなキャリアビヒクルの表面に、共有結合的または非共有結合的のいずれかで結合され得る。

【0233】

潜在的な薬物キャリアとしてのリポソームおよびリポソーム様調製物の形成および使用は一般に、当業者に公知である (例えば、Lasica, Trends Biotechnol 1998 Jul; 16(7): 307-21; Takakura, Nippon Rinsho 1998 Mar; 56(3): 691-5; Chandranら、Indian J Exp Biol. 1997 Aug; 35(8): 801-9; Margalit, Crit Rev Ther Drug Carrier Syst. 1995; 12(2-3): 233-61; 米国特許第5,567,434号; 米国特許第5,552,157号; 米国特許第5,565,213号; 米国特許第5,738,868号および米国特許第5,795,587号 (これらの各々は、その全体が本明細書中で参考として詳細に援用される) を参照のこと)。

20

【0234】

リポソームは、T細胞懸濁液、初代肝細胞培養物およびPC12細胞を含む他の手順によるトランスフェクションが通常困難である多くの細胞型とともに、首尾よく使用されている (Renneisenら、J Biol Chem. 1990 Sep 25; 265(27): 16337-42; Mullerら、DNA Cell Biol. 1990 Apr; 9(3): 221-9)。さらに、リポソームは、ウイルスベクターの送達系に典型的なDNA長の制限がない。リポソームは、遺伝子、種々の薬物、放射線治療剤、酵素、ウイルス、転写因子およびアロステリックエフェクターなどを、種々の培養細胞株および動物に導入するために効果的に使用されている。さらに、リポソームの使用は、全身送達後の、自己免疫応答にも受容不可能な毒性にも関連しないようである。

30

【0235】

特定の実施形態において、リポソームは、水性媒体中に分散したリン脂質から形成され、そして多層同心性二重層小胞 (多層小胞 (MLV) と呼ばれる) を自発的に形成する。

40

【0236】

あるいは、他の実施形態において、本発明は、本発明の組成物の薬学的に受容可能なナノカプセル処方物を提供する。ナノカプセルは一般に、化合物を安定かつ再現可能な様式で補足し得る (例えば、Quintanar-Guerreroら、Drug Dev Ind Pharm. 1998 Dec; 24(12): 1113-28を参照のこと)。細胞内ポリマー過負荷に起因する副作用を回避するために、このような超微粒子 (約0.1 μmのサイズ) は、インビボで分解され得るポリマーを使用して設計され得る。このような粒子は、例えば、Couvreurら、Crit Rev Ther Drug Carrier Syst. 1988; 5(1): 1-20; zur Muhlenら、Eur J Pharm Biopharm. 1998 Mar; 45(2): 149-5

50

5 ; Zambauxら、J Controlled Release. 1998 Jan 2 ; 50 (1 - 3) : 31 - 40 ; および米国特許第 5 , 145 , 684 号に記載されるようにして作製され得る。

【0237】

(癌治療法)

本発明のさらなる局面において、本明細書に記載される薬学的組成物は、癌の処置、特に、乳癌の免疫治療のために使用され得る。このような方法において、本明細書中に記載される薬学的組成物は、患者、代表的には、温血動物、好ましくはヒトに投与される。患者は、癌に罹患していてもしていなくてもよい。従って、上記の薬学的組成物は、癌の発生を予防するため、または癌に罹患した患者を処置するために使用され得る。薬学的組成物およびワクチンは、原発性腫瘍の外科的除去および／または放射線治療剤もしくは従来の化学療法剤の投与のような処置の前、あるいはその後のいずれかに投与され得る。上記のように、薬学的組成物の投与は、任意の適切な方法（静脈内、腹腔内、筋肉内、皮下、鼻腔内、皮内、肛門内、膈内、局所のおよび経口経路による投与を含む）によってであり得る。

10

【0238】

特定の実施形態において、免疫療法は、能動的免疫療法であり得、この療法において処置は、免疫応答改変剤（例えば、本明細書中で提供されたポリペプチドおよびポリヌクレオチド）の投与で腫瘍に対して反応する内因性宿主免疫系のインビボ刺激に依存する。

【0239】

他の実施形態において、免疫療法は、受動的免疫療法であり得、この療法において処置は、確立された腫瘍免疫反応性を有する因子（例えば、エフェクター細胞または抗体）（抗腫瘍効果を直接的または間接的に媒介し得、そしてインタクトな宿主免疫系に依存する必要はない）の送達を含む。エフェクター細胞の例としては、上記のようなT細胞、本明細書中に提供されるポリペプチドを発現するTリンパ球（例えば、CD8⁺細胞傷害性Tリンパ球およびCD4⁺Tヘルパー腫瘍浸潤性リンパ球）、キラー細胞（例えば、ナチュラルキラー細胞およびリンホカイン活性化キラー細胞）、B細胞および抗原提示細胞（例えば、樹状細胞およびマクロファージ）が挙げられる。本明細書中に列挙されるポリペプチドに特異的なT細胞レセプターおよび抗体レセプターは、養子免疫療法のために他のベクターまたはエフェクター細胞中にクローニングされ、発現され、そして移入され得る。本明細書に提供されるポリペプチドはまた、受動免疫療法のための抗体または抗イディオタイプ抗体（上記および米国特許第4,918,164号に記載される）を生成するために用いられ得る。

20

30

【0240】

エフェクター細胞は、一般に、本明細書中に記載されるように、インビトロでの増殖により養子免疫治療のために十分な量で得られ得る。単一の抗原特異的なエフェクター細胞を、インビボでの抗原認識を保持しながら数十億まで増殖させるための培養条件は当該分野で周知である。このようなインビトロの培養条件は代表的に、しばしばサイトカイン（例えば、IL-2）および分裂しない支持細胞の存在下で、抗原での間欠刺激を用いる。上で述べたように、本明細書中で提供される免疫反応性ポリペプチドは、抗原特異的なT細胞培養物を急速に増殖するために用いられ、免疫治療に十分な数の細胞を生成し得る。詳細には、抗原提示細胞（例えば、樹状細胞、マクロファージ、単球、繊維芽細胞、および／またはB細胞）は、当該分野で周知の標準的技術を用いて、免疫反応性ポリペプチドでパルスされ得るか、または1つ以上のポリヌクレオチドでトランスフェクトされ得る。例えば、抗原提示細胞は、組換えウイルスまたは他の発現系における発現を増大するのに適切なプロモーターを有するポリヌクレオチドでトランスフェクトされ得る。治療において使用するための培養されたエフェクター細胞は、インビボで増殖されかつ広範に分散され得、そして長期間生存され得なければならない。培養されたエフェクター細胞が、インビボで増殖し、そしてIL-2を補充された抗原での反復刺激によって、長期間、多数生存しするように誘導され得ることが研究で示されている（例えば、Cheeverら、Immu

40

50

n o l o g i c a l R e v i e w s 1 5 7 : 1 7 7、1 9 9 7を参照のこと）。

【0241】

あるいは、本明細書において列挙されるポリペプチドを発現するベクターは、患者から得られた抗原提示細胞に導入され得、そして同じ患者に戻す移植のためにエキソビボでクローニング的に増殖され得る。トランスフェクトされた細胞は、当該分野で公知の任意の手段（好ましくは、静脈投与、腔内投与、腹腔内投与、または腫瘍内投与による滅菌形態）を用いて患者に再導入され得る。

【0242】

本明細書中に記載される治療的組成物の投与の経路および頻度、ならびに投薬量は、個々人で異なり、そして標準的技術を用いて容易に確立され得る。概して、薬学的組成物およびワクチンは、注射（例えば、皮内、筋肉内、静脈内、または皮下）により、経鼻的に（例えば、吸引により）または経口的に、投与され得る。好ましくは、52週間にわたって1～10用量の間が投与され得る。好ましくは、1ヶ月の間隔で6用量が投与され、そしてブースター（追加）ワクチン接種がその後定期的に与えられ得る。交互のプロトコールが個々の患者に適切であり得る。適切な用量は、上記のように投与された場合、抗腫瘍免疫応答を促進し得、そして基底（すなわち、未処置）レベルより少なくとも10～50%上である、化合物の量である。このような応答は、患者内の抗腫瘍抗体を測定することによってか、または患者の腫瘍細胞をインビトロで殺傷し得る細胞溶解性エフェクター細胞のワクチン依存性の生成によってモニターされ得る。このようなワクチンはまた、ワクチン接種されていない患者と比較すると、ワクチン接種された患者において、改善された臨床的結果（例えば、より頻繁な寛解、完全もしくは部分的に疾患を有さないか、またはより長く疾患を有さない生存）を導く免疫応答を生じ得るはずである。一般に、1つ以上のポリペプチドを含む薬学的組成物およびワクチンについて、用量中に存在する各ポリペプチドの量は、宿主の体重（kg）あたり、約25 μg～5 mgの範囲である。適切な用量サイズは、患者の大きさで変化するが、代表的には約0.1 mL～約5 mLの範囲である。

【0243】

一般に、適切な投薬量および処置レジメンは、治療的および／または予防的利点を提供するのに十分な量の活性化化合物を提供する。このような応答は、処置されていない患者と比較して、処置された患者において、改善された臨床的結果（例えば、より頻繁な寛解、完全なまたは部分的な、あるいはより長い疾患なしでの生存）を確立することによってモニターされ得る。腫瘍タンパク質に対する既存の免疫応答における増加は、一般的に、改善された臨床的結果と関連する。このような免疫応答は、一般的に、標準的な増殖アッセイ、細胞障害性アッセイまたはサイトカインアッセイを使用して評価され得、これは、処置の前または後に患者から得られるサンプルを使用して行われ得る。

【0244】

（癌の検出および診断組成物、方法およびキット）

一般的に、癌は、患者から得られた生物学的サンプル（例えば、血液、血清、痰、尿、および／または腫瘍生検）における1つ以上の乳房腫瘍タンパク質および／またはこのようなタンパク質をコードするポリヌクレオチドの存在に基づいて患者において検出され得る。言い換えると、このようなタンパク質は、乳癌のような癌の存在または非存在を示すためのマーカーとして使用され得る。さらに、このようなタンパク質は、他の癌の検出に有用であり得る。本明細書に提供される結合剤が、一般的に、生物学的サンプル中の薬剤に結合する抗原のレベルの検出を可能にする。ポリヌクレオチドプライマーおよびプローブは、腫瘍タンパク質をコードするmRNAのレベルを検出するために使用され得、これもまた、癌の存在または非存在を示す。一般に、乳房腫瘍の配列は、正常な組織におけるよりも、腫瘍組織において少なくとも3倍高いレベルで存在するはずである。

【0245】

サンプル中のポリペプチドマーカーを検出するために結合剤を使用するための、当業者に公知の種々のアッセイ型式が存在する。例えば、HarlowおよびLane、Anti

bodies: A Laboratory Manual, Cold Spring Harbor Laboratory, 1988を参照のこと。一般的に、患者における癌の存在または非存在は、(a) 患者から得られた生物学的サンプルを結合剤と接触させる工程；(b) 結合剤に結合するポリペプチドのレベルをサンプルにおいて検出する工程；および(c) ポリペプチドのレベルと所定のカットオフ値とを比較する工程によって決定され得る。

【0246】

好ましい実施形態において、このアッセイは、ポリペプチドに結合するためおよびサンプルの残りからポリペプチドを除くために固体支持体上に固定化された結合剤の使用を含む。次いで、結合されたポリペプチドは、レポーター基を含み、結合剤／ポリペプチド複合体に特異的に結合する検出試薬を使用して検出され得る。このような検出試薬は、例えば、ポリペプチドまたは抗体に特異的に結合する結合剤あるいは結合剤に特異的に結合する他の薬剤（例えば、抗免疫グロブリン、プロテインG、プロテインAまたはレクチン）を含み得る。あるいは、競合アッセイが、使用され得、ここで、ポリペプチドは、レポーター基で標識され、そしてサンプルと結合剤のインキュベーション後にその固定化結合剤に結合される。サンプルの成分が、標識ポリペプチドの結合剤への結合を阻害する程度は、サンプルの固定化結合剤との反応性を示す。このようなアッセイにおける使用に適切なポリペプチドは、上記のような、全長乳房腫瘍タンパク質および結合剤が結合するそのポリペプチド部分を含む。

10

【0247】

固体支持体は、腫瘍タンパク質が付着され得る当業者に公知の任意の材料であり得る。例えば、固体支持体は、マイクロタイタープレートにおける試験ウェルあるいはニトロセルロースまたは他の適切な膜であり得る。あるいは、その支持体は、ビーズまたはディスク（例えば、ガラス）、ファイバークラス、ラテックス、またはプラスチック物質（例えば、ポリスチレン、またはポリ塩化ビニル）であり得る。その支持体はまた、磁気粒子または光ファイバーセンサー（例えば、米国特許第5,359,681号に記載のような）であり得る。この結合剤は、当業者に公知の種々の技術を使用して固体支持体上に固定化され得、これらは特許および科学文献に十分に記載されている。本発明の状況において、用語「固定化」とは、非共有結合的な会合（例えば、吸着）および共有結合的な付着（これは、薬剤と支持体上の官能基との間で直接結合され得るかまたは架橋剤を用いる結合であり得る）の両方をいう。マイクロタイタープレートのウェル、または膜への吸着による固定化は好ましい。このような場合、吸着は、適切な緩衝液中で固体支持体を用いて適切な時間量で結合剤に接触させることによって達成され得る。接触時間は、温度によって変化するが、代表的には、約1時間から約1日の間である。一般的には、約10 ng～約10 μg、そして好ましくは約100 ng～約1 μgの範囲の量の結合剤とプラスチックマイクロタイタープレート（例えば、ポリスチレンまたはポリ塩化ビニル）のウェルを接触させることは、適切な量の結合剤を固定化するのに十分である。

20

30

【0248】

固体支持体への結合剤の共有結合的付着は、一般に、支持体および結合剤上の官能基（例えば、水酸基またはアミノ基）の両方と反応する二官能性試薬と支持体を最初に反応させることによって達成され得る。例えば、この結合剤は、ベンゾキノンを用いるかまたは結合パートナー上のアミンおよび活性水素での支持体上のアルデヒド基の縮合によって、適切なポリマーコーティングを有する支持体に、共有結合的に付着され得る（例えば、Pierce Immunotechnology Catalog and Handbook, 1991, A12-A13を参照のこと）。

40

【0249】

特定の実施形態において、このアッセイは、2抗体サンドイッチアッセイである。本アッセイは、最初に、固体支持体（一般に、マイクロタイタープレートのウェル）上で固定化されている抗体をサンプルに接触させて、その結果、サンプル内のポリペプチドを固定化抗体に結合させることによって実施され得る。次いで、非結合サンプルは固定化ポリペ

50

チドー抗体複合体から除去され、そして検出試薬（好ましくは、そのポリペプチド上の異なる部位に結合し得る第2の抗体（レポーター基を含む））が添加される。次いで、固体支持体に結合したままである検出試薬の量が、特定のレポーター基に関して適切な方法を用いて決定される。

【0250】

より詳細には、一旦抗体が上記のように支持体上に固定化されると、支持体上の残りのタンパク質結合部位は、典型的にはブロックされる。任意の適切なブロック剤（例えば、ウシ血清アルブミンまたはTween 20TM（Sigma Chemical Co., St. Louis, MO））は、当業者に公知である。固定化抗体は次いで、サンプルとインキュベートされ、そしてポリペプチドをこの抗体に結合させる。インキュベーションの前に、このサンプルは適切な希釈液（例えば、リン酸緩衝化生理食塩水（PBS））で希釈され得る。概して、適切な接触時間（すなわち、インキュベーション時間）は、乳癌を有する個体から得られたサンプル内のポリペプチドの存在を検出するのに十分な時間である。好ましくは、この接触時間は、結合ポリペプチドと非結合ポリペプチドとの間の平衡で達成されるレベルの少なくとも約95%の結合レベルを達成するのに十分な時間である。当業者は、ある時間にわたって起こる結合レベルをアッセイすることによって、平衡に達するまでに必要な時間が容易に決定され得ることを認識する。室温では、一般に、約30分間のインキュベーション時間で十分である。

10

【0251】

次いで、非結合サンプルが、適切な緩衝液（例えば、0.1% Tween 20TMを含むPBS）を用いて固体支持体を洗浄することによって除去される。レポーター基を含む第2の抗体が次いで、固体支持体に添加され得る。好ましいレポーター基は、上記の基を含む。

20

【0252】

次いで、検出試薬が、結合されたポリペプチドを検出するのに十分な量の時間、固定化抗体-ポリペプチド複合体とインキュベートされる。適切な量の時間は、一般に、ある時間にわたって起こる結合のレベルをアッセイすることによって決定され得る。次いで、非結合の検出試薬は除去され、そして結合した検出試薬は、レポーター基を用いて検出される。レポーター基を検出するために使用される方法は、レポーター基の性質に依存する。放射性基について、一般的には、シンチレーション計数法またはオートラジオグラフィ法が適切である。分光法は、色素、発光基および蛍光基を検出するために使用され得る。ビチオンは、異なるレポーター基（一般に、放射性もしくは蛍光基または酵素）に結合されたアビジンを使用して検出され得る。酵素レポーター基は、一般に、基質の添加（一般には、特定の時間の間）、続いて反応産物の分光分析または他の分析により検出され得る。

30

【0253】

癌（例えば、乳癌）の存在または非存在を決定するために、固体支持体に結合したままのレポーター基から検出されるシグナルが、一般に、所定のカットオフ値と対応するシグナルと比較される。1つの好ましい実施形態において、癌の検出のためのカットオフ値は、固定化抗体を、癌を有さない患者由来のサンプルとインキュベートした際に得られる平均シグナル値である。概して、所定のカットオフ値を3標準偏差上回るシグナルを生じるサンプルが、癌に対して陽性とみなされる。代替りの好ましい実施形態において、このカットオフ値は、Sackettら、Clinical Epidemiology: A Basic Science for Clinical Medicine, Little Brown and Co., 1985, 106~7頁の方法に従って、レシーバーオペレーターカーブ（Receiver Operator Curve）を使用して決定される。簡単に言うと、本実施形態において、このカットオフ値は、診断試験結果についての各可能なカットオフ値に対応する真の陽性割合（すなわち、感度）および偽陽性割合（100%-特異性）の対のプロットから決定され得る。プロット上の上方左手角に最も近いカットオフ値（すなわち、最大領域を囲む値）が、最も正確なカットオフ値であり、そして本方法によって決定されたカットオフ値より高いシグナルを生ずるサンプルが陽

40

50

性と見なされ得る。あるいは、カットオフ値は、偽陽性割合を最小にするためにプロットに沿って左へシフトされ得るか、または偽陰性割合を最小にするために右へシフトされ得る。概して、本方法によって決定されたカットオフ値より高いシグナルを生ずるサンプルが、癌に対して陽性で見なされる。

【0254】

関連の実施形態において、このアッセイは、フロースルー試験形式またはストリップ試験形式で実行される（ここで、結合剤は、ニトロセルロースのような膜上で固定化される）。フロースルー試験では、サンプル内のポリペプチドは、サンプルが膜を通過するにつれて固定化結合剤に結合する。次いで、第2の標識化された結合剤が、この第2の結合剤を含む溶液がその膜を介して流れるにつれて、結合剤-ポリペプチド複合体と結合する。次いで、結合した第2の結合剤の検出は、上記のように実行され得る。ストリップ試験形式では、結合剤が結合される膜の一端をサンプルを含む溶液中に浸す。このサンプルは、膜に沿って、第2の結合剤を含む領域を通して、そして固定化結合剤の領域まで移動する。固定化抗体の領域での第2の結合剤の濃度が、癌の存在を示す。代表的には、その部位での第2の結合剤の濃度は、視覚的に読みとられ得るパターン（例えば、線）を生成する。このようなパターンを示さないことは陰性の結果を示す。概して、この膜上に固定化される結合剤の量は、生物学的サンプルが、上記の形式において、2抗体サンドイッチアッセイにおいて陽性シグナルを生じるのに十分であるレベルのポリペプチドを含む場合、視覚的に識別可能なパターンを生じるように選択される。このようなアッセイにおける使用に好ましい結合剤は、抗体およびその抗原結合フラグメントである。好ましくは、膜上に固定化される抗体の量は、約25 ng～約1 μgの範囲であり、そしてより好ましくは、約50 ng～約500 ngの範囲である。このような試験は、代表的には、非常に少ない量の生物学的サンプルを用いて実行され得る。

【0255】

もちろん、本発明の腫瘍タンパク質または結合剤との使用に適する多数の他のアッセイ手順が存在する。上記の記載は、例示のみを意図する。例えば、上記手順が、腫瘍ポリペプチドを使用するために容易に改変され得て、生物学的サンプル内でこのようなポリペプチドに結合する抗体を検出し得ることが当業者に明かである。このような腫瘍タンパク質特異的抗体の検出は、癌の存在と相関する。

【0256】

癌もまた、あるいは癌が、生物学的サンプル中の腫瘍タンパク質と特異的に反応するT細胞の存在に基づいて検出され得る。特定の方法では、患者から単離されたCD4⁺T細胞および／またはCD8⁺T細胞を含む生物学的サンプルは、腫瘍ポリペプチド、このようなポリペプチドをコードするポリヌクレオチドおよび／またはこのようなポリペプチドの少なくとも免疫原性部分を発現するAPCとともにインキュベートされ、そしてT細胞の特異的活性化の存在または非存在が検出される。適切な生物学的サンプルとしては、単離されたT細胞が挙げられるがこれらに限定されない。例えば、T細胞は、慣用技術によって（例えば、末梢血リンパ球のFicoll/Hypaque密度勾配遠心分離法によって）患者から単離され得る。T細胞は、2～9日間（代表的に4日間）、37℃にてポリペプチド（例えば、5～25 μg/ml）とともにインビトロでインキュベートされ得る。T細胞サンプルの別のアリコート、コントロールとして役立てるために、腫瘍ポリペプチドの非存在下でインキュベートすることが所望され得る。CD4⁺T細胞に関して、活性化は好ましくは、T細胞の増殖を評価することによって検出される。CD8⁺T細胞に関しては、活性化は好ましくは、細胞溶解活性を評価することによって検出される。疾患のない患者におけるよりも少なくとも2倍高い増殖レベルおよび／または少なくとも20%高い細胞溶解活性レベルは、患者における癌の存在を示す。

【0257】

上記のように、癌はまた、あるいは、生物学的サンプル中の腫瘍タンパク質をコードするmRNAレベルに基づいて検出され得る。例えば、少なくとも2つのオリゴヌクレオチドプライマーをポリメラーゼ連鎖反応（PCR）に基づくアッセイにおいて用いて、生物学

10

20

30

40

50

的サンプルに由来する腫瘍 c D N A の一部を増幅し得、ここで、オリゴヌクレオチドプライマーのうちの少なくとも1つは、腫瘍タンパク質をコードするポリヌクレオチドに特異的である（すなわち、ハイブリダイズする）。次いで、増幅された c D N A が、当該分野で周知の技術（例えば、ゲル電気泳動）を用いて分離され、そして検出される。同様に、腫瘍タンパク質をコードするポリヌクレオチドに特異的にハイブリダイズするオリゴヌクレオチドプローブをハイブリダイゼーションアッセイにおいて用いて、生物学的サンプル中でこの腫瘍タンパク質をコードするポリヌクレオチドの存在を検出し得る。

【0258】

アッセイ条件下でのハイブリダイゼーションを可能にするために、オリゴヌクレオチドのプライマーおよびプローブは、少なくとも10ヌクレオチド、そして好ましくは少なくとも20ヌクレオチドの長さの、本発明の腫瘍タンパク質をコードするポリヌクレオチドの一部に対して少なくとも約60%、好ましくは少なくとも約75%、そしてより好ましくは少なくとも約90%の同一性を有するオリゴヌクレオチド配列を含むべきである。好ましくは、オリゴヌクレオチドプライマーおよび／またはプローブは、上記に規定されるような、中程度にストリンジェントな条件下で、本明細書中に記載されるポリペプチドをコードするポリヌクレオチドにハイブリダイズする。本明細書中に記載される診断方法において有用に用いられ得るオリゴヌクレオチドのプライマーおよび／またはプローブは、好ましくは、少なくとも10～40ヌクレオチドの長さである。好ましい実施形態において、オリゴヌクレオチドプライマーは、本明細書中に開示される配列を有するDNA分子の少なくとも10の連続するヌクレオチド、より好ましくは少なくとも15の連続するヌクレオチドを含む。PCRに基づくアッセイおよびハイブリダイゼーションアッセイの両方についての技術は、当該分野で周知である（例えば、Mullisら, Cold Spring Harbor Symp. Quant. Biol., 51:263, 1987; Erlich編, PCR Technology, Stockton Press, NY, 1989を参照のこと）。

【0259】

1つの好ましいアッセイは、RT-PCRを用い、ここで、PCRは、逆転写と組み合わせて適用される。代表的に、RNAは生物学的サンプル（例えば、生検組織）から抽出され、そして逆転写されてcDNA分子を生成する。少なくとも1つの特異的プライマーを用いるPCR増幅は、cDNA分子を生成し、このcDNA分子は、例えば、ゲル電気泳動を用いて分離および可視化され得る。増幅は、試験患者および癌に罹患していない個体から採取された生物学的サンプルについて行われ得る。増幅反応は、2桁の大きさに及ぶいくつかのcDNA希釈物について行われ得る。試験患者サンプルのいくつかの希釈物における発現の増加が、癌のないサンプルの同一希釈の物と比較して2倍以上である場合、これは、代表的に、陽性とみなされる。

【0260】

別の実施形態において、本明細書中に記載された組成物は、癌の進行についてのマーカーとして使用され得る。この実施形態において、癌の診断について上記に記載されるようなアッセイは、経時的に実行され得、そして反応性ポリペプチドまたはポリヌクレオチドのレベルの変化を評価し得る。例えば、このアッセイは、6ヶ月～1年の期間、24～72時間毎に実行され得、そしてその後、必要に応じて実行され得る。一般に、癌は、検出されるこのポリペプチドまたはポリヌクレオチドのレベルが経時的に増大する患者において進行している。対照的に、癌は、反応性ポリペプチドまたはポリヌクレオチドのレベルが一定のままであるか、または時間とともに減少するかのいずれかである場合、進行していない。

【0261】

特定のインビボ診断アッセイは、腫瘍上で直接実施され得る。1つのこのようなアッセイは、腫瘍細胞を結合剤と接触させる工程を包含する。次いで、結合された結合剤は、レポーター基によって直接的または間接的に検出され得る。このような結合剤はまた、組織学的な用途において使用され得る。あるいは、ポリヌクレオチドプローブは、このような用

途において使用され得る。

【0262】

上記のように、感度を改善するために、複数の腫瘍タンパク質マーカーは、所定のサンプル内でアッセイされ得る。本明細書中に提供される種々のタンパク質に特異的な結合剤が単一のアッセイにおいて組み合わせられ得ることは明かである。さらに、複数のプライマーまたはプローブが同時に用いられ得る。腫瘍タンパク質マーカーの選択は、最適な感度をもたらす組み合わせを決定する慣用実験に基づき得る。さらに、または代替として、本明細書中に提供される腫瘍タンパク質のアッセイは、他の公知の腫瘍抗原に対するアッセイと組み合わせられ得る。

【0263】

本発明はさらに、上記の診断方法のうちのいずれかにおいて使用するためのキットを提供する。このようなキットは代表的に、診断アッセイを行うのに必要な2以上の構成要素を備える。構成要素は、化合物、試薬、容器および／または器具であり得る。例えば、キット内の1つの容器は、腫瘍タンパク質に特異的に結合するモノクローナル抗体またはそのフラグメントを含み得る。このような抗体またはフラグメントは、上記のように支持体材料に付着されて提供され得る。1以上のさらなる容器は、アッセイにおいて使用される要素（例えば、試薬または緩衝液）を封入し得る。このようなキットはまた、あるいは、抗体結合の直接的または間接的な検出に適切なレポーター基を含む上記のような検出試薬を備え得る。

【0264】

あるいは、キットは、生物学的サンプル中の腫瘍タンパク質をコードするmRNAレベルを検出するように設計され得る。このようなキットは一般に、腫瘍タンパク質をコードするポリヌクレオチドにハイブリダイズする、上記のような、少なくとも1つのオリゴヌクレオチドのプローブまたはプライマーを備える。このようなオリゴヌクレオチドは、例えば、PCRまたはハイブリダイゼーションアッセイにおいて用いられ得る。このようなキット内に存在し得るさらなる構成要素としては、腫瘍タンパク質をコードするポリヌクレオチドの検出を容易にする、第2のオリゴヌクレオチドおよび／または診断試薬もしくは容器が挙げられる。

【0265】

以下の実施例は、例示のために提供され、そして限定のためではない。

【0266】

（実施例1）

（ディファレンシャルディスプレイRT-PCRを使用する、乳房腫瘍特異的cDNA配列の調製）

本実施例は、ディファレンシャルディスプレイスクリーニングを使用する、乳房腫瘍特異的ポリペプチドをコードするcDNA分子の調製を例示する。

【0267】

（A. B18Ag1 cDNAの調製およびmRNA発現の特徴づけ）

組織サンプルを、乳癌を患う患者の乳房腫瘍および正常組織（患者から取り出した後に病理学によって確認された）から調製した。正常なRNAおよび腫瘍RNAを、このサンプルから抽出して、そしてmRNAを単離して、そして(dT)₁₂AG（配列番号130）アンカー3'プライマーを使用して、cDNAに変換した。次いで、ディファレンシャルディスプレイPCRを、無作為に選択されたプライマー（CTTCAACCTC）（配列番号103）を使用して行った。増幅条件は、1.5mM MgCl₂、20pmolのプライマー、500pmol dNTPおよび1ユニットのTaq DNAポリメラーゼ（Perkin-Elmer、Branchburg, NJ）を含む標準的な緩衝液であった。40サイクルの増幅を、94℃で30秒間の変性、42℃で1分間のアニーリングおよび72℃で30秒間の伸長を使用して行った。76の増幅された産物を含むRNAフィンガープリントを得た。乳房腫瘍組織のRNAフィンガープリントは、正常な乳房組織のRNAフィンガープリントと98%を越えて同一であったが、バンドは、腫瘍のRN

10

20

30

40

50

Aフィンガープリントパターンに特異的であることが繰り返し観察された。このバンドを、銀染色ゲルから切り出して、Tベクター (Novagen, Madison, WI) にサブクローニングして、そして配列決定した。

【0268】

cDNAの配列 (B18Ag1と呼ばれる) を、配列番号1に示す。GENBANKおよびEMBLのデータベース検索は、最初にクローニングされたB18Ag1フラグメントが、内因性ヒトレトロウイルスエレメントS71 (これは、Simian Sarcoma Virus (SSV) に相同な短縮レトロウイルスエレメントである) に77%同一であることを明かにした。S71は、不完全gag遺伝子、pol遺伝子の一部、および3'末端におけるLTR様構造を含む (Wernerら、Virology 174: 225-238 (1990) を参照のこと)。B18Ag1はまた、P30 (gag) 遺伝子座に対応する領域において、SSVと64%同一である。B18Ag1は、哺乳動物に感染するレトロウイルスの幅広い種々のgagタンパク質とかなりの相同性を共有する領域をカバーする、3つの別々の不完全なリーディングフレームを含む。さらに、S71に対する相同性は、gag遺伝子内のみではなく、LTRを含む配列の数kbにわたる。

【0269】

B18Ag1特異的PCRプライマーは、コンピュータ分析ガイドラインを使用して合成された。RT-PCR増幅 (94℃、30秒; 60℃→42℃、30秒; 72℃、30秒を40サイクル) によって、B18Ag1が患者の乳房腫瘍組織において比較的高いレベルで存在する実際のmRNA配列を表すことが確認された。増幅に使用されるプライマーは、3.5mMのマグネシウム濃度および8.5のpHでの、B18Ag1-1 (CTG CCT GAG CCA CAA ATG) (配列番号128) およびB18Ag1-4 (CCG GAG GAG GAA GCT AGA GGA ATA) (配列番号129)、ならびに2mMのマグネシウム、pH9.5でのB18Ag1-2 (ATG GCT ATT TTC GGG GCC TGA CA) (配列番号126) およびB18Ag1-3 (CCG GTA TCT CCT CGT GGG TAT T) (配列番号127) であった。この患者の正常な乳房組織において、同じ実験は、非常に低いレベルから存在しないレベルの発現を示した (図1を参照のこと)。次いで、RT-PCR実験を使用して、B18Ag1 mRNAが、9個の他の乳房腫瘍サンプル (ブラジル人の患者およびアメリカ人の患者由来) において存在したが、それぞれの癌患者に対応する正常な乳房腫瘍組織において存在しないかまたは非常に低いレベルであることを示した。RT-PCR分析はまた、B18Ag1転写物が、種々の正常組織 (リンパ節、心筋および肝臓を含む) に存在せず、そしてPBM Cおよび肺組織において比較的低いレベルで存在することを示した。乳房腫瘍サンプルにおけるB18Ag1 mRNAの存在、および正常な乳房組織でのその非存在は、図2に示されるように、ノーザンブロット分析によって確認された。

【0270】

乳房腫瘍組織におけるB18Ag1のディファレンシャル発現はまた、RNase保護アッセイによって確認された。図3は、4つの異なるRNase保護アッセイにおいて決定される種々の組織型におけるB18Ag1 mRNAのレベルを示す。レーン1~12は、種々の正常な乳房組織サンプルを表し、レーン13~25は、種々の乳房腫瘍サンプルを表し; レーン26~27は、正常な前立腺サンプルを表し; レーン28~29は、前立腺腫瘍サンプルを表し; レーン30~32は、結腸腫瘍サンプルを表し; レーン33は、正常な大動脈を表し; レーン34は、正常な小腸を表し; レーン35は、正常な皮膚を表し; レーン36は、正常なリンパ節を表し; レーン37は、正常な卵巣を表し; レーン38は、正常な肝臓を表し; レーン39は、正常な骨格筋を表し; レーン40は、正常な第1の正常な胃サンプルを表し; レーン41は、正常な第2の正常な胃サンプルを表し; レーン42は、正常な肺を表し; レーン43は、正常な腎臓を表し; そしてレーン44は、正常な脾臓を表す。実験間の比較は、それぞれのアッセイにおいて公知のβアクチンメッセージの量のポジティブコントロールRNAを含み、そしてこのポジティブコントロール

に対して異なるアッセイの結果を規格化することによって容易にされた。

【0271】

R T-P C Rおよびサザンブロット分析は、B 1 8 A g 1 遺伝子座が、単一コピー内因性レトロウイルスエレメントとしてヒトゲノムD N Aに存在することを示した。最初のB 1 8 A g 1 配列をプローブとして使用して、約1 2 ~ 1 8 k bのゲノムクローンを単離した。4つのさらなるサブクローンをまた、X b a I 消化によって単離した。これらのクローンのX b a I 消化物の末端から得たさらなるレトロウイルス配列（図4に示されるように位置する）を、配列番号3 ~ 配列番号1 0として示し、ここで、配列番号3は、図4において、1 0と標識された配列の位置を示し、配列番号4は、1 1 - 2 9と標識された配列の位置を示し、配列番号5は、3と標識された配列の位置を示し、配列番号6は、6と標識された配列の位置を示し、配列番号7は、1 2と標識された配列の位置を示し、配列番号8は、1 3と標識された配列の位置を示し、配列番号9は、1 4と標識された配列の位置を示し、そして配列番号1 0は、1 1 - 2 2と標識された配列の位置を示す。

10

【0272】

引き続く研究によって、1 2 ~ 1 8 k bのゲノムクローンが図5 Aおよび5 Bに示される約7. 7 5 k bのレトロウイルスエレメントを含むことが、示された。このレトロウイルスエレメントの配列は、配列番号1 4 1に示される。図5 Aの上部に数字を付けられた線は、このレトロウイルスのゲノムクローンのセンス鎖配列を表す。この線の下囲みは、選択された制限部位の位置を示す。矢印は、レトロウイルスエレメントを配列決定するために使用された異なる重複クローンを示す。矢印の方向は、シングルパスサブクローン配列がセンス鎖に対応するかまたはアンチセンス鎖に対応するかを示す。図5 Bは、B 1 8 A g 1を含むレトロウイルスエレメントの概略図であり、エレメント内のウイルス遺伝子の構成を示す。白抜きの囲みは、推定リーディングフレームに対応し、メチオニンから始まり、エレメント全体にわたって見出される。6個の同様なリーディングフレームのそれぞれが、囲みの左に示されるように示され、フレーム1 ~ 3が、センス鎖において見出されたリーディングフレームに対応する。

20

【0273】

配列番号1のc D N Aをプローブとして使用して、より長いc D N A（配列番号2 2 7）を得た。これは、配列番号1 4 1に示されるゲノム配列と比較して、小さなヌクレオチドの差異（1 %未満）を含んだ。

30

【0274】

（B. 他の乳房腫瘍特異的ポリペプチドをコードするc D N A分子の調製）
正常なR N Aおよび腫瘍R N Aを調製し、そしてm R N Aを単離し、そして上記のように、（d T）_{1 2} A Gアンカー3'プライマーを使用してc D N Aに変換した。次いで、ディファレンシャルディスプレイP C Rを、配列番号8 7 ~ 1 2 5の無作為に選択したプライマーを使用して行った。増幅条件は、上記の通りであり、そして腫瘍のR N Aフィンガープリントパターンに特異的であると観察されたバンドを、銀染色ゲルから切り出して、Tベクター（N o v a g e n, M a d i s o n, W I）またはp C R I Iベクター（I n v i t o r g e n, S a n D i e g o, C A）のいずれかにサブクローニングして、そして配列決定した。配列を配列番号1 1 ~ 配列番号8 6に示す。単離された7 9の配列のうち、6 7個が、新規であることが見出された（配列番号1 1 ~ 2 6および2 8 ~ 7 7）（図6 ~ 2 0も参照のこと）。

40

【0275】

抗原B 1 5 A g 1（配列番号2 7において提供される、最初に同定された部分配列）に対して伸長したD N A配列（配列番号2 9 0）をさらなる研究において得た。上記のように、配列番号2 9 0の配列とジーンバンクにおける配列との比較により、既知のヒトβ-Aアクチビン遺伝子に対して相同であることが明らかになった。さらなる研究により、抗原B 2 1 G T 2（B 3 1 1 Dともいわれる；配列番号5 6において提供される、最初に同定された部分c D N A配列）についての全長c D N A配列の単離がもたらされた。この全長配列は、配列番号3 0 7において提供され、対応するアミノ酸配列は配列番号3 0 8に提

50

供される。さらなる研究により、B 3 1 1 Dのスプライス改変体の単離が導かれた。配列番号3 1 6のB 3 1 1 Dクローンを配列決定し、そしてこのクローン由来のX h o I / N o t I フラグメントをゲル精製し、そしてさらなるスクリーニングのためのプローブとして使用するために、ランダムプライミングにより3 2 P - c D T P 標識して、さらなるB 3 1 1 D遺伝子配列を得た。ヒト乳房腫瘍c D N Aの細菌ライブラリーの2つの画分を、標準的技術を用いてスクリーニングした。この様式で単離されたクローンのうちの1つは、ポリA+テイル (t a i l) を含むさらなる配列を生じた。このクローンの決定されたc D N A配列 (B 3 1 1 D _ B T 1 _ 1 Aといわれる) を配列番号3 1 7に提供する。配列番号3 1 6および3 1 7の配列は、4 6 4 b pの領域に対して同一性を共有することが見出され、この配列は、配列番号3 1 7のポリA+配列近辺で分岐している。

10

【0 2 7 6】

引き続き研究により、さらなる1 4 6の配列 (配列番号1 4 2 ~ 2 8 9) を同定した。このうち1 1 5が新規であるようである (配列番号1 4 2、1 4 3、1 4 6 ~ 1 5 2、1 5 4 ~ 1 6 6、1 6 8 ~ 1 7 6、1 7 8 ~ 1 9 2、1 9 4 ~ 1 9 8、2 0 0 ~ 2 0 4、2 0 6、2 0 7、2 0 9 ~ 2 1 4、2 1 6、2 1 8、2 1 9、2 2 1 ~ 2 4 0、2 4 3 ~ 2 4 5、2 4 7、2 5 0、2 5 1、2 5 3、2 5 5、2 5 7 ~ 2 6 6、2 6 8、2 6 9、2 7 1 ~ 2 7 3、2 7 5、2 7 6、2 7 8、2 8 0、2 8 1、2 8 4、2 8 8および2 9 1) 。本発明者らが知る限りでは、これまでに、以前に同定された配列のいずれも、正常乳房組織よりも高いレベルでヒト乳房腫瘍組織において発現されることは示されていない。

【0 2 7 7】

さらなる研究において、抗原B 1 1 A g 1 (B 3 0 5 Dともいわれる) のいくつかの異なるスプライス形態を単離した (種々のスプライス形態の各々は、B 1 1 A g 1 コードフレームのわずかに異なるバージョンを含む) 。スプライス接合部配列は個々のエキソンを規定し、このエキソンは、種々のパターンおよび配置において種々のスプライス形態を作製する。プライマーを設計して、下記のようにR T - P C Rを使用してエキソンの各々の発現パターンを試験した。各エキソンが、最初のB 1 1 A g 1 クローンと同じ発現パターン (この発現は、乳房腫瘍特異的、正常前立腺特異的および正常精巣特異的である) を示すことを見出した。単離したタンパク質をコードするエキソンについて決定したc D N A配列を、それぞれ配列番号2 9 2 ~ 2 9 8に提供する。配列番号2 9 2および2 9 8の配列に対応する推定アミノ酸配列を、配列番号2 9 9および3 0 0に提供する。c D N A末端の迅速な増幅 (R A C E)、上記に提供したB 1 1 A g 1 のスプライス形態のうちの1つに対する5 ' 特異的プライマーおよび乳房腺癌を用いたさらなる研究により、アイソフォームB 1 1 C - 1 5、B 1 1 C - 8およびB 1 1 C - 9、1 6といわれる、3つのさらなる関連したスプライス形態の単離がもたらされた。これらのアイソフォームについての決定されたc D N A配列を、配列番号3 0 1 ~ 3 0 3に提供する。対応する推定アミノ酸配列を、配列番号3 0 4 ~ 3 0 6に提供する。

20

30

【0 2 7 8】

B 1 1 C - 1 5のタンパク質コード領域 (配列番号3 0 1 ; B 3 0 5 DアイソフォームCとも呼ばれる) を、G e n b a n k D N AデータベースのB L A S T N検索における問い合わせ配列として使用した。第2 1染色体由来のゲノムクローン (登録番号A P 0 0 1 4 6 5) に対する一致を見出した。B L A S T N出力において提供される対の整列を、第2 1染色体の配列 (これは、B 3 0 5 D配列に対応する) の推定エキソン、またはコード配列を同定するために使用した。B l a s t Nの対整列に基づいて、G e n B a n k 記録A P 0 0 1 4 6 5の以下の部分をまとめた：塩基対6 7 9 7 8 - 6 8 4 9 9、7 2 8 7 0 - 7 2 9 8 7、7 3 1 4 4 - 7 3 3 3 5、7 6 0 8 5 - 7 6 2 0 6、7 7 9 0 5 - 7 8 0 8 5、8 0 5 2 0 - 8 0 6 2 4、8 7 6 0 2 - 8 7 6 3 3。次いで、この配列を、D N A S t a r S e q m a nプログラムを使用してB 3 0 5 DアイソフォームC配列と整列させ、そしてB 3 0 5 Dに最も類似の配列を維持するような様式で、過剰の配列を除去した。第2 1染色体配列の最終編集形態は、B 3 0 5 Dに9 6 . 5 % 同一であった。次いで、第2 1染色体から得られた編集した配列を翻訳し、B 3 0 5 Dについての終止コドンと

40

50

同じ位置の最終終止コドン以外は、終止コドンを含まないことを見出した。B 3 0 5 D の場合、この第 2 1 染色体配列（配列番号 3 2 5 において提供される）は、3 8 4 アミノ酸を有するタンパク質（配列番号 3 2 6）をコードした。このタンパク質と B 3 0 5 D アイソフォーム C タンパク質（配列番号 3 0 4）との整列は、9 0 % のアミノ酸同一性を示した。

【0 2 7 9】

B 3 0 5 D アイソフォーム C の c D N A 配列（配列番号 3 0 1）を、High Throughput Genome Sequencing (HTGS) データベース (NCBI, National Institutes for Health, Bethesda, MD) を検索することによって、ホモログを同定するために使用した。ホモログを、第 2 染色体 (Clone ID 9 8 3 8 1 8 1)、第 1 0 染色体 (Clone ID 1 0 9 3 3 0 2 2)、第 1 5 染色体 (Clone ID 1 1 5 6 0 2 8 4) において同定した。これらのホモログは、核酸レベルで B 3 0 5 D アイソフォーム C と 9 0 % より高い同一性を共有した。これらのホモログ 3 つ全ては、3 8 4 アミノ酸の ORF（これは、配列番号 3 0 4 のアミノ酸配列と 9 0 % より高い同一性を共有する）をコードする。配列番号 3 0 1 の配列を用いる GenBank データベースのさらなる検索によって、第 2 2 染色体 (Clone ID 5 9 3 1 5 0 7) での部分的配列ホモログが得られた。第 2、1 0、1 5 および 2 2 染色体のホモログについての c D N A 配列を、B 3 0 5 D アイソフォーム C およびイントロン-エキソン接合部で保存される配列との相同性に基づいて構築した。第 2 2、2、1 5 および 1 0 染色体のホモログについての c D N A 配列を、それぞれ配列番号 3 2 7 ~ 3 3 0 に提供し、対応するアミノ酸配列を、それぞれ配列番号 3 3 1、3 3 4、3 3 3 および 3 3 2 に提供する。

【0 2 8 0】

B 3 0 5 D アイソフォーム A（配列番号 2 9 2 に提供される c D N A 配列）についての引き続き研究において、この c D N A 配列（配列番号 3 1 3 に提供される）は 8 8 4 位にさらなるグアニン残基を含むことが見出され、これにより、オープンリーディングフレームにおいてフレームシフトがもたらされていた。この ORF の決定された D N A 配列を、配列番号 3 1 4 に提供する。このフレームシフトは、B 3 0 5 D の他のアイソフォームに共通する C 末端ドメインを含むが N 末端領域において異なる 2 9 3 アミノ酸のタンパク質配列（配列番号 3 1 5 に提供される）を生成する。

【0 2 8 1】

（実施例 2）

（ヒトゲノム D N A 由来の B 1 8 A g 1 D N A の調製）

本実施例は、ヒトゲノム D N A から増幅による B 1 8 A g 1 D N A の調製を例示する。

【0 2 8 2】

B 1 8 A g 1 D N A を、2 0 p m o l の B 1 8 A g 1 特異的プライマー、5 0 0 p m o l の d N T P および 1 ユニットの Taq D N A ポリメラーゼ (Perkin Elmer, Branchburg, NJ) を使用して、以下の増幅パラメーターを使用して 2 5 0 n g のヒトゲノム D N A から調製し得る：9 4 °C で 3 0 秒間の変性、6 0 °C ~ 4 2 °C への 3 0 秒間のタッチダウン (touchdown) アニールリング（2 サイクル毎に 2 °C 増加）および 7 2 °C で 3 0 秒間の伸長。最後の増加（4 2 °C のアニールリング温度）は、2 5 回のサイクルで行うべきである。プライマーを、コンピュータ分析を使用して選択した。合成したプライマーは、B 1 8 A g 1 - 1、B 1 8 A g 1 - 2、B 1 8 A g 1 - 3、および B 1 8 A g 1 - 4 であった。使用され得るプライマー対は、1 + 3、1 + 4、2 + 3 および 2 + 4 である。

【0 2 8 3】

ゲル電気泳動の後に、B 1 8 A g 1 D N A に対応するバンドを切り出し、そして適切なベクター中にクローニングし得る。

【0 2 8 4】

10

20

30

40

50

(実施例 3)

(乳房腫瘍 cDNA 由来の B18Ag1 DNA の調製)

本実施例は、ヒト乳房腫瘍 cDNA からの増幅による、B18Ag1 DNA の調製を例示する。

【0285】

第1鎖 (first strand) cDNA を、500 ng のポリ A+RNA、200 pmol のプライマー (T)₁₂AG (すなわち、TTT TTT TTT TTT AG) (配列番号130)、1×第1鎖逆転写酵素緩衝液、6.7 mM の DTT、500 mmol の dNTP、および1ユニットのAMVまたはMMLV逆転写酵素 (任意の供給業者 (例えば、Gibco-BRL (Grand Island, NY)) から) を30 μl の最終容量で含む反応混合液において、ヒト乳房腫瘍組織から調製したRNAから合成した。第1鎖合成の後に、このcDNAを、およそ25倍に希釈して、そして1 μl を、実施例2に記載されるような増幅に使用した。いくつかのプライマー対が、転写物の不均一な集団を生じ得るが、プライマーB18Ag1-2 (5' ATG GCT ATT T TC GGG GGC TGA CA) (配列番号126) およびB18Ag1-3 (5' CCG GTA TCT CCT CGT GGG TAT T) (配列番号127) は、単一の151 bp の増幅産物を生じる。

【0286】

(実施例 4)

(B18Ag1 の B 細胞および T 細胞のエピトープの同定)

本実施例は、B18Ag1 エピトープの同定を例示する。

【0287】

B18Ag1 配列を、種々のコンピュータアルゴリズムを使用してスクリーニングし得る。B細胞エピトープを決定するために、この配列を、Hopp, Prog. Clin. Biol. Res. 172B:367~77 (1985) またはCeaseら、J. Exp. Med. 164:1779~84 (1986) もしくはSpougeら、J. Immunol. 138:204~12 (1987) の方法を使用して、疎水性および親水性の値についてスクリーニングし得る。さらなるクラスII MHC (抗体またはB細胞) エピトープを、AMPHIのようなプログラム (例えば、Margalitら、J. Immunol. 138:2213 (1987)) またはRothbardおよびTaylorの方法 (例えば、EMBO J. 7:93 (1988)) を使用して、推定し得る。

【0288】

一旦、これらの技術を使用してペプチド (15~20 アミノ酸長) を同定すると、個々のペプチドを、自動化ペプチド合成装置 (製造業者 (例えば、Perkin Elmer / Applied Biosystems Division, Foster City, CA) から入手可能) およびメリーフィールド合成のような技術を使用して合成し得る。合成の後に、このペプチドを使用して、乳癌を患う患者がこのペプチドと反応する抗体を保有するか否かを決定するために、健常者または乳癌患者のいずれかから収集した血清をスクリーニングし得る。乳癌患者におけるこのような抗体の存在により、問題となる特異的B細胞エピトープの免疫原性が確認される。このペプチドをまた、インビボでの免疫後に動物 (マウス、ラット、ウサギ、チンパンジー (chimps) など) において血清学的免疫または液性免疫を生じるそれらの能力について試験し得る。このような免疫の後のペプチド特異的抗血清の生成により、問題となる特異的B細胞エピトープの免疫原性がされに確認される。

【0289】

T細胞エピトープを同定するために、B18Ag1 配列を、HLAクラスI MHC分子に結合し得るB18Ag1 配列内の8~10個のアミノ酸モチーフを同定するのに有用である、異なるコンピュータアルゴリズムを使用してスクリーニングし得る (例えば、Ramnenseeら、Immunogenetics 41:178~228 (1995) を参照のこと)。合成後に、このようなペプチドを、標準的な結合アッセイ (例えば、S

e t t e ら、J. Immunol. 153:5586~92 (1994)) を使用して、クラス I MHC に結合するそれらの能力について試験し得、そしてより重要なことに、患者または正常な末梢単核細胞のインビトロ刺激の後に、抗原反応性細胞傷害性 T 細胞を生成するそれらの能力について、例えば、Bakker ら、Cancer Res. 55:5330-34 (1995); Visseren ら、J. Immunol. 154:3991-98 (1995); Kawakami ら、J. Immunol. 154:3961-68 (1995); および Kast ら、J. Immunol. 152:3904-12 (1994) の方法を使用して、試験し得る。インビトロでのペプチド刺激の後の、自系 (同じクラス I MHC 分子を保有する) 腫瘍細胞を殺傷し得る T 細胞のインビトロでの首尾よい生成により、B18Ag1 抗原の免疫原性がさらに確認される。さらに、この
10
ようなペプチドを使用して、特定のヒト MHC クラス I ハプロタイプを発現するようにトランスジェニックされたマウスにおけるインビボでの免疫の後に、マウスペプチドおよび B18Ag1 反応性細胞傷害性 T 細胞を生成し得る (Vitellio ら、J. Exp. Med. 173:1007~15 (1991))。

【0290】

推定 HLA クラス I MHC 結合抗原に従って破壊された、推定 B18Ag1 B 細胞および T 細胞のエピトープの代表的な表を、以下に示す：

推定 Th モチーフ (B 細胞エピトープ) (配列番号 131~133)

SSGGRTFDDFHRYL LVGI

QGAAQKP INLSKXIEVVQGHDE

SPGVFLEHLQEAYRIYTPFDLSA

推定 HLA A2.1 モチーフ (T 細胞エピトープ) (配列番号 134~140)

YLLVGIQGA

GAAQKP INL

NLSKXIEVV

EVVQGHDES

HLQEAYRIY

NLA FVAQAA

FVAQAAPDS

。

【0291】

(実施例 5)

(B11Ag1 の T 細胞エピトープの同定)

本実施例は、B11Ag1 (B305D ともいわれる) エピトープの同定を例示する。B11-8、B11-1、B11-5 および B11-12 といわれる 4 つのペプチド (それぞれ、配列番号 309~312) を、B11Ag1 遺伝子から得た。

【0292】

ヒト CD8 T 細胞を、Van Tsai ら (Critical Reviews in Immunology 18:65-75, 1998) のプロトコルに従って、樹状細胞を使用してペプチド B11-8 に対してインビトロでプライムした。得られた CD8
40
T 細胞培養物を、B-LCL 株 JY によって提示された、B11-8 ペプチドまたは陰性コントロールペプチドを認識するそれらの能力について試験した。簡潔には、T 細胞を、 $10 \mu\text{g}/\text{ml}$ のペプチド、 $10 \text{ ng}/\text{ml}$ の IL-7 および $10 \mu\text{g}/\text{ml}$ の IL-2 の存在下で自系単球とともにインキュベートし、そして標準的な 51-Cr 放出アッセイにおいて標的細胞を特異的に溶解するそれらの能力についてアッセイした。図 22 に示されるように、バルク培養株は、B11-8 ペプチドの強力な認識、およびペプチド B11-1 の弱い認識を実証した。

【0293】

この CTL 株由来のクローンを、モノクローナル抗体 OKT3 およびヒト IL-2 を使用する迅速な増殖後に単離した。図 23 に示されるように、このクローン (A1 といわれる
50

）は、特定のペプチドを認識し得ることに加えて、B 1 1 A g 1 遺伝子で形質導入された J Y L C L を認識した。このデータは、B 1 1 - 8 が B 1 1 A g 1 遺伝子の天然にプロセスされるエピトープであることを実証する。さらに、これらの T 細胞は、H L A - A 2 制限様式において、B 1 1 A g 1 を天然に発現している樹立腫瘍細胞株を認識しそして溶解することがさらに見出された（図 2 4）。この T 細胞は、H L A - A 2 で形質導入された B 1 1 A g 1 を天然に発現している肺腺癌（L T - 1 4 0 - 2 2）、および B 1 1 A g 1 で形質導入された A 2 + 乳癌（C A M A - 1）を強力に認識するが、非形質導入株または別の陰性腫瘍株（S W 6 2 0）を強力に認識しない。

【0294】

これらのデータは、これらのヒト T 細胞が B 1 1 特異的ペプチドのみでなく、形質導入された細胞および天然に発現する腫瘍株もまた認識することを明確に実証する。 10

【0295】

上記の手順を使用して、抗原 B 1 1 - 5 および B 1 1 - 1 2 に対して惹起された C T L 株は、対応するペプチドでコーティングされた標的を認識することが見出された。

【0296】

（実施例 6）

（ディファレンシャルディスプレイ P C R によって発見された乳房腫瘍遺伝子の特徴付け）

ディファレンシャルディスプレイ P C R によって発見された乳房腫瘍遺伝子の特異性および感受性を、R T - P C R を使用して決定した。この手順によって、多量の R N A を使用することなく、乳房腫瘍遺伝子 m R N A 発現の半定量的な迅速な評価が可能となった。遺伝子特異的プライマーを使用して、種々の組織（8つの乳房腫瘍、5つの正常乳房、2つの前立腺腫瘍、2つの結腸腫瘍、1つの肺腫瘍、および14の他の正常な成人組織（正常な前立腺、結腸、腎臓、肝臓、肺、卵巣、脾臓、骨格筋、皮膚、胃および精巣を含む））における m R N A 発現レベルを試験した。 20

【0297】

R T - P C R の半定量的性質を確実にするために、 β -アクチンを、試験した組織の各々に対する内部コントロールとして使用した。第1鎖 c D N A の系列希釈物を調製し、そして R T - P C R アッセイを、 β -アクチン特異的プライマーを使用して実施した。次いで、 β -アクチンテンプレートの線形範囲増幅を可能にし、そして最初のコピー数の差異を反映するに十分に感受性であった希釈物を選択した。この条件を使用して、 β -アクチンレベルを、各組織からの各逆転写反応に対して決定した。D N A 汚染を、D N a s e 処理によって、そして逆転写酵素を添加することなく調製された第1鎖 c D N A を使用する場合はネガティブな結果を保証することによって、最小化した。 30

【0298】

遺伝子特異的プライマーを使用して、m R N A 発現レベルを種々の組織において決定した。現在まで、38の遺伝子が、R T - P C R によって首尾よく試験され、これらのうちの5つが、乳房腫瘍に対して良好な特異性および感受性を示す（B 1 5 A G - 1、B 3 1 G A 1 b、B 3 8 G A 2 a、B 1 1 A 1 a および B 1 8 A G 1 a）。図 2 1 A および 2 1 B は、これらの遺伝子のうちの3つ（B 1 5 A G - 1（配列番号27）、B 3 1 G A 1 b（配列番号148）および B 3 8 G A 2 a（配列番号157））に対する結果を示す。表 I は、正常乳房組織および乳房腫瘍、ならびに他の組織においてもまた試験した、全ての遺伝子の発現レベルを要約する。 40

【0299】

【表 2】

表 I
種々の組織において発現される乳癌抗原の割合

乳房組織	乳房腫瘍において過剰発現される	84%
	正常と腫瘍とで等しく発現される	16%
他の組織	乳房腫瘍において過剰発現されるが いずれの正常組織においても過剰発現されない	9%
	乳房腫瘍において過剰発現されるが いくつかの正常組織において発現される	30%
	乳房腫瘍において過剰発現されるが 他の全ての組織において等しく発現される	61%

10

(実施例 7)

(乳房腫瘍ポリペプチドに対する抗体の調製および特徴付け)

20

乳房腫瘍抗原 B 3 0 5 D に対するポリクローナル抗体を、以下のように調製した。

【0300】

E. coli 組換え発現系で発現される乳房腫瘍抗原を、振とうインキュベーター中で 37℃で、適切な抗生物質と共に LB ブロス中で一晚増殖させた。次の朝、一晚培養物の 10ml を、2L のバッフル付 Erlenmeyer フラスコ中の 500ml の 2×YT および適切な抗生物質に添加した。培養物の光学濃度 (560nm で) が 0.4~0.6 に達した時点で、細胞を IPTG (1mM) によって誘導した。IPTG による誘導の 4 時間後、細胞を遠心分離によって回収した。次いで、この細胞をリン酸緩衝化生理食塩水で洗浄し、そして再び遠心分離した。上清を捨て、そして細胞をさらなる使用のために凍結させるかまたは直ぐに処理した。溶解緩衝液 20ml を、細胞ペレットに添加し、そしてボルテックスした。E. coli 細胞を破壊して開くために、次いで、この混合物を、16,000psi の圧力で French Press に通した。次いで、細胞を、再び遠心分離し、そして上清およびペレットを、組換えタンパク質を分離するための SDS-PAGE によってチェックした。細胞ペレット中に位置するタンパク質について、このペレットを 10mM Tris pH 8.0、1% CHAPS 中に再懸濁し、そして封入体ペレットを洗浄し、そして再び遠心分離した。この手順をもう 2 回繰り返した。洗浄された封入体ペレットを、8M 尿素または 6M グアニジン HCl (10mM Tris pH 8.0 および 10mM イミダゾールを含む) で可溶化した。可溶化されたタンパク質を、5ml のニッケルキレート樹脂 (Qiagen) に添加し、そして連続して攪拌しながら室温で 45 分~1 時間インキュベートした。インキュベーション後、この樹脂およびタンパク質混合物を、使い捨てのカラムを通して注ぎ、フロースルーを回収した。次いで、カラムの 10~20 容量の可溶化緩衝液でこのカラムを洗浄した。次いで、抗原を、8M 尿素、10mM Tris pH 8.0 および 300mM イミダゾールを使用してカラムから溶出し、そして 3ml の画分に集めた。SDS-PAGE ゲルを、どの画分をさらなる精製のためにプールするかを決定するために、電気泳動した。

30

40

【0301】

最終精製工程として、強アニオン交換樹脂 (例えば、HiPrep Q (Biorad)) を、適切な緩衝液で平衡化して、そして上記からプールされた画分をカラム上にロードした。抗原を、漸増する塩勾配でカラムから溶出した。このカラムに流れるにつれて画分が収集され、そしてカラムからのどの画分をプールするかを決定するために、別の SDS-PAGE

50

A G Eゲルを電気泳動した。プールされた画分を、10 mM T r i s p H 8 . 0 に対して透析した。次いで、タンパク質を、0 . 22 ミクロンフィルターを介した濾過後にバイアルにし (v i a l e d)、そして抗原を免疫のために必要になるまで凍結した。

【0302】

400 μ g の B 3 0 5 D 抗原を、100 μ g のムラミルジペプチド (M D P) と合わせた。等量の不完全フロイントアジュバント (I F A) と混合した100 μ g で、4週間ごとにウサギをブーストした。各々のブーストの7日後、動物を採血した。血液を4℃で12～24時間インキュベートし、次いで遠心分離することによって血清を作製した。

【0303】

96ウェルのプレートに、50 μ l (代表的には1 μ g) の組換えタンパク質と共に4℃ 10
で20時間インキュベートすることによってB305D抗原で被覆した。250 μ l のB
S A ブロッキング緩衝液をウェルに添加し、そして室温で2時間インキュベートした。プ
レートに、P B S / 0 . 01 % T w e e n で6回洗浄した。ウサギ血清を、P B S 中に
希釈した。50 μ l の希釈した血清を各ウェルに添加し、そして室温で30分インキュベ
ートした。プレートを上記のように洗浄し、その後、1 : 10000 の希釈で50 μ l の
ヤギ抗ウサギ西洋ワサビペルオキシダーゼ (H R P) を添加し、そして室温で30分イン
キュベートした。プレートを再び上記のように洗浄し、そして100 μ l のT M B マイク
ロウェルペルオキシダーゼ基質を各ウェルに添加した。室温での暗所での15分間のイン
キュベーション後、比色反応を100 μ l の1 N H ₂ S O ₄ で停止させ、450 nm で 20
直ぐに読み取った。このポリクローナル抗体は、B305Dに対して免疫活性を示した。

【0304】

乳癌標本および正常乳房標本におけるB305D発現の免疫組織化学的 (I H C) 分析を
、以下のように実施した。パラフィンに埋包したホルマール固定した組織を、8ミクロン
の切片にスライスした。0 . 1 M クエン酸ナトリウム緩衝液 (p H 6 . 0) における蒸気
熱誘導エピトープ想起 (S t e a m h e a t i n d u c e d e p i t o p e r e
t r i e v a l) (S H I E R) を、最適な染色条件として使用した。切片を10%血清
/ P B S と共に5分間インキュベートした。一次抗体を、示される濃度において25分間
、各切片に添加し、続いて抗ウサギビオチン化抗体または抗マウスビオチン化抗体のい
ずれかと共に、25分間インキュベートした。内因性ペルオキシダーゼ活性を、過酸化水素 30
と共に1 . 5分間3回インキュベートすることによって、ブロックした。アビジンビオチ
ン複合体 / 西洋ワサビペルオキシダーゼ (A B C / H R P) 系を、D A B 色素原 (c h r
o m a g e n) と共に使用して、抗原発現を可視化した。スライドをヘマトキシリンで対
比染色した。B305D発現は、乳房腫瘍と正常乳房組織との両方において検出された。
しかし、染色の強度は、腫瘍サンプルにおいてより正常サンプルにおいてずっと低く、そ
してB305Dの表面発現は、乳房腫瘍組織においてのみ観察された。

【0305】

正常組織の広範なパネルにおけるB305D発現のリアルタイムPCRおよび免疫組織化
学的分析の要約を、以下の表I I に示す。これらの結果は、精巣においてB305Dの最
小の発現を示し、胆嚢において不十分な結果を示し、そして試験した他の全ての組織にお
いては検出されない。

【0306】

【表3】

表 II

mRNA	IHC染色	組織型	要約
中程度に陽性	陽性	精巣	精子細胞の少量の核染色；精子陰性；精上皮腫陰性
陰性	陰性	胸腺	発現なし
N/A	陰性	動脈	発現なし
陰性	陰性	骨格筋	発現なし
陰性	陽性(弱い染色)	小腸	発現なし
陰性	陽性(弱い染色)	卵巣	発現なし
陰性		下垂体	発現なし
陰性	陽性(弱い染色)	胃	発現なし
陰性	陰性	脊髄	発現なし
陰性	陰性	脾臓	発現なし
陰性	陰性	尿管	発現なし
N/A	陰性	胆嚢	発現なし
N/A	陰性	胎盤	未決定
陰性	陰性	甲状腺	発現なし
陰性	陰性	心臓	発現なし
陰性	陰性	腎臓	発現なし
陰性	陰性	肝臓	発現なし
陰性	陰性	脳-小脳	発現なし
陰性	陰性	結腸	発現なし
陰性	陰性	皮膚	発現なし
陰性	陰性	骨髓	発現なし
N/A	陰性	上皮小体	発現なし
陰性	陰性	肺	発現なし
陰性	陰性	食道	発現なし
陰性	陽性(弱い染色)	子宮	発現なし
陰性	陰性	副腎	発現なし
陰性	陰性	脾臓	発現なし
N/A	陰性	リンパ節	発現なし
陰性	陰性	脳-皮質	発現なし
N/A	陰性	ファローピウス管	発現なし
陰性	陽性(弱い染色)	膀胱	発現なし
陰性	N/A	骨	発現なし
陰性	N/A	唾液腺	発現なし
陰性	N/A	活性化PBMC	発現なし
陰性	N/A	休止PBMC	発現なし
陰性	N/A	気管	発現なし
陰性	N/A	大静脈	発現なし
陰性	N/A	網膜	発現なし
陰性	N/A	軟骨	発現なし

10

20

30

40

50

(実施例 8)

(乳房腫瘍抗原のタンパク質発現)

本実施例は、E. coli および哺乳動物細胞における乳房腫瘍抗原 B305D の発現および精製を記載する。

【0307】

B305D アイソフォーム C-15 (配列番号 301 ; 384 のアミノ酸に翻訳された) の、E. coli における発現を、M. tuberculosis 抗原 Ra12 (配列番号 318) の最初の 30 アミノ酸の下流の、B305D アイソフォーム C-15 のオープンリーディングフレームを pET17b にクローニングすることによって、達成した。第一に、B305D ORF における内部 EcoRI 部位を、タンパク質配列を変化させることなく変異させ、その結果、この遺伝子は、EcoRI 部位において Ra12 でクローニングされ得た。部位特異的変異誘発のために使用される PCR プライマーを、配列番号 319 (AW012 と称する) および配列番号 320 (AW013 と称する) に示す。次

いで、E c o R I 部位改変 B 3 0 5 D の O R F を、プライマー A W 0 1 4 (配列番号 3 2 1) および A W 0 1 5 (配列番号 3 2 2) を使用する P C R によって増幅した。この P C R 産物を E c o R I で消化し、そして E c o R I 部位において R a 1 2 / p E T 1 7 b ベクターに連結した。得られた融合構築物 (R a 1 2 m B 1 1 C と称する) の配列を、D N A 配列決定によって確認した。この融合構築物に対して決定した c D N A 配列を、配列番号 3 2 3 に提供し、そしてアミノ酸配列を配列番号 3 2 4 に提供する。

【0308】

この融合構築物を B L 2 1 (D E 3) C o d o n P l u s - R I L E . c o l i (S t r a t a g e n e) に形質転換し、そしてカナマイシンと共に L B ブロス中で一晚増殖させた。得られた培養物に、I P T G を導入した。タンパク質を P V D F 膜 (m e m b e r 10)) に移し、そして 5 % の脱脂乳 (P B S - T w e e n 緩衝液中) でブロックし、3 回洗浄し、そしてマウス抗 H i s タグ抗体 (C l o n t e c h) と共に 1 時間インキュベートした。この膜を 3 回洗浄し、そして H R P - プロテイン A (Z y m e d) で 3 0 分間プローブした。最後に、この膜を 3 回洗浄し、そして E C L (A m e r s h a m) で現像した。発現を、ウェスタンブロットによって決定した。

【0309】

哺乳動物細胞における組換え発現のために、B 3 0 5 D アイソフォーム C - 1 5 (配列番号 3 0 1 ; 3 8 4 のアミノ酸に翻訳された) を、哺乳動物発現ベクター p C E P 4 および p c D N A 3 . 1 (I n v i t r o g e n) にサブクローニングした。これらの構築物を、F u g e n e 6 試薬 (R o c h e) を使用して H E K 2 9 3 細胞 (A T C C) にトランスフェクトした。簡単に言えば、H E K 細胞を、1 0 % F B S (H y c l o n e) を含む D M E M (G i b c o) 中 1 0 0 , 0 0 0 細胞 / m l の密度でプレートし、そして一晚増殖させた。翌日、2 μ l の F u g e n e 6 を、F B S を含まない 1 0 0 μ l の D M E M に添加し、そして室温で 1 5 分間インキュベートした。この F u g e n e 6 / D M E M 混合物を、1 μ g の B 3 0 5 D / p C E P 4 または B 3 0 5 D / p c D N A プラスミド D N A に添加し、そして室温で 1 5 分間インキュベートした。次いで、F u g e n e / D N A 混合物を H E K 2 9 3 細胞に添加し、そして 7 % C O ₂ で 3 7 °C で 4 8 ~ 7 2 時間インキュベートした。細胞を P B S でリンスし、収集し、そして遠心分離によってペレット化した。

【0310】

ウェスタンブロット分析のために、全細胞溶解物を、T r i t o n - X 1 0 0 含有溶解緩衝液中で細胞を 3 0 分間氷上でインキュベートすることによって、生成した。次いで、溶解物を 1 0 , 0 0 0 r p m で 5 分間、4 °C で遠心分離することによって、清澄化した。サンプルを、 β -メルカプトエタノールを含む S D S _ P A G E 負荷緩衝液で希釈し、そして 1 0 分間煮沸し、その後、S D S _ P A G E ゲルを負荷した。タンパク質をニトロセルロースに移し、そしてプロテイン A で精製した抗 B 3 0 5 D ウサギポリクローナル抗血清 (上記のように調製した) を使用して、1 μ g / m l の濃度でプローブした。ブロットを、H R P に結合体化したヤギ抗ウサギ I g で明らかにし、次いで、E C L 基質中でインキュベートした。B 3 0 5 D の発現は、B 3 0 5 D でトランスフェクトした H E K 2 9 3 溶解物中で検出されたが、ベクター単独でトランスフェクトしたコントロール H E K 2 9 3 40 細胞においては検出されなかった。

【0311】

F A C S 分析のために、細胞をさらに氷冷染色緩衝液で洗浄し、次いでヤギ抗ウサギ I g (H + L) - F I T C 試薬 (S o u t h e r n B i o t e c h n o l o g y) の 1 : 1 0 0 の希釈物と共に 3 0 分間氷上でインキュベートした。3 回の洗浄に続いて、これらの細胞を、P r o p i d i u m I o d i d e (P I) および透過性細胞の同定を可能にするウイルス染料を含む染色緩衝液中に再懸濁させ、次いで F A C S によって分析した。この F A C S 分析は、B 3 0 5 D タンパク質の表面発現を示した。

【0312】

前述から、本発明の特定の実施形態が、説明の目的で本明細書中に記載されたが、本発明 50

の意図および範囲から逸脱することなく、種々の改変がなされ得ることが、理解される。従って、本発明は、添付の特許請求の範囲による以外には、限定されない。

【図面の簡単な説明】

【図 1】

図 1 は、ディファレンシャルディスプレイ P C R 産物を示す。この P C R 産物は、ゲル電気泳動によって分離され、正常な乳房組織から調製された c D N A（レーン 1 およびレーン 2）から、および同じ患者の乳房腫瘍組織から調製された c D N A（レーン 3 およびレーン 4）から得た。矢印は、B 1 8 A g 1 に対応するバンドを示す。

【図 2】

図 2 は、乳房腫瘍組織における B 1 8 A g 1 m R N A のレベル（レーン 1）と、正常な乳房組織の m R N A のレベルとを比較するノーザンブロットである。 10

【図 3】

図 3 は、R N a s e 保護アッセイによって決定された、様々な正常な組織および非腫瘍組織における B 1 8 A g 1 m R N A のレベルを比較した場合の、乳房腫瘍組織における B 1 8 A g 1 m R N A のレベルを示す。

【図 4】

図 4 は、B 1 8 A g 1 に対する X b a I 制限消化物の末端から得られたさらなるレトロウイルス配列（配列番号 3 ～ 配列番号 1 0 として提供される）の位置を示すゲノムクローンマップである。

【図 5】

図 5 A および図 5 B は、レトロウイルスエレメントに対する配列決定ストラテジー、ゲノム構成、および B 1 8 A g 1 を含む推定されたオープンリーディングフレームを示す。 20

【図 6】

図 6 は、代表的な乳房腫瘍特異的 c D N A B 1 8 A g 1 のヌクレオチド配列を示す。

【図 7】

図 7 は、代表的な乳房腫瘍特異的 c D N A B 1 7 A g 1 のヌクレオチド配列を示す。

【図 8】

図 8 は、代表的な乳房腫瘍特異的 c D N A B 1 7 A g 2 のヌクレオチド配列を示す。

【図 9】

図 9 は、代表的な乳房腫瘍特異的 c D N A B 1 3 A g 2 a のヌクレオチド配列を示す。 30

【図 1 0】

図 1 0 は、代表的な乳房腫瘍特異的 c D N A B 1 3 A g 1 b のヌクレオチド配列を示す。

【図 1 1】

図 1 1 は、代表的な乳房腫瘍特異的 c D N A B 1 3 A g 1 a のヌクレオチド配列を示す。

【図 1 2】

図 1 2 は、代表的な乳房腫瘍特異的 c D N A B 1 1 A g 1 のヌクレオチド配列を示す。

【図 1 3】

図 1 3 は、代表的な乳房腫瘍特異的 c D N A B 3 C A 3 c のヌクレオチド配列を示す。 40

【図 1 4】

図 1 4 は、代表的な乳房腫瘍特異的 c D N A B 9 C G 1 のヌクレオチド配列を示す。

【図 1 5】

図 1 5 は、代表的な乳房腫瘍特異的 c D N A B 9 C G 3 のヌクレオチド配列を示す。

【図 1 6】

図 1 6 は、代表的な乳房腫瘍特異的 c D N A B 2 C A 2 のヌクレオチド配列を示す。

【図 1 7】

図 1 7 は、代表的な乳房腫瘍特異的 c D N A B 3 C A 1 のヌクレオチド配列を示す。

【図 1 8】

図 1 8 は、代表的な乳房腫瘍特異的 c D N A B 3 C A 2 のヌクレオチド配列を示す。 50

【図 19】

図 19 は、代表的な乳房腫瘍特異的 cDNA B3CA3 のヌクレオチド配列を示す。

【図 20】

図 20 は、代表的な乳房腫瘍特異的 cDNA B4CA1 のヌクレオチド配列を示す。

【図 21A】

図 21A は、乳房腫瘍組織における乳房腫瘍遺伝子（レーン 1～8）および正常な乳房組織における乳房腫瘍遺伝子（レーン 9～13）ならびに H₂O（レーン 14）の RT-PCR 分析を示す。

【図 21B】

図 21B は、以下における乳房腫瘍遺伝子の RT-PCR 分析を示す：前立腺腫瘍（レーン 1、2）、結腸腫瘍（レーン 3）、肺腫瘍（レーン 4）、正常な前立腺（レーン 5）、正常な結腸（レーン 6）、正常な腎臓（レーン 7）、正常な肝臓（レーン 8）、正常な肺（レーン 9）、正常な卵巣（レーン 10、18）、正常な脾臓（レーン 11、12）、正常な骨格筋（レーン 13）、正常な皮膚（レーン 14）、正常な胃（レーン 15）、正常な精巣（レーン 16）、正常な小腸（レーン 17）、HBL-100（レーン 19）、MCF-12A（レーン 20）、乳房腫瘍（レーン 21～23）、H₂O（レーン 24）および結腸腫瘍（レーン 25）。 10

【図 22】

図 22 は、抗 B11-8 CTL 株による、（B11-8 と称される）B11Ag1 ペプチドの認識を示す。 20

【図 23】

図 23 は、B11-8 特異的クローン A1 による、抗原 B11Ag1 で形質転換された細胞株の認識を示す。

【図 24】

図 24 は、B11-8 特異的クローン A1 による、肺腺癌株（LT-140-22）および乳房腺癌株（CAMA-1）の認識を示す。

【図 1】



Fig. 1

【図 2】



Fig. 2

【図 3】

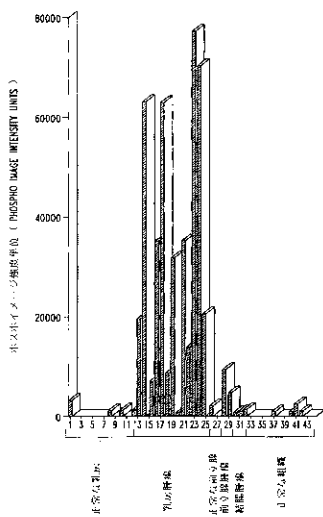


Fig. 3

【図 4】

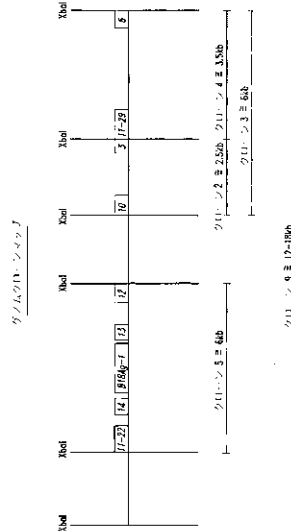


Fig. 4

【図 6】

代表的な乳癌腫瘍特異的 cDNA
U13A の 1 のメクレイフ配列

TTA GAG ACC CAA TTG GGA CCG AAT TGG GAC GCA AAT TTC TCA AGT GGA 43
Leu Glu Thr Glu Leu Gln Pro Asp Pro Asn Phe Ser Ser Gln
5 10 15

GGG AGA ACT TTT GAC GAT TTC CAC CGG TAT CTC CTG GGT ATT CAG 96
Gly Arg Thr Phe Asp Asp Phe His Arg Tyr Leu Leu Val Gln Ile Gln
20 25 30

GGA GCT GCG CAG AAA CCG ATA AAC TGG TGT AAG GCG ATT GAA GTC GTC 144
Gly Ala Thr Gln Lys Pro Ile Asn Leu Ser Lys Ala Ile Gln Val Val
35 40 45

CAG GGG CAT GAT GAG TCA CCA GGA CTG TTT TTA GAG CAC CTC GAG GAG 192
Gln Gly His Asp Glu Ser Pro Gln Val Phe Leu Glu His Leu Gln Gln
50 55 60

GCT TAT GCG ATT TAC ACC CCG TTT GAC CTG GCA GCG CGG GAA AAT AGC 240
Ala Tyr Arg Ile Tyr Thr Pro Phe Asp Leu Ala Ala Pro Glu Asn Ser
65 70 75 80

CAT GCT CTT AAT TTG GGA TTT GTC GCT CAG GCA GCG CCA GAT AGT AAA 288
His Ala Leu Asn Leu Ala Phe Val Ala Gln Ala Ala Pro Asp Ser Lys
85 90 95

AGG AAA CTC CAA AAA CTA GAG GGA TTT TCG TGG AAT GAA TAC CAG TCA 336
Arg Lys Leu Glu Lys Leu Glu Gln Phe Cys Trp Asn Glu Tyr Gln Ser
100 105 110

GCT TTT AGA GAT AGC CTA AAA GGT TTT 363
Ala Phe Arg Asp Ser Leu Lys Gln Phe
115 120

Fig. 6

【図 7】

代表的な乳癌腫瘍特異的 cDNA
U13A の 1 のメクレイフ配列

GC TGGGACAGT GGCCTAACC TGTAACTCTG ACCGTTTCAG AGGCTCAGGT 60

CG CTTAGAGCCA AGATTTCAG ACTAGTCTGG GTAACTAGT GAGAGCCTAT 120

AA AATATAAAA ATGACCTCTG TGTAGTGGCA CACACAGCT GAGGAGGAG 180

CT AGGAGA 196

Fig. 7

【図 8】

代表的な乳癌腫瘍特異的 cDNA
U13A の 2 のメクレイフ配列

GC TTGGGGGTC TGACAGAAA TTAAGGAACT TGGATTTA AGTCAACTG 30
ACTTACCTCTG GAGCCCAATA AACTCTCTCT TCTCTATCC CTCCTATTA 120
AA GGAAGAGCAT GCTCTCTAT AGCCAGTCA GATATCTTAA AGAGAGTAC 180
AT TAAATATCAG AATGTAAAC TGGGAACTA GGTTCAGC CTGGATTAA 240
CA AGAAGAGTGA AGAGTATAC TGGGAAAGC GTGAGAGGCA AATATGTGA 300
TT GAGAGAGTGG TGGGAGATG AATCTCTGT GGCAGTCC CAGCTCACT 360
CT CCTTATAC CTAGAGA 388

Fig. 8

【図 11】

代表的な乳癌腫瘍特異的 cDNA
U13A の 1 のメクレイフ配列

TA ATAACTTAAA TATATTTTGA TTAACTACTG GGGTGAATGA AATATAGATA 60

TT TCCAAAAGC ATAAAGCAAA AATATATATC CAATCAATCT TATATCTCT 120

CC GCACTGAATC TTAACCTCTT AATGTCTAC ATAAAGCAT TTAATCTCT 180

GG TCGCTCTCTA CTACTCTCT TTTTCTTTT TTTTCTTGG AGATGAGTCC 240

CA GCGCAGGGGT GAGTACATAT GCGCAAACT CAGCTCACTG NAAGCTCCG 300

TT CATAGAGATTC TCGGATCTA GCTTCTCCAG TADGTGGGAC TACAGAGTCTG 360

TG CCTGATTAAT CTTTCTTTGA TTTTGGGTAG AGATGGGGT TTATCATGTT 420

TG GTTGTGAGAT CCGTACCTGA AGTGAATCAC CAGCTCAGG CTCGAAGGT 480

TA CAGAGATGAG CAGCTGAGGC CAGCTCTGTT GATCTCTAC TCTCTAGGG 540

Fig. 11

【図 12】

代表的な乳癌腫瘍特異的 cDNA
U13A の 1 のメクレイフ配列

TG CAGATCAGCA ATATCTATC GATCTCTGAT CATTATCTCA TTTCTATGCG 60

AG CTTATCTCA AGATGAGTAT TTAGAGAGAA TTGATTATAG CATAGAGCAA 120

GC ACTCTGACTA CAGCAATCTG TTGAGATCTG ATGGATTAT GACAGTCTAT 180

GA GATTATTAAG TGAATTATT TTAAGGGATC CATTAAATCC AGATATCTCT 240

TC AAGATGATAT AGAATATGAA CAGAGAGAGA CACAAATGA AGATGATCA 300

TA TTGAGAGGCT TATAGTAGAA AATCAATTAG CTGATTTAT TADCTTACA 360

TT TCTCTGATGA ATCTATATT CAGCATGGA CATAGCATTA CTGTATGGG 420

GA ATAAATAGAA CTGGTCTGG GGTATTGAT GAATCATCC NAGTAATAT 480

AC AATATATAG TGGATTGAT TTGATGATG GAATACTAAA TCTGCAAAA 540

GG AGCTATGAT AGCTCTCTT TTGAGATGAG AATATCTTCT TTAGCGTTT 600

CT ACTTTAGGGA TATTTGAGCA TATGAGGA 638

Fig. 12

【図 9】

代表的な乳癌腫瘍特異的 cDNA
U13A の 2 のメクレイフ配列

GC CTAAATCAT GTCTCTATT ATTTCACAT TTTATTAAGC AATTTCTGTT 60

AA AATATGAGGG AATATATGA AACAGGAGG CAAATGTGAG AATATGATC 120

TG ATTCTACAT CAGATGCTCT TCTCTTCTCT GTTATTCTCC TTCTATTTC 180

GG TCGAATGTAA TAGCTTGTG TCAAGCAGAA GTTTTGGGAG TTCTCTAGC 240

CT GCTCATGCT CTGAGGATCT ATTCTACTT TAGGAGGTGT GGTGGAGAG 300

CT ATTTTTCTCA TATTGGGA ACTACTA 337

Fig. 9

【図 10】

代表的な乳癌腫瘍特異的 cDNA
U13A の 1 のメクレイフ配列

GC CATACAGTGC CTTCCTATT ATTTCACCTC CACTGAAATG GCA*AACTG 60

GC TGGTGTCT TT TATGTAAAC AATAAGGAGA CTCTGTCTT CATTTAAAC 120

AT TTCTATATTT ACCTCTGAGG GTTTTACCG GTTCTTTTA ACCTCTTA 180

TT TAAGTGTCTT GGAACAAGAT ATTTTCTCT TCTCTCAGC TTTTCAACTT 240

TT TGTGATGGG GGACTGCTG TCACTGTTCT TCAAGTCTC AATCAAGGC 300

CC AGAAAAAAA AACTTTTGT TTTATTTGA AACTGAGAGC GATAAAGCT 360

CG GCTCTGAT ATAG TTTAA ATGTTTAT TT GACTCTCTT AAGTTCTCT 420

GC GGGGTTTGT NATAGAAAT TTTTNTTCAE AATGCTCAAG GACTTTTMT 480

NA CTGAGCTAAA AAGGCTGTT TT TGGGTGG GGGAGATGA AGGCTCAGC 540

TC TCTTACAGGG GGAATCTCT A 571

Fig. 10

【図 13】

代表的な乳癌腫瘍特異的 cDNA
U3CA の 1 のメクレイフ配列

AGTGATGAT GTGCGCGSAG GCGAGGGGCC TTATCTGATG CTGCGCTGCC TGTCTGTAT 60

GTGCGCGGCC ATTGGCTGT TTATCTCAAA CAGCGCGCAG GGGGTGCTGA TGGCGCTAT 120

TGCTTAGAG GCGCGGAAT CATGCGGGGT CTCACCTCAT CTTTGGCCA TGGTGTGCC 180

GATGCGGGCT TGGCGGGGT TTATGACGCC GGTCTCTGCT CCGGTAAACA CCGTGTGCT 240

TGGCGCTGCC AAGTACTCAT TTAGGATTT TGTCAAAATA GCGCTG 296

Fig. 13

【図 14】

代表的な乳癌腫瘍特異的 cDNA
U3CA の 1 のメクレイフ配列

AG CAGCGCTTC TTCTAATT CATCTGCAAC TACCTGTG TABTATCTCA 60

CA TTTTATAGC CTCCTCTG GTCTCTCTTT TGAATTTCTT GACTGAATC 120

AL ATAACTGCA GTAAATATTT TAAAGCTG GTTATGCTGA TGTACTCTCT 180

AA ATAGTTTCA TTAACTCTTT AATAAATTC GATTGTCT CTCTATTAT 240

CA CCAAGAGCG AA 262

Fig. 14

【図 15】

代表的な乳癌腫瘍特異的 cDNA
U3CA の 1 のメクレイフ配列

AG CAAAGCAGT GGTCTGACT CTCTACTCTG TAACTCTCA AACCAAGGCC 60

TA AATGAGGCA GATTTTAT TATAAGATG TACCTCTCA AATTTCTAT 120

GA TATATCTCT TACATTAACA CATAAATAAT AATCTATTT TAAAGCTCA 180

AG TATAGTGA GTTCTTAA TGAGGGGATA AGCTATAAT CAGAGCTCAA 240

TG CCTATGAGCG A 261

Fig. 15

【図 16】

代表的な乳癌腫瘍特異的 c-DNA
B31GA1 のヌクレオチド配列

```

CGAGCTGGT AAAATGGAG ATGAAGACAC CGGTGGT TTTCTGTGAG GGAATAGGCG 60
GCGCACCCA GAGGAATGGT TGGCTGGATG GCGAAGGAG CCGAGTCTT TGAGAGAG 120
TATGAATCTT GTGTGAAAA TACTGCGCCG CTCTGTGTGA GAGATGGAG TGAATATCTT 180
AATCATGGT GAGCGGATG CTGCGCGGA AGCCG 276

```

Fig. 16

【図 17】

代表的な乳癌腫瘍特異的 c-DNA
B33GA1 のヌクレオチド配列

```

CCAGGTCAA CAGGCTGCA ACAGAGGCT CTTGGATTG GCGAGAGAG AGGCTTGG 60
TGTTTTCAG GAATTTCAC CAGTGGGCTT GGGCTTCTC ATGAGAGAG AGGCTTGG 120
CGAGTCTTT TCCAGGGCC TCTATTGGC TGGAAAAAG CACATCTCA CCGCAAGA 180
CATCTTTGA AATGAGAGT TCCGAGATT CCAGATANT GAGAGAGAG TTTGCGAAT 240
AGGTCA TCTAAGCTTCA TGCACACAT GTGAGCAAT GAAATCTTA ACAATCAAC 300
TAGTGGCAT AANAAGGAA GGAATCTCC AATAGCTGA AGAGCTCTTT TG 352

```

Fig. 17

【図 18】

代表的な乳癌腫瘍特異的 c-DNA
B33GA1 のヌクレオチド配列

```

CTGATGCTC AGGGCTGGA TTACGGGTC GTGAGGTC ACCCTGAGA GCTGCGGCTC 60
TCCGAGACA GCGCAGAGG GATGTTGTTGG GCAAGATAA CCGCTGAGC ATGCTGGA 120
CCTATATGCT TACTGGGAG GAGAGGCTG TGGAGGGGG ATTGAGTTT AGCAGTGGT 180
GGAGCTTTC CACCGGCTC ATGATATGAG ACTGCTGTT TCTAGGGGT CTGCTGGA 240
AGCTGATGG TACGCTGAG CCGAGCATG TGACTTTTG 280

```

Fig. 18

【図 19】

代表的な乳癌腫瘍特異的 c-DNA
B33GA1 のヌクレオチド配列

```

AG GAGAGAGGA GAGGAGTGG AGAGGCTGAG CCTGGTCTG GCTAGAGCT 60
CT GTGCTGGGG ATGCTGAGCA ACTGAGGCG GACCTCTGAG AGCTCTGCG 120
TC GCCGACAGG AGGAGGCTCT TTEGTGATC TGGAGAGAG: GGGGGAGAG 180
TC ATGCTGAGCA TCCC 284

```

Fig. 19

【図 20】

代表的な乳癌腫瘍特異的 c-DNA
B33GA1 のヌクレオチド配列

```

TC AGAGCGGGT AGAGTGGAC CATTGAGGCG ATATTCAAA ATATTATTT 60
TG ATAGTCTG AGTTTCTCT TGCAGCAGA GTTATATGG AGTTATTTT 120
CC AATGCTATG ACATGTTAGA CTATTCTTCT GTTAATGAT TCTATTTTTA 180
GA TTGAGAAAT TGCTTNTAT TATATCAAT TTTGCTATT GTTGAGTTT 240
GC TTAGTATG AGCA 284

```

Fig. 20

【図 21 A】

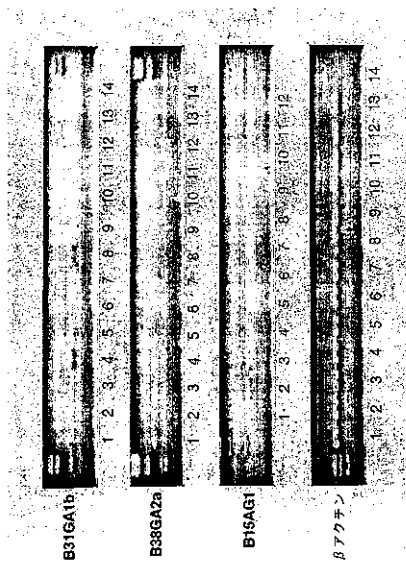


Fig. 21A

【図 21 B】



Fig. 21B

【図 22】

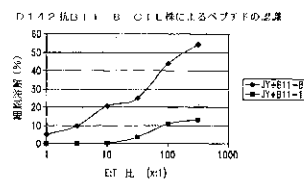


Fig. 22

【図 23】

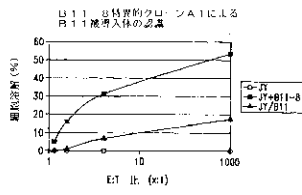


Fig. 23

【図 24】

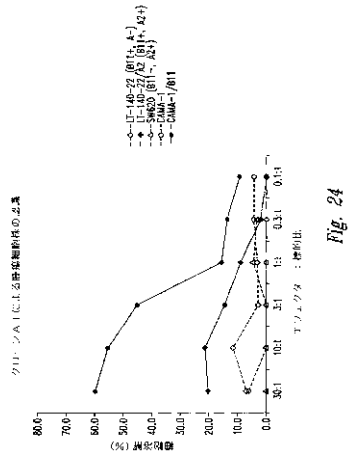


Fig. 24

【国際公開パンフレット】

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

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
29 November 2001 (29.11.2001)

PCT

(10) International Publication Number
WO 01/90152 A2

(51) International Patent Classification: C07K 14/00

(21) International Application Number: PCT/US01/16776

(22) International Filing Date: 22 May 2001 (22.05.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

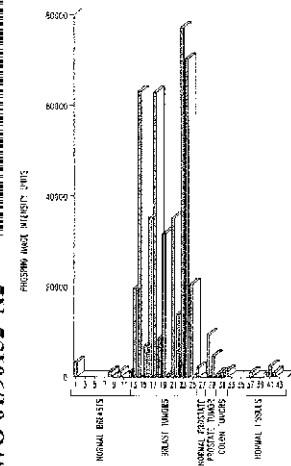
09/577,505	24 May 2000 (24.05.2000)	US
09/590,583	8 June 2000 (08.06.2000)	US
09/609,293	26 October 2000 (26.10.2000)	US
09/810,936	16 March 2001 (16.03.2001)	US

(71) Applicant (for all designated States except US): CORIXA CORPORATION [US/US]; 1124 Columbia Street, Suite 200, Seattle, WA 98104 (US).

(72) Inventors; and
(75) Inventors/Applicants (for US only): ERUDAKIS, Tony, N. [US/US]; 7937 Broadmoor Pines Blvd., Sarasota, FL 34215 (US); REED, Steven, G. [US/US]; 2843 122nd Place N.E., Bellevue, WA 98005 (US); SMITH, John, M. [US/US]; 1101 Pierce Ter N.E., Columbia Heights, MN 55421 (US); MISHK, Lynda, E. [US/US]; 5043 16th Avenue N.E., Seattle, WA 98105 (US); DILLON, David, C. [US/US]; 18112 N.W. Montreux Drive, Issaquah, WA 98027 (US); RETTER, Marc, W. [US/US]; 33402 N.E. 43rd Place, Carnation, WA 98014 (US); WANG, Aijun [CN/US]; 3106 213th Place S.E., Issaquah, WA 98029 (US); SKEEDY, Yasin, A., W. [LB/US]; 15106 S.E. 47th Place, Bellevue, WA 98006 (US); HARLOCKER, Susan, L. [US/US]; 7522 13th Avenue W., Seattle, WA 98117 (US); DAY, Craig, H. [US/US]; 11501 Stone Avenue N., CL22, Seattle, WA 98133 8317 (US).

(Continued on next page)

(54) Title: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS OF BREAST CANCER



(57) Abstract: Compositions and methods for the therapy and diagnosis of cancer, particularly breast cancer, are disclosed. Illustrative compositions comprise one or more breast (tumor) polypeptides, immunogenic portions thereof, polynucleotides that encode such polypeptides, antigen presenting cell that expresses such polypeptides, and T cells that are specific for cells expressing such polypeptides. The disclosed compositions are useful, for example, in the diagnosis, prevention and/or treatment of diseases, particularly breast cancer.

WO 01/90152 A2

WO 01/90152 A2



(74) Agents: **POTTER, Jane, E., R.**, Seed Intellectual Property Law Group PLLC, Suite 6300, 701 Fifth Avenue, Seattle, WA 98104-7092 et al. (US).

(81) Designated States (*optional*): AH, 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, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NG, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TH, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

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

Published:

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

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

WO 01/90152

PCT/US01/16776

COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
OF BREAST CANCER

TECHNICAL FIELD OF THE INVENTION

The present invention relates generally to therapy and diagnosis of
5 cancer, such as breast cancer. The invention is more specifically related to
polypeptides, comprising at least a portion of a breast tumor protein, and to
polynucleotides encoding such polypeptides. Such polypeptides and polynucleotides
are useful in pharmaceutical compositions, e.g., vaccines, and other compositions for
the diagnosis and treatment of breast cancer.

10 BACKGROUND OF THE INVENTION

Breast cancer is a significant health problem for women in the United
States and throughout the world. Although advances have been made in detection and
treatment of the disease, breast cancer remains the second leading cause of cancer-
related deaths in women, affecting more than 180,000 women in the United States each
15 year. For women in North America, the life-time odds of getting breast cancer are now
one in eight.

No vaccine or other universally successful method for the prevention or
treatment of breast cancer is currently available. Management of the disease currently
relies on a combination of early diagnosis (through routine breast screening procedures)
20 and aggressive treatment, which may include one or more of a variety of treatments
such as surgery, radiotherapy, chemotherapy and hormone therapy. The course of
treatment for a particular breast cancer is often selected based on a variety of prognostic
parameters, including an analysis of specific tumor markers. See, e.g., Porter-Jordan
and Lippman, *Breast Cancer* 8:73-100 (1994). However, the use of established markers
25 often leads to a result that is difficult to interpret, and the high mortality observed in
breast cancer patients indicates that improvements are needed in the treatment,
diagnosis and prevention of the disease.

WO 01/90152

PCT/US01/16776

2

Accordingly, there is a need in the art for improved methods for therapy and diagnosis of breast cancer. The present invention fulfills these needs and further provides other related advantages.

SUMMARY OF THE INVENTION

5 In one aspect, the present invention provides polynucleotide compositions comprising a sequence selected from the group consisting of:

- (a) sequences provided in SEQ ID NO: 1, 3-86, 142-298, 301-303, 307, 313, 314, 316, 317, 325 and 327-330;
- (b) complements of the sequences provided in SEQ ID NO: 1, 3-86, 142-298, 301-303, 307, 313, 314, 316, 317, 325 and 327-330;
- (c) sequences consisting of at least 20 contiguous residues of a sequence provided in SEQ ID NO: 1, 3-86, 142-298, 301-303, 307, 313, 314, 316, 317, 325 and 327-330;
- (d) sequences that hybridize to a sequence provided in SEQ ID NO: 1, 3-86, 142-298, 301-303, 307, 313, 314, 316, 317, 325 and 327-330, under moderately stringent conditions;
- (e) sequences having at least 75% identity to a sequence of SEQ ID NO: 1, 3-86, 142-298, 301-303, 307, 313, 314, 316, 317, 325 and 327-330;
- (f) sequences having at least 90% identity to a sequence of SEQ ID NO: 1, 3-86, 142-298, 301-303, 307, 313, 314, 316, 317, 325 and 327-330; and
- (g) degenerate variants of a sequence provided in SEQ ID NO: 1, 3-86, 142-298, 301-303, 307, 313, 314, 316, 317, 325 and 327-330.

25 In one preferred embodiment, the polynucleotide compositions of the invention are expressed in at least about 20%, more preferably in at least about 30%, and most preferably in at least about 50% of breast tumors samples tested, at a level that is at least about 2-fold, preferably at least about 5-fold, and most preferably at least about 10-fold higher than that for normal tissues.

WO 01/90152

PCT/US01/16776

3

The present invention, in another aspect, provides polypeptide compositions comprising an amino acid sequence that is encoded by a polynucleotide sequence described above.

The present invention further provides polypeptide compositions comprising an amino acid sequence selected from the group consisting of sequences recited in SEQ ID NO: 299, 300, 304-306, 308-312, 314, 326 and 331-334.

In certain preferred embodiments, the polypeptides and/or polynucleotides of the present invention are immunogenic, *i.e.*, they are capable of eliciting an immune response, particularly a humoral and/or cellular immune response, as further described herein.

The present invention further provides fragments, variants and/or derivatives of the disclosed polypeptide and/or polynucleotide sequences, wherein the fragments, variants and/or derivatives preferably have a level of immunogenic activity of at least about 50%, preferably at least about 70% and more preferably at least about 90% of the level of immunogenic activity of a polypeptide sequence set forth in SEQ ID NOs: 299, 300, 304-306, 308-312, 314, 326 and 331-334 or a polypeptide sequence encoded by a polynucleotide sequence set forth in SEQ ID NOs: 1, 3-86, 142-298, 301-303, 307, 313, 314, 316, 317, 325 and 327-330.

The present invention further provides polynucleotides that encode a polypeptide described above, expression vectors comprising such polynucleotides and host cells transformed or transfected with such expression vectors.

Within other aspects, the present invention provides pharmaceutical compositions comprising a polypeptide or polynucleotide as described above and a physiologically acceptable carrier.

Within a related aspect of the present invention, the pharmaceutical compositions, *e.g.*, vaccine compositions, are provided for prophylactic or therapeutic applications. Such compositions generally comprise an immunogenic polypeptide or polynucleotide of the invention and an immunostimulant, such as an adjuvant.

The present invention further provides pharmaceutical compositions that comprise: (a) an antibody or antigen-binding fragment thereof that specifically binds to

WO 01/90152

PCT/US01/16776

4

a polypeptide of the present invention, or a fragment thereof; and (b) a physiologically acceptable carrier.

Within further aspects, the present invention provides pharmaceutical compositions comprising: (a) an antigen presenting cell that expresses a polypeptide as described above and (b) a pharmaceutically acceptable carrier or excipient. Illustrative antigen presenting cells include dendritic cells, macrophages, monocytes, fibroblasts and B cells.

Within related aspects, pharmaceutical compositions are provided that comprise: (a) an antigen presenting cell that expresses a polypeptide as described above and (b) an immunostimulant.

The present invention further provides, in other aspects, fusion proteins that comprise at least one polypeptide as described above, as well as polynucleotides encoding such fusion proteins, typically in the form of pharmaceutical compositions, e.g., vaccine compositions, comprising a physiologically acceptable carrier and/or an immunostimulant. The fusion proteins may comprise multiple immunogenic polypeptides or portions/variants thereof, as described herein, and may further comprise one or more polypeptide segments for facilitating the expression, purification and/or immunogenicity of the polypeptide(s).

Within further aspects, the present invention provides methods for stimulating an immune response in a patient, preferably a T cell response in a human patient, comprising administering a pharmaceutical composition described herein. The patient may be afflicted with breast cancer, in which case the methods provide treatment for the disease, or patient considered at risk for such a disease may be treated prophylactically.

Within further aspects, the present invention provides methods for inhibiting the development of a cancer in a patient, comprising administering to a patient a pharmaceutical composition as recited above. The patient may be afflicted with breast cancer, in which case the methods provide treatment for the disease, or patient considered at risk for such a disease may be treated prophylactically.

WO 01/90152

PCT/US01/16776

5

The present invention further provides, within other aspects, methods for removing tumor cells from a biological sample, comprising contacting a biological sample with T cells that specifically react with a polypeptide of the present invention, wherein the step of contacting is performed under conditions and for a time sufficient to permit the removal of cells expressing the protein from the sample.

Within related aspects, methods are provided for inhibiting the development of a cancer in a patient, comprising administering to a patient a biological sample treated as described above.

Methods are further provided, within other aspects, for stimulating and/or expanding T cells specific for a polypeptide of the present invention, comprising contacting T cells with one or more of: (i) a polypeptide as described above; (ii) a polynucleotide encoding such a polypeptide; and/or (iii) an antigen presenting cell that expresses such a polypeptide; under conditions and for a time sufficient to permit the stimulation and/or expansion of T cells. Isolated T cell populations comprising T cells prepared as described above are also provided.

Within further aspects, the present invention provides methods for inhibiting the development of a cancer in a patient, comprising administering to a patient an effective amount of a T cell population as described above.

The present invention further provides methods for inhibiting the development of a cancer in a patient, comprising the steps of: (a) incubating CD4⁺ and/or CD8⁺ T cells isolated from a patient with one or more of: (i) a polypeptide comprising at least an immunogenic portion of polypeptide disclosed herein; (ii) a polynucleotide encoding such a polypeptide; and (iii) an antigen-presenting cell that expressed such a polypeptide; and (b) administering to the patient an effective amount of the proliferated T cells, and thereby inhibiting the development of a cancer in the patient. Proliferated cells may, but need not, be cloned prior to administration to the patient.

Within further aspects, the present invention provides methods for determining the presence or absence of a cancer, preferably a breast cancer, in a patient comprising: (a) contacting a biological sample obtained from a patient with a binding

WO 01/90152

PCT/US01/16776

6

agent that binds to a polypeptide as recited above; (b) detecting in the sample an amount of polypeptide that binds to the binding agent; and (c) comparing the amount of polypeptide with a predetermined cut-off value, and therefrom determining the presence or absence of a cancer in the patient. Within preferred embodiments, the binding agent
5 is an antibody, more preferably a monoclonal antibody.

The present invention also provides, within other aspects, methods for monitoring the progression of a cancer in a patient. Such methods comprise the steps of: (a) contacting a biological sample obtained from a patient at a first point in time with a binding agent that binds to a polypeptide as recited above; (b) detecting in the
10 sample an amount of polypeptide that binds to the binding agent; (c) repeating steps (a) and (b) using a biological sample obtained from the patient at a subsequent point in time; and (d) comparing the amount of polypeptide detected in step (c) with the amount detected in step (b) and therefrom monitoring the progression of the cancer in the patient.

The present invention further provides, within other aspects, methods for determining the presence or absence of a cancer in a patient, comprising the steps of: (a) contacting a biological sample obtained from a patient with an oligonucleotide that hybridizes to a polynucleotide that encodes a polypeptide of the present invention; (b)
15 detecting in the sample a level of a polynucleotide, preferably mRNA, that hybridizes to the oligonucleotide; and (c) comparing the level of polynucleotide that hybridizes to the oligonucleotide with a predetermined cut-off value, and therefrom determining the presence or absence of a cancer in the patient. Within certain embodiments, the amount of mRNA is detected via polymerase chain reaction using, for example, at least one oligonucleotide primer that hybridizes to a polynucleotide encoding a polypeptide as
20 recited above, or a complement of such a polynucleotide. Within other embodiments, the amount of mRNA is detected using a hybridization technique, employing an oligonucleotide probe that hybridizes to a polynucleotide that encodes a polypeptide as recited above, or a complement of such a polynucleotide.

In related aspects, methods are provided for monitoring the progression
30 of a cancer in a patient, comprising the steps of: (a) contacting a biological sample

WO 01/90152

PCT/US01/16776

7

obtained from a patient with an oligonucleotide that hybridizes to a polynucleotide that encodes a polypeptide of the present invention; (b) detecting in the sample an amount of a polynucleotide that hybridizes to the oligonucleotide; (c) repeating steps (a) and (b) using a biological sample obtained from the patient at a subsequent point in time; and
5 (d) comparing the amount of polynucleotide detected in step (c) with the amount detected in step (b) and therefrom monitoring the progression of the cancer in the patient.

Within further aspects, the present invention provides antibodies, such as monoclonal antibodies, that bind to a polypeptide as described above, as well as
10 diagnostic kits comprising such antibodies. Diagnostic kits comprising one or more oligonucleotide probes or primers as described above are also provided.

These and other aspects of the present invention will become apparent upon reference to the following detailed description and attached drawings. All references disclosed herein are hereby incorporated by reference in their entirety as if
15 each was incorporated individually.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows the differential display PCR products, separated by gel electrophoresis, obtained from cDNA prepared from normal breast tissue (lanes 1 and 2) and from cDNA prepared from breast tumor tissue from the same patient (lanes 3 and
20 4). The arrow indicates the band corresponding to B18Ag1.

Figure 2 is a northern blot comparing the level of B18Ag1 mRNA in breast tumor tissue (lane 1) with the level in normal breast tissue.

Figure 3 shows the level of B18Ag1 mRNA in breast tumor tissue compared to that in various normal and non-breast tumor tissues as determined by
25 RNase protection assays.

Figure 4 is a genomic clone map showing the location of additional retroviral sequences obtained from ends of XbaI restriction digests (provided in SEQ ID NO:3 - SEQ ID NO:10) relative to B18Ag1.

WO 01/90152

PCT/US01/16776

8

Figures 5A and 5B show the sequencing strategy, genomic organization and predicted open reading frame for the retroviral element containing B18Ag1.

Figure 6 shows the nucleotide sequence of the representative breast tumor-specific cDNA B18Ag1.

5 Figure 7 shows the nucleotide sequence of the representative breast tumor-specific cDNA B17Ag1.

Figure 8 shows the nucleotide sequence of the representative breast tumor-specific cDNA B17Ag2.

Figure 9 shows the nucleotide sequence of the representative breast
10 tumor-specific cDNA B13Ag2a.

Figure 10 shows the nucleotide sequence of the representative breast tumor-specific cDNA B13Ag1b.

Figure 11 shows the nucleotide sequence of the representative breast tumor-specific cDNA B13Ag1a.

15 Figure 12 shows the nucleotide sequence of the representative breast tumor-specific cDNA B11Ag1.

Figure 13 shows the nucleotide sequence of the representative breast tumor-specific cDNA B3CA3c.

Figure 14 shows the nucleotide sequence of the representative breast
20 tumor-specific cDNA B9CG1.

Figure 15 shows the nucleotide sequence of the representative breast tumor-specific cDNA B9CG3.

Figure 16 shows the nucleotide sequence of the representative breast tumor-specific cDNA B2CA2.

25 Figure 17 shows the nucleotide sequence of the representative breast tumor-specific cDNA B3CA1.

Figure 18 shows the nucleotide sequence of the representative breast tumor-specific cDNA B3CA2.

Figure 19 shows the nucleotide sequence of the representative breast
30 tumor-specific cDNA B3CA3.

WO 01/90152

PCT/US01/16776

9

Figure 20 shows the nucleotide sequence of the representative breast tumor-specific cDNA B4CA1.

Figure 21A depicts RT-PCR analysis of breast tumor genes in breast tumor tissues (lanes 1-8) and normal breast tissues (lanes 9-13) and H₂O (lane 14).

5 Figure 21B depicts RT-PCR analysis of breast tumor genes in prostate tumors (lane 1, 2), colon tumors (lane 3), lung tumor (lane 4), normal prostate (lane 5), normal colon (lane 6), normal kidney (lane 7), normal liver (lane 8), normal lung (lane 9), normal ovary (lanes 10, 18), normal pancreases (lanes 11, 12), normal skeletal muscle (lane 13), normal skin (lane 14), normal stomach (lane 15), normal testes (lane 10 16), normal small intestine (lane 17), HBL-100 (lane 19), MCF-12A (lane 20), breast tumors (lanes 21-23), H₂O (lane 24), and colon tumor (lane 25).

Figure 22 shows the recognition of a B11Ag1 peptide (referred to as B11-8) by an anti-B11-8 CTL line.

Figure 23 shows the recognition of a cell line transduced with the antigen B11Ag1 by the B11-8 specific clone A1.

Figure 24 shows recognition of a lung adenocarcinoma line (LT-140-22) and a breast adenocarcinoma line (CAMA-1) by the B11-8 specific clone A1.

DETAILED DESCRIPTION OF THE INVENTION

20 The present invention is directed generally to compositions and their use in the therapy and diagnosis of cancer, particularly breast cancer. As described further below, illustrative compositions of the present invention include, but are not restricted to, polypeptides, particularly immunogenic polypeptides, polynucleotides encoding such polypeptides, antibodies and other binding agents, antigen presenting cells (APCs) and immune system cells (e.g., T cells).

25 The practice of the present invention will employ, unless indicated specifically to the contrary, conventional methods of virology, immunology, microbiology, molecular biology and recombinant DNA techniques within the skill of the art, many of which are described below for the purpose of illustration. Such techniques are explained fully in the literature. See, e.g., Sambrook, et al. Molecular

WO 01/90152

PCT/US01/16776

10

Cloning: A Laboratory Manual (2nd Edition, 1989); Maniatis et al. Molecular Cloning: A Laboratory Manual (1982); DNA Cloning: A Practical Approach, vol. I & II (D. Glover, ed.); Oligonucleotide Synthesis (N. Gait, ed., 1984); Nucleic Acid Hybridization (B. Hames & S. Higgins, eds., 1985); Transcription and Translation (B. Hames & S. Higgins, eds., 1984); Animal Cell Culture (R. Freshney, ed., 1986); Perbal, A Practical Guide to Molecular Cloning (1984).

All publications, patents and patent applications cited herein, whether *supra* or *infra*, are hereby incorporated by reference in their entirety.

As used in this specification and the appended claims, the singular forms "a," "an" and "the" include plural references unless the content clearly dictates otherwise.

Polypeptide Compositions

As used herein, the term "polypeptide" is used in its conventional meaning, *i.e.*, as a sequence of amino acids. The polypeptides are not limited to a specific length of the product; thus, peptides, oligopeptides, and proteins are included within the definition of polypeptide, and such terms may be used interchangeably herein unless specifically indicated otherwise. This term also does not refer to or exclude post-expression modifications of the polypeptide, for example, glycosylations, acetylations, phosphorylations and the like, as well as other modifications known in the art, both naturally occurring and non-naturally occurring. A polypeptide may be an entire protein, or a subsequence thereof. Particular polypeptides of interest in the context of this invention are amino acid subsequences comprising epitopes, *i.e.*, antigenic determinants substantially responsible for the immunogenic properties of a polypeptide and being capable of evoking an immune response.

Particularly illustrative polypeptides of the present invention comprise those encoded by a polynucleotide sequence set forth in any one of SEQ ID NOs: 1, 3-86, 142-298, 301-303, 307, 313, 314, 316, 317, 325 and 327-330, or a sequence that hybridizes under moderately stringent conditions, or, alternatively, under highly stringent conditions, to a polynucleotide sequence set forth in any one of SEQ ID NOs:

WO 01/90152

PCT/US01/16776

11

1, 3-86, 142-298, 301-303, 307, 313, 314, 316, 317, 325 and 327-330. Certain other illustrative polypeptides of the invention comprise amino acid sequences as set forth in any one of SEQ ID NOs: 299, 300, 304-306, 308-312, 314, 326 and 331-334.

The polypeptides of the present invention are sometimes herein referred to as breast tumor proteins or breast tumor polypeptides, as an indication that their identification has been based at least in part upon their increased levels of expression in breast tumor samples. Thus, a "breast tumor polypeptide" or "breast tumor protein," refers generally to a polypeptide sequence of the present invention, or a polynucleotide sequence encoding such a polypeptide, that is expressed in a substantial proportion of breast tumor samples, for example preferably greater than about 20%, more preferably greater than about 30%, and most preferably greater than about 50% or more of breast tumor samples tested, at a level that is at least two fold, and preferably at least five fold, greater than the level of expression in normal tissues, as determined using a representative assay provided herein. A breast tumor polypeptide sequence of the invention, based upon its increased level of expression in tumor cells, has particular utility both as a diagnostic marker as well as a therapeutic target, as further described below.

In certain preferred embodiments, the polypeptides of the invention are immunogenic, i.e., they react detectably within an immunoassay (such as an ELISA or T-cell stimulation assay) with antisera and/or T-cells from a patient with breast cancer. Screening for immunogenic activity can be performed using techniques well known to the skilled artisan. For example, such screens can be performed using methods such as those described in Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988. In one illustrative example, a polypeptide may be immobilized on a solid support and contacted with patient sera to allow binding of antibodies within the sera to the immobilized polypeptide. Unbound sera may then be removed and bound antibodies detected using, for example, ¹²⁵I-labeled Protein A.

As would be recognized by the skilled artisan, immunogenic portions of the polypeptides disclosed herein are also encompassed by the present invention. An "immunogenic portion," as used herein, is a fragment of an immunogenic polypeptide

WO 01/90152

PCT/US01/16776

12

of the invention that itself is immunologically reactive (*i.e.*, specifically binds) with the B-cells and/or T-cell surface antigen receptors that recognize the polypeptide. Immunogenic portions may generally be identified using well known techniques, such as those summarized in Paul, *Fundamental Immunology*, 3rd ed., 243-247 (Raven Press, 5 1993) and references cited therein. Such techniques include screening polypeptides for the ability to react with antigen-specific antibodies, antisera and/or T-cell lines or clones. As used herein, antisera and antibodies are "antigen-specific" if they specifically bind to an antigen (*i.e.*, they react with the protein in an ELISA or other immunoassay, and do not react detectably with unrelated proteins). Such antisera and 10 antibodies may be prepared as described herein, and using well-known techniques.

In one preferred embodiment, an immunogenic portion of a polypeptide of the present invention is a portion that reacts with antisera and/or T-cells at a level that is not substantially less than the reactivity of the full-length polypeptide (*e.g.*, in an ELISA and/or T-cell reactivity assay). Preferably, the level of immunogenic activity of 15 the immunogenic portion is at least about 50%, preferably at least about 70% and most preferably greater than about 90% of the immunogenicity for the full-length polypeptide. In some instances, preferred immunogenic portions will be identified that have a level of immunogenic activity greater than that of the corresponding full-length polypeptide, *e.g.*, having greater than about 100% or 150% or more immunogenic 20 activity.

In certain other embodiments, illustrative immunogenic portions may include peptides in which an N-terminal leader sequence and/or transmembrane domain have been deleted. Other illustrative immunogenic portions will contain a small N- and/or C-terminal deletion (*e.g.*, 1-30 amino acids, preferably 5-15 amino acids), 25 relative to the mature protein.

In another embodiment, a polypeptide composition of the invention may also comprise one or more polypeptides that are immunologically reactive with T cells and/or antibodies generated against a polypeptide of the invention, particularly a polypeptide having an amino acid sequence disclosed herein, or to an immunogenic 30 fragment or variant thereof.

WO 01/90152

PCT/US01/16776

13

In another embodiment of the invention, polypeptides are provided that comprise one or more polypeptides that are capable of eliciting T cells and/or antibodies that are immunologically reactive with one or more polypeptides described herein, or one or more polypeptides encoded by contiguous nucleic acid sequences contained in the polynucleotide sequences disclosed herein, or immunogenic fragments or variants thereof, or to one or more nucleic acid sequences which hybridize to one or more of these sequences under conditions of moderate to high stringency.

The present invention, in another aspect, provides polypeptide fragments comprising at least about 5, 10, 15, 20, 25, 50, or 100 contiguous amino acids, or more, including all intermediate lengths, of a polypeptide compositions set forth herein, such as those set forth in SEQ ID NOs: 299, 300, 304-306, 308-312, 314, 326 and 331-334, or those encoded by a polynucleotide sequence set forth in a sequence of SEQ ID NOs: 1, 3-86, 142-298, 301-303, 307, 313, 314, 316, 317, 325 and 327-330.

In another aspect, the present invention provides variants of the polypeptide compositions described herein. Polypeptide variants generally encompassed by the present invention will typically exhibit at least about 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% or more identity (determined as described below), along its length, to a polypeptide sequences set forth herein.

In one preferred embodiment, the polypeptide fragments and variants provide by the present invention are immunologically reactive with an antibody and/or T-cell that reacts with a full-length polypeptide specifically set forth herein.

In another preferred embodiment, the polypeptide fragments and variants provided by the present invention exhibit a level of immunogenic activity of at least about 50%, preferably at least about 70%, and most preferably at least about 90% or more of that exhibited by a full-length polypeptide sequence specifically set forth herein.

A polypeptide "variant," as the term is used herein, is a polypeptide that typically differs from a polypeptide specifically disclosed herein in one or more substitutions, deletions, additions and/or insertions. Such variants may be naturally

WO 01/90152

PCT/US01/16776

14

occurring or may be synthetically generated, for example, by modifying one or more of the above polypeptide sequences of the invention and evaluating their immunogenic activity as described herein and/or using any of a number of techniques well known in the art.

5 For example, certain illustrative variants of the polypeptides of the invention include those in which one or more portions, such as an N-terminal leader sequence or transmembrane domain, have been removed. Other illustrative variants include variants in which a small portion (*e.g.*, 1-30 amino acids, preferably 5-15 amino acids) has been removed from the N- and/or C-terminal of the mature protein.

10 In many instances, a variant will contain conservative substitutions. A "conservative substitution" is one in which an amino acid is substituted for another amino acid that has similar properties, such that one skilled in the art of peptide chemistry would expect the secondary structure and hydrophobic nature of the polypeptide to be substantially unchanged. As described above, modifications may be
15 made in the structure of the polynucleotides and polypeptides of the present invention and still obtain a functional molecule that encodes a variant or derivative polypeptide with desirable characteristics, *e.g.*, with immunogenic characteristics. When it is desired to alter the amino acid sequence of a polypeptide to create an equivalent, or even an improved, immunogenic variant or portion of a polypeptide of the invention,
20 one skilled in the art will typically change one or more of the codons of the encoding DNA sequence according to Table 1.

For example, certain amino acids may be substituted for other amino acids in a protein structure without appreciable loss of interactive binding capacity with structures such as, for example, antigen-binding regions of antibodies or binding sites
25 on substrate molecules. Since it is the interactive capacity and nature of a protein that defines that protein's biological functional activity, certain amino acid sequence substitutions can be made in a protein sequence, and, of course, its underlying DNA coding sequence, and nevertheless obtain a protein with like properties. It is thus contemplated that various changes may be made in the peptide sequences of the

WO 01/90152

PCT/US01/16776

15

disclosed compositions, or corresponding DNA sequences which encode said peptides without appreciable loss of their biological utility or activity.

TABLE I

5

Amino Acids			Codons						
Alanine	Ala	A	GCA	GCC	GCG	GCU			
Cysteine	Cys	C	UGC	UGU					
Aspartic acid	Asp	D	GAC	GAU					
Glutamic acid	Glu	E	GAA	GAG					
Phenylalanine	Phe	F	UUC	UUU					
Glycine	Gly	G	GGA	GGC	GGG	GGU			
Histidine	His	H	CAC	CAU					
Isoleucine	Ile	I	AUA	AUC	AUU				
Lysine	Lys	K	AAA	AAG					
Leucine	Leu	L	UUA	UUG	CUA	CUC	CUG	CUU	
Methionine	Met	M	AUG						
Asparagine	Asn	N	AAC	AAU					
Proline	Pro	P	CCA	CCC	CCG	CCU			
Glutamine	Gln	Q	CAA	CAG					
Arginine	Arg	R	AGA	AGG	CGA	CGC	CGG	CGU	
Serine	Ser	S	AGC	AGU	UCA	UCC	UCG	UCU	
Threonine	Thr	T	ACA	ACC	ACG	ACU			
Valine	Val	V	GUA	GUC	GUG	GUU			
Tryptophan	Trp	W	UGG						
Tyrosine	Tyr	Y	UAC	UAU					

In making such changes, the hydropathic index of amino acids may be considered. The importance of the hydropathic amino acid index in conferring interactive biologic function on a protein is generally understood in the art (Kyte and Doolittle, 1982, incorporated herein by reference). It is accepted that the relative hydropathic character of the amino acid contributes to the secondary structure of the

WO 01/90152

PCT/US01/16776

16

resultant protein, which in turn defines the interaction of the protein with other molecules, for example, enzymes, substrates, receptors, DNA, antibodies, antigens, and the like. Each amino acid has been assigned a hydrophobic index on the basis of its hydrophobicity and charge characteristics (Kyte and Doolittle, 1982). These values are:

5 isoleucine (+4.5); valine (+4.2); leucine (+3.8); phenylalanine (+2.8); cysteine/cystine (+2.5); methionine (+1.9); alanine (+1.8); glycine (-0.4); threonine (-0.7); serine (-0.8); tryptophan (-0.9); tyrosine (-1.3); proline (-1.6); histidine (-3.2); glutamate (-3.5); glutamine (-3.5); aspartate (-3.5); asparagine (-3.5); lysine (-3.9); and arginine (-4.5).

It is known in the art that certain amino acids may be substituted by other
10 amino acids having a similar hydrophobic index or score and still result in a protein with similar biological activity, i.e. still obtain a biologically equivalent protein. In making such changes, the substitution of amino acids whose hydrophobic indices are within ± 2 is preferred, those within ± 1 are particularly preferred, and those within ± 0.5 are even more particularly preferred. It is also understood in the art that the substitution
15 of like amino acids can be made effectively on the basis of hydrophilicity. U. S. Patent 4,554,101 (specifically incorporated herein by reference in its entirety), states that the greatest local average hydrophilicity of a protein, as governed by the hydrophilicity of its adjacent amino acids, correlates with a biological property of the protein.

As detailed in U. S. Patent 4,554,101, the following hydrophilicity values
20 have been assigned to amino acid residues: arginine (+3.0); lysine (+3.0); aspartate (+3.0 $\pm$ 1); glutamate (+3.0 $\pm$ 1); serine (+0.3); asparagine (+0.2); glutamine (+0.2); glycine (0); threonine (-0.4); proline (-0.5 $\pm$ 1); alanine (-0.5); histidine (-0.5); cysteine (-1.0); methionine (-1.3); valine (-1.5); leucine (-1.8); isoleucine (-1.8); tyrosine (-2.3); phenylalanine (-2.5); tryptophan (-3.4). It is understood that an amino acid can be
25 substituted for another having a similar hydrophilicity value and still obtain a biologically equivalent, and in particular, an immunologically equivalent protein. In such changes, the substitution of amino acids whose hydrophilicity values are within ± 2 is preferred, those within ± 1 are particularly preferred, and those within ± 0.5 are even more particularly preferred.

WO 01/90152

PCT/US01/16776

17

As outlined above, amino acid substitutions are generally therefore based on the relative similarity of the amino acid side-chain substituents, for example, their hydrophobicity, hydrophilicity, charge, size, and the like. Exemplary substitutions that take various of the foregoing characteristics into consideration are well known to those of skill in the art and include: arginine and lysine; glutamate and aspartate; serine and threonine; glutamine and asparagine; and valine, leucine and isoleucine.

In addition, any polynucleotide may be further modified to increase stability *in vivo*. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends; the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages in the backbone; and/or the inclusion of nontraditional bases such as inosine, quosine and wybutosine, as well as acetyl-methyl-, thio- and other modified forms of adenine, cytidine, guanine, thymine and uridine.

Amino acid substitutions may further be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity and/or the amphipathic nature of the residues. For example, negatively charged amino acids include aspartic acid and glutamic acid; positively charged amino acids include lysine and arginine; and amino acids with uncharged polar head groups having similar hydrophilicity values include leucine, isoleucine and valine; glycine and alanine; asparagine and glutamine; and serine, threonine, phenylalanine and tyrosine. Other groups of amino acids that may represent conservative changes include: (1) ala, pro, gly, glu, asp, gln, asn, ser, thr; (2) cys, ser, tyr, thr; (3) val, ile, leu, met, ala, phe; (4) lys, arg, his; and (5) phe, tyr, trp, his. A variant may also, or alternatively, contain nonconservative changes. In a preferred embodiment, variant polypeptides differ from a native sequence by substitution, deletion or addition of five amino acids or fewer. Variants may also (or alternatively) be modified by, for example, the deletion or addition of amino acids that have minimal influence on the immunogenicity, secondary structure and hydrophobic nature of the polypeptide.

As noted above, polypeptides may comprise a signal (or leader) sequence at the N-terminal end of the protein, which co-translationally or post-translationally

WO 01/90152

PCT/US01/16776

18

directs transfer of the protein. The polypeptide may also be conjugated to a linker or other sequence for ease of synthesis, purification or identification of the polypeptide (e.g., poly-His), or to enhance binding of the polypeptide to a solid support. For example, a polypeptide may be conjugated to an immunoglobulin Fc region.

5 When comparing polypeptide sequences, two sequences are said to be "identical" if the sequence of amino acids in the two sequences is the same when aligned for maximum correspondence, as described below. Comparisons between two sequences are typically performed by comparing the sequences over a comparison window to identify and compare local regions of sequence similarity. A "comparison
10 window" as used herein, refers to a segment of at least about 20 contiguous positions, usually 30 to about 75, 40 to about 50, in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned.

Optimal alignment of sequences for comparison may be conducted using
15 the Megalign program in the Lasergene suite of bioinformatics software (DNASTAR, Inc., Madison, WI), using default parameters. This program embodies several alignment schemes described in the following references: Dayhoff, M.O. (1978) A model of evolutionary change in proteins - Matrices for detecting distant relationships. In Dayhoff, M.O. (ed.) Atlas of Protein Sequence and Structure, National Biomedical
20 Research Foundation, Washington DC Vol. 5, Suppl. 3, pp. 345-358; Hein J. (1990) Unified Approach to Alignment and Phylogenies pp. 626-645 *Methods in Enzymology* vol. 183, Academic Press, Inc., San Diego, CA; Higgins, D.G. and Sharp, P.M. (1989) *CABIOS* 5:151-153; Myers, E.W. and Muller W. (1988) *CABIOS* 4:11-17; Robinson, E.D. (1971) *Comb. Theor* 11:105; Santou, N. Nes, M. (1987) *Mol. Biol. Evol.* 4:406-
25 425; Sneath, P.H.A. and Sokal, R.R. (1973) *Numerical Taxonomy - the Principles and Practice of Numerical Taxonomy*, Freeman Press, San Francisco, CA; Wilbur, W.J. and Lipman, D.J. (1983) *Proc. Natl. Acad. Sci. USA* 80:726-730.

Alternatively, optimal alignment of sequences for comparison may be conducted by the local identity algorithm of Smith and Waterman (1981) *Add. Appl. Math* 2:482, by the identity alignment algorithm of Needleman and Wunsch (1970) *J.*
30

WO 01/90152

PCT/US01/16776

19

Mol. Biol. 48:443, by the search for similarity methods of Pearson and Lipman (1988) *Proc. Natl. Acad. Sci. USA* 85: 2444, by computerized implementations of these algorithms (GAP, BESTFIT, BLAST, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group (GCG), 575 Science Dr., Madison, WI),
5 or by inspection.

One preferred example of algorithms that are suitable for determining percent sequence identity and sequence similarity are the BLAST and BLAST 2.0 algorithms, which are described in Altschul et al. (1977) *Nucl. Acids Res.* 25:3389-3402 and Altschul et al. (1990) *J. Mol. Biol.* 215:403-410, respectively. BLAST and BLAST
10 2.0 can be used, for example with the parameters described herein, to determine percent sequence identity for the polynucleotides and polypeptides of the invention. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information. For amino acid sequences, a scoring matrix can be used to calculate the cumulative score. Extension of the word hits in each direction are halted
15 when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T and X determine the sensitivity and speed of the alignment.

20 In one preferred approach, the "percentage of sequence identity" is determined by comparing two optimally aligned sequences over a window of comparison of at least 20 positions, wherein the portion of the polypeptide sequence in the comparison window may comprise additions or deletions (*i.e.*, gaps) of 20 percent or less, usually 5 to 15 percent, or 10 to 12 percent, as compared to the reference
25 sequences (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical amino acid residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the reference sequence (*i.e.*, the window size) and multiplying the results by
30 100 to yield the percentage of sequence identity.

WO 01/90152

PCT/US01/16776

20

Within other illustrative embodiments, a polypeptide may be a fusion polypeptide that comprises multiple polypeptides as described herein, or that comprises at least one polypeptide as described herein and an unrelated sequence, such as a known tumor protein. A fusion partner may, for example, assist in providing T helper epitopes (an immunological fusion partner), preferably T helper epitopes recognized by humans, or may assist in expressing the protein (an expression enhancer) at higher yields than the native recombinant protein. Certain preferred fusion partners are both immunological and expression enhancing fusion partners. Other fusion partners may be selected so as to increase the solubility of the polypeptide or to enable the polypeptide to be targeted to desired intracellular compartments. Still further fusion partners include affinity tags, which facilitate purification of the polypeptide.

Fusion polypeptides may generally be prepared using standard techniques, including chemical conjugation. Preferably, a fusion polypeptide is expressed as a recombinant polypeptide, allowing the production of increased levels, relative to a non-fused polypeptide, in an expression system. Briefly, DNA sequences encoding the polypeptide components may be assembled separately, and ligated into an appropriate expression vector. The 3' end of the DNA sequence encoding one polypeptide component is ligated, with or without a peptide linker, to the 5' end of a DNA sequence encoding the second polypeptide component so that the reading frames of the sequences are in phase. This permits translation into a single fusion polypeptide that retains the biological activity of both component polypeptides.

A peptide linker sequence may be employed to separate the first and second polypeptide components by a distance sufficient to ensure that each polypeptide folds into its secondary and tertiary structures. Such a peptide linker sequence is incorporated into the fusion polypeptide using standard techniques well known in the art. Suitable peptide linker sequences may be chosen based on the following factors: (1) their ability to adopt a flexible extended conformation; (2) their inability to adopt a secondary structure that could interact with functional epitopes on the first and second polypeptides; and (3) the lack of hydrophobic or charged residues that might react with the polypeptide functional epitopes. Preferred peptide linker sequences contain Gly,

Asn and Ser residues. Other near neutral amino acids, such as Thr and Ala may also be used in the linker sequence. Amino acid sequences which may be usefully employed as linkers include those disclosed in Maratea et al., *Gene* 40:39-46, 1985; Murphy et al., *Proc. Natl. Acad. Sci. USA* 83:8258-8262, 1986; U.S. Patent No. 4,935,233 and U.S.

- 5 Patent No. 4,751,180. The linker sequence may generally be from 1 to about 50 amino acids in length. Linker sequences are not required when the first and second polypeptides have non-essential N-terminal amino acid regions that can be used to separate the functional domains and prevent steric interference.

- The ligated DNA sequences are operably linked to suitable
10 transcriptional or translational regulatory elements. The regulatory elements responsible for expression of DNA are located only 5' to the DNA sequence encoding the first polypeptides. Similarly, stop codons required to end translation and transcription termination signals are only present 3' to the DNA sequence encoding the second polypeptide.

- 15 The fusion polypeptide can comprise a polypeptide as described herein together with an unrelated immunogenic protein, such as an immunogenic protein capable of eliciting a recall response. Examples of such proteins include tetanus, tuberculosis and hepatitis proteins (see, for example, Stoute et al. *New Engl. J. Med.*, 336:86-91, 1997).

- 20 In one preferred embodiment, the immunological fusion partner is derived from a *Mycobacterium* sp., such as a *Mycobacterium tuberculosis*-derived Ra12 fragment. Ra12 compositions and methods for their use in enhancing the expression and/or immunogenicity of heterologous polynucleotide/polypeptide sequences is described in U.S. Patent Application 60/158,585, the disclosure of which is
25 incorporated herein by reference in its entirety. Briefly, Ra12 refers to a polynucleotide region that is a subsequence of a *Mycobacterium tuberculosis* MTB32A nucleic acid. MTB32A is a serine protease of 32 KD molecular weight encoded by a gene in virulent and avirulent strains of *M. tuberculosis*. The nucleotide sequence and amino acid sequence of MTB32A have been described (for example, U.S. Patent Application
30 60/158,585; see also, Skeiky et al., *Infection and Immunity*. (1999) 67:3998-4007,

WO 01/90152

PCT/US01/16776

22

incorporated herein by reference). C-terminal fragments of the MTB32A coding sequence express at high levels and remain as a soluble polypeptides throughout the purification process. Moreover, Ra12 may enhance the immunogenicity of heterologous immunogenic polypeptides with which it is fused. One preferred Ra12 fusion polypeptide comprises a 14 KD C-terminal fragment corresponding to amino acid residues 192 to 323 of MTB32A. Other preferred Ra12 polynucleotides generally comprise at least about 15 consecutive nucleotides, at least about 30 nucleotides, at least about 60 nucleotides, at least about 100 nucleotides, at least about 200 nucleotides, or at least about 300 nucleotides that encode a portion of a Ra12 polypeptide. Ra12 polynucleotides may comprise a native sequence (*i.e.*, an endogenous sequence that encodes a Ra12 polypeptide or a portion thereof) or may comprise a variant of such a sequence. Ra12 polynucleotide variants may contain one or more substitutions, additions, deletions and/or insertions such that the biological activity of the encoded fusion polypeptide is not substantially diminished, relative to a fusion polypeptide comprising a native Ra12 polypeptide. Variants preferably exhibit at least about 70% identity, more preferably at least about 80% identity and most preferably at least about 90% identity to a polynucleotide sequence that encodes a native Ra12 polypeptide or a portion thereof.

Within other preferred embodiments, an immunological fusion partner is derived from protein D, a surface protein of the gram-negative bacterium *Haemophilus influenza B* (WO 91/18926). Preferably, a protein D derivative comprises approximately the first third of the protein (*e.g.*, the first N-terminal 100-110 amino acids), and a protein D derivative may be lipidated. Within certain preferred embodiments, the first 109 residues of a Lipoprotein D fusion partner is included on the N-terminus to provide the polypeptide with additional exogenous T-cell epitopes and to increase the expression level in *E. coli* (thus functioning as an expression enhancer). The lipid tail ensures optimal presentation of the antigen to antigen presenting cells. Other fusion partners include the non-structural protein from influenzae virus, NS1 (hemagglutinin). Typically, the N-terminal 81 amino acids are used, although different fragments that include T-helper epitopes may be used.

WO 01/90152

PCT/US01/16776

23

In another embodiment, the immunological fusion partner is the protein known as LYTA, or a portion thereof (preferably a C-terminal portion). LYTA is derived from *Streptococcus pneumoniae*, which synthesizes an N-acetyl-L-alanine amidase known as amidase LYTA (encoded by the *LytA* gene; *Gene* 43:265-292, 1986).

5 LYTA is an autolysin that specifically degrades certain bonds in the peptidoglycan backbone. The C-terminal domain of the LYTA protein is responsible for the affinity to the choline or to some choline analogues such as DEAE. This property has been exploited for the development of *E. coli* C-LYTA expressing plasmids useful for expression of fusion proteins. Purification of hybrid proteins containing the C-LYTA
10 fragment at the amino terminus has been described (see *Biotechnology* 10:795-798, 1992). Within a preferred embodiment, a repeat portion of LYTA may be incorporated into a fusion polypeptide. A repeat portion is found in the C-terminal region starting at residue 178. A particularly preferred repeat portion incorporates residues 188-305.

Yet another illustrative embodiment involves fusion polypeptides, and
15 the polynucleotides encoding them, wherein the fusion partner comprises a targeting signal capable of directing a polypeptide to the endosomal/lysosomal compartment, as described in U.S. Patent No. 5,633,234. An immunogenic polypeptide of the invention, when fused with this targeting signal, will associate more efficiently with MHC class II molecules and thereby provide enhanced in vivo stimulation of CD4⁺ T-cells specific
20 for the polypeptide.

Polypeptides of the invention are prepared using any of a variety of well known synthetic and/or recombinant techniques, the latter of which are further described below. Polypeptides, portions and other variants generally less than about 150 amino acids can be generated by synthetic means, using techniques well known to
25 those of ordinary skill in the art. In one illustrative example, such polypeptides are synthesized using any of the commercially available solid-phase techniques, such as the Merrifield solid-phase synthesis method, where amino acids are sequentially added to a growing amino acid chain. See Merrifield, *J. Am. Chem. Soc.* 85:2149-2146, 1963. Equipment for automated synthesis of polypeptides is commercially available from

suppliers such as Perkin Elmer/Applied BioSystems Division (Foster City, CA), and may be operated according to the manufacturer's instructions.

In general, polypeptide compositions (including fusion polypeptides) of the invention are isolated. An "isolated" polypeptide is one that is removed from its original environment. For example, a naturally-occurring protein or polypeptide is isolated if it is separated from some or all of the coexisting materials in the natural system. Preferably, such polypeptides are also purified, *e.g.*, are at least about 90% pure, more preferably at least about 95% pure and most preferably at least about 99% pure.

10 Polynucleotide Compositions

The present invention, in other aspects, provides polynucleotide compositions. The terms "DNA" and "polynucleotide" are used essentially interchangeably herein to refer to a DNA molecule that has been isolated free of total genomic DNA of a particular species. "Isolated," as used herein, means that a polynucleotide is substantially away from other coding sequences, and that the DNA molecule does not contain large portions of unrelated coding DNA, such as large chromosomal fragments or other functional genes or polypeptide coding regions. Of course, this refers to the DNA molecule as originally isolated, and does not exclude genes or coding regions later added to the segment by the hand of man.

As will be understood by those skilled in the art, the polynucleotide compositions of this invention can include genomic sequences, extra-genomic and plasmid-encoded sequences and smaller engineered gene segments that express, or may be adapted to express, proteins, polypeptides, peptides and the like. Such segments may be naturally isolated, or modified synthetically by the hand of man.

As will be also recognized by the skilled artisan, polynucleotides of the invention may be single-stranded (coding or antisense) or double-stranded, and may be DNA (genomic, cDNA or synthetic) or RNA molecules. RNA molecules may include HnRNA molecules, which contain introns and correspond to a DNA molecule in a one-to-one manner, and mRNA molecules, which do not contain introns. Additional coding

WO 01/90152

PCT/US01/16776

25

or non-coding sequences may, but need not, be present within a polynucleotide of the present invention, and a polynucleotide may, but need not, be linked to other molecules and/or support materials.

Polynucleotides may comprise a native sequence (*i.e.*, an endogenous
5 sequence that encodes a polypeptide/protein of the invention or a portion thereof) or may comprise a sequence that encodes a variant or derivative, preferably and immunogenic variant or derivative, of such a sequence.

Therefore, according to another aspect of the present invention, polynucleotide compositions are provided that comprise some or all of a polynucleotide
10 sequence set forth in any one of SEQ ID NOs: 1, 3-86, 142-298, 301-303, 307, 313, 314, 316, 317, 325 and 327-330, complements of a polynucleotide sequence set forth in any one of SEQ ID NOs: 1, 3-86, 142-298, 301-303, 307, 313, 314, 316, 317, 325 and 327-330, and degenerate variants of a polynucleotide sequence set forth in any one of SEQ ID NOs: 1, 3-86, 142-298, 301-303, 307, 313, 314, 316, 317, 325 and 327-330. In
15 certain preferred embodiments, the polynucleotide sequences set forth herein encode immunogenic polypeptides, as described above.

In other related embodiments, the present invention provides polynucleotide variants having substantial identity to the sequences disclosed herein in SEQ ID NOs: 1, 3-86, 142-298, 301-303, 307, 313, 314, 316, 317, 325 and 327-330, for
20 example those comprising at least 70% sequence identity, preferably at least 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% or higher, sequence identity compared to a polynucleotide sequence of this invention using the methods described herein, (*e.g.*, BLAST analysis using standard parameters, as described below). One skilled in this art will recognize that these values can be appropriately adjusted to determine
25 corresponding identity of proteins encoded by two nucleotide sequences by taking into account codon degeneracy, amino acid similarity, reading frame positioning and the like.

Typically, polynucleotide variants will contain one or more substitutions, additions, deletions and/or insertions, preferably such that the immunogenicity of the
30 polypeptide encoded by the variant polynucleotide is not substantially diminished

WO 01/90152

PCT/US01/6776

26

relative to a polypeptide encoded by a polynucleotide sequence specifically set forth herein). The term "variants" should also be understood to encompass homologous genes of xenogenic origin.

In additional embodiments, the present invention provides
5 polynucleotide fragments comprising various lengths of contiguous stretches of sequence identical to or complementary to one or more of the sequences disclosed herein. For example, polynucleotides are provided by this invention that comprise at least about 10, 15, 20, 30, 40, 50, 75, 100, 150, 200, 300, 400, 500 or 1000 or more
10 contiguous nucleotides of one or more of the sequences disclosed herein as well as all intermediate lengths there between. It will be readily understood that "intermediate lengths", in this context, means any length between the quoted values, such as 16, 17, 18, 19, *etc.*; 21, 22, 23, *etc.*; 30, 31, 32, *etc.*; 50, 51, 52, 53, *etc.*; 100, 101, 102, 103, *etc.*; 150, 151, 152, 153, *etc.*; including all integers through 200-500; 500-1,000, and the like.

15 In another embodiment of the invention, polynucleotide compositions are provided that are capable of hybridizing under moderate to high stringency conditions to a polynucleotide sequence provided herein, or a fragment thereof, or a complementary sequence thereof. Hybridization techniques are well known in the art of molecular biology. For purposes of illustration, suitable moderately stringent conditions for
20 testing the hybridization of a polynucleotide of this invention with other polynucleotides include prewashing in a solution of 5 X SSC, 0.5% SDS, 1.0 mM EDTA (pH 8.0); hybridizing at 50°C-60°C, 5 X SSC, overnight; followed by washing twice at 65°C for 20 minutes with each of 2X, 0.5X and 0.2X SSC containing 0.1% SDS. One skilled in the art will understand that the stringency of hybridization can be readily manipulated,
25 such as by altering the salt content of the hybridization solution and/or the temperature at which the hybridization is performed. For example, in another embodiment, suitable highly stringent hybridization conditions include those described above, with the exception that the temperature of hybridization is increased, *e.g.*, to 60-65°C or 65-70°C.

WO 01/90152

PCT/US01/16776

27

In certain preferred embodiments, the polynucleotides described above, e.g., polynucleotide variants, fragments and hybridizing sequences, encode polypeptides that are immunologically cross-reactive with a polypeptide sequence specifically set forth herein. In other preferred embodiments, such polynucleotides encode polypeptides that have a level of immunogenic activity of at least about 50%, preferably at least about 70%, and more preferably at least about 90% of that for a polypeptide sequence specifically set forth herein.

The polynucleotides of the present invention, or fragments thereof, regardless of the length of the coding sequence itself, may be combined with other DNA sequences, such as promoters, polyadenylation signals, additional restriction enzyme sites, multiple cloning sites, other coding segments, and the like, such that their overall length may vary considerably. It is therefore contemplated that a nucleic acid fragment of almost any length may be employed, with the total length preferably being limited by the ease of preparation and use in the intended recombinant DNA protocol. For example, illustrative polynucleotide segments with total lengths of about 10,000, about 5000, about 3000, about 2,000, about 1,000, about 500, about 200, about 100, about 50 base pairs in length, and the like, (including all intermediate lengths) are contemplated to be useful in many implementations of this invention.

When comparing polynucleotide sequences, two sequences are said to be "identical" if the sequence of nucleotides in the two sequences is the same when aligned for maximum correspondence, as described below. Comparisons between two sequences are typically performed by comparing the sequences over a comparison window to identify and compare local regions of sequence similarity. A "comparison window" as used herein, refers to a segment of at least about 20 contiguous positions, usually 30 to about 75, 40 to about 50, in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned.

Optimal alignment of sequences for comparison may be conducted using the Megalign program in the Lasergene suite of bioinformatics software (DNASTAR, Inc., Madison, WI), using default parameters. This program embodies several

WO 01/90152

PCT/US01/16776

28

- alignment schemes described in the following references: Dayhoff, M.O. (1978) A model of evolutionary change in proteins. Matrices for detecting distant relationships. In Dayhoff, M.O. (ed.) *Atlas of Protein Sequence and Structure*, National Biomedical Research Foundation, Washington DC. Vol. 5, Suppl. 3, pp. 345-358; Hein J. (1990)
- 5 Unified Approach to Alignment and Phylogenies pp. 626-645 *Methods in Enzymology* vol. 183, Academic Press, Inc., San Diego, CA; Higgins, D.G. and Sharp, P.M. (1989) *CABIOS* 5:151-153; Myers, E.W. and Muller W. (1988) *CABIOS* 4:11-17; Robinson, E.D. (1971) *Comb. Theor* 11:105; Santou, N. Nes, M. (1987) *Mol. Biol. Evol.* 4:406-425; Sneath, P.H.A. and Sokal, R.R. (1973) *Numerical Taxonomy - the Principles and*
- 10 *Practice of Numerical Taxonomy*, Freeman Press, San Francisco, CA; Wilbur, W.J. and Lipman, D.J. (1983) *Proc. Natl. Acad. Sci. USA* 80:726-730.

Alternatively, optimal alignment of sequences for comparison may be conducted by the local identity algorithm of Smith and Waterman (1981) *Adv. APL. Math* 2:482, by the identity alignment algorithm of Needleman and Wunsch (1970) *J.*

15 *Mol. Biol.* 48:443, by the search for similarity methods of Pearson and Lipman (1988) *Proc. Natl. Acad. Sci. USA* 85: 2444, by computerized implementations of these algorithms (GAP, BESTFIT, BLAST, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group (GCG), 575 Science Dr., Madison, WI), or by inspection.

- 20 One preferred example of algorithms that are suitable for determining percent sequence identity and sequence similarity are the BLAST and BLAST 2.0 algorithms, which are described in Altschul et al. (1990) *Nucl. Acids Res.* 18:3500-3502 and Altschul et al. (1990) *J. Mol. Biol.* 215:403-410, respectively. BLAST and BLAST 2.0 can be used, for example with the parameters described herein, to determine percent
- 25 sequence identity for the polynucleotides of the invention. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information. In one illustrative example, cumulative scores can be calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always >0) and N (penalty score for mismatching residues; always <0). Extension of
- 30 the word hits in each direction are halted when: the cumulative alignment score falls off

WO 01/90152

PCT/US01/16776

29

by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, and expectation (E) of 10, and the BLOSUM62 scoring matrix (see Henikoff and Henikoff (1989) *Proc. Natl. Acad. Sci. USA* 89:10915) alignments, (B) of 50, expectation (E) of 10, M=5, N=-4 and a comparison of both strands.

Preferably, the "percentage of sequence identity" is determined by comparing two optimally aligned sequences over a window of comparison of at least 20 positions, wherein the portion of the polynucleotide sequence in the comparison window may comprise additions or deletions (*i.e.*, gaps) of 20 percent or less, usually 5 to 15 percent, or 10 to 12 percent, as compared to the reference sequences (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical nucleic acid bases occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the reference sequence (*i.e.*, the window size) and multiplying the results by 100 to yield the percentage of sequence identity.

It will be appreciated by those of ordinary skill in the art that, as a result of the degeneracy of the genetic code, there are many nucleotide sequences that encode a polypeptide as described herein. Some of these polynucleotides bear minimal homology to the nucleotide sequence of any native gene. Nonetheless, polynucleotides that vary due to differences in codon usage are specifically contemplated by the present invention. Further, alleles of the genes comprising the polynucleotide sequences provided herein are within the scope of the present invention. Alleles are endogenous genes that are altered as a result of one or more mutations, such as deletions, additions and/or substitutions of nucleotides. The resulting mRNA and protein may, but need not, have an altered structure or function. Alleles may be identified using standard techniques (such as hybridization, amplification and/or database sequence comparison).

WO 01/90152

PCT/US01/16776

30

Therefore, in another embodiment of the invention, a mutagenesis approach, such as site-specific mutagenesis, is employed for the preparation of immunogenic variants and/or derivatives of the polypeptides described herein. By this approach, specific modifications in a polypeptide sequence can be made through mutagenesis of the underlying polynucleotides that encode them. These techniques provides a straightforward approach to prepare and test sequence variants, for example, incorporating one or more of the foregoing considerations, by introducing one or more nucleotide sequence changes into the polynucleotide.

Site-specific mutagenesis allows the production of mutants through the use of specific oligonucleotide sequences which encode the DNA sequence of the desired mutation, as well as a sufficient number of adjacent nucleotides, to provide a primer sequence of sufficient size and sequence complexity to form a stable duplex on both sides of the deletion junction being traversed. Mutations may be employed in a selected polynucleotide sequence to improve, alter, decrease, modify, or otherwise change the properties of the polynucleotide itself, and/or alter the properties, activity, composition, stability, or primary sequence of the encoded polypeptide.

In certain embodiments of the present invention, the inventors contemplate the mutagenesis of the disclosed polynucleotide sequences to alter one or more properties of the encoded polypeptide, such as the immunogenicity of a polypeptide vaccine. The techniques of site-specific mutagenesis are well-known in the art, and are widely used to create variants of both polypeptides and polynucleotides. For example, site-specific mutagenesis is often used to alter a specific portion of a DNA molecule. In such embodiments, a primer comprising typically about 14 to about 25 nucleotides or so in length is employed, with about 5 to about 10 residues on both sides of the junction of the sequence being altered.

As will be appreciated by those of skill in the art, site-specific mutagenesis techniques have often employed a phage vector that exists in both a single stranded and double stranded form. Typical vectors useful in site-directed mutagenesis include vectors such as the M13 phage. These phage are readily commercially-available and their use is generally well-known to those skilled in the art. Double-stranded

plasmids are also routinely employed in site directed mutagenesis that eliminates the step of transferring the gene of interest from a plasmid to a phage.

In general, site-directed mutagenesis in accordance herewith is performed by first obtaining a single-stranded vector or melting apart of two strands of a double-stranded vector that includes within its sequence a DNA sequence that encodes the desired peptide. An oligonucleotide primer bearing the desired mutated sequence is prepared, generally synthetically. This primer is then annealed with the single-stranded vector, and subjected to DNA polymerizing enzymes such as *E. coli* polymerase I Klenow fragment, in order to complete the synthesis of the mutation-bearing strand. Thus, a heteroduplex is formed wherein one strand encodes the original non-mutated sequence and the second strand bears the desired mutation. This heteroduplex vector is then used to transform appropriate cells, such as *E. coli* cells, and clones are selected which include recombinant vectors bearing the mutated sequence arrangement.

The preparation of sequence variants of the selected peptide-encoding DNA segments using site-directed mutagenesis provides a means of producing potentially useful species and is not meant to be limiting as there are other ways in which sequence variants of peptides and the DNA sequences encoding them may be obtained. For example, recombinant vectors encoding the desired peptide sequence may be treated with mutagenic agents, such as hydroxylamine, to obtain sequence variants. Specific details regarding these methods and protocols are found in the teachings of Maloy *et al.*, 1994; Segal, 1976; Prokop and Bajpai, 1991; Kuby, 1994; and Maniatis *et al.*, 1982, each incorporated herein by reference, for that purpose.

As used herein, the term "oligonucleotide directed mutagenesis procedure" refers to template-dependent processes and vector-mediated propagation which result in an increase in the concentration of a specific nucleic acid molecule relative to its initial concentration, or in an increase in the concentration of a detectable signal, such as amplification. As used herein, the term "oligonucleotide directed mutagenesis procedure" is intended to refer to a process that involves the template-dependent extension of a primer molecule. The term template dependent process refers to nucleic acid synthesis of an RNA or a DNA molecule wherein the

sequence of the newly synthesized strand of nucleic acid is dictated by the well-known rules of complementary base pairing (see, for example, Watson, 1987). Typically, vector mediated methodologies involve the introduction of the nucleic acid fragment into a DNA or RNA vector, the clonal amplification of the vector, and the recovery of the amplified nucleic acid fragment. Examples of such methodologies are provided by U. S. Patent No. 4,237,224, specifically incorporated herein by reference in its entirety.

In another approach for the production of polypeptide variants of the present invention, recursive sequence recombination, as described in U.S. Patent No. 5,837,458, may be employed. In this approach, iterative cycles of recombination and screening or selection are performed to "evolve" individual polynucleotide variants of the invention having, for example, enhanced immunogenic activity.

In other embodiments of the present invention, the polynucleotide sequences provided herein can be advantageously used as probes or primers for nucleic acid hybridization. As such, it is contemplated that nucleic acid segments that comprise a sequence region of at least about 15 nucleotide long contiguous sequence that has the same sequence as, or is complementary to, a 15 nucleotide long contiguous sequence disclosed herein will find particular utility. Longer contiguous identical or complementary sequences, *e.g.*, those of about 20, 30, 40, 50, 100, 200, 500, 1000 (including all intermediate lengths) and even up to full length sequences will also be of use in certain embodiments.

The ability of such nucleic acid probes to specifically hybridize to a sequence of interest will enable them to be of use in detecting the presence of complementary sequences in a given sample. However, other uses are also envisioned, such as the use of the sequence information for the preparation of mutant species primers, or primers for use in preparing other genetic constructions.

Polynucleotide molecules having sequence regions consisting of contiguous nucleotide stretches of 10-14, 15-20, 30, 50, or even of 100-200 nucleotides or so (including intermediate lengths as well), identical or complementary to a polynucleotide sequence disclosed herein, are particularly contemplated as hybridization probes for use in, *e.g.*, Southern and Northern blotting. This would allow a gene

product, or fragment thereof, to be analyzed, both in diverse cell types and also in various bacterial cells. The total size of fragment, as well as the size of the complementary stretch(es), will ultimately depend on the intended use or application of the particular nucleic acid segment. Smaller fragments will generally find use in hybridization embodiments, wherein the length of the contiguous complementary region may be varied, such as between about 15 and about 100 nucleotides, but larger contiguous complementarity stretches may be used, according to the length complementary sequences one wishes to detect.

The use of a hybridization probe of about 15-25 nucleotides in length allows the formation of a duplex molecule that is both stable and selective. Molecules having contiguous complementary sequences over stretches greater than 15 bases in length are generally preferred, though, in order to increase stability and selectivity of the hybrid, and thereby improve the quality and degree of specific hybrid molecules obtained. One will generally prefer to design nucleic acid molecules having complementary stretches of 15 to 25 contiguous nucleotides, or even longer where desired.

Hybridization probes may be selected from any portion of any of the sequences disclosed herein. All that is required is to review the sequences set forth herein, or to any continuous portion of the sequences, from about 15-25 nucleotides in length up to and including the full length sequence, that one wishes to utilize as a probe or primer. The choice of probe and primer sequences may be governed by various factors. For example, one may wish to employ primers from towards the termini of the total sequence.

Small polynucleotide segments or fragments may be readily prepared by, for example, directly synthesizing the fragment by chemical means, as is commonly practiced using an automated oligonucleotide synthesizer. Also, fragments may be obtained by application of nucleic acid reproduction technology, such as the PCRTM technology of U. S. Patent 4,683,202 (incorporated herein by reference), by introducing selected sequences into recombinant vectors for recombinant production, and by other

recombinant DNA techniques generally known to those of skill in the art of molecular biology.

The nucleotide sequences of the invention may be used for their ability to selectively form duplex molecules with complementary stretches of the entire gene or gene fragments of interest. Depending on the application envisioned, one will typically desire to employ varying conditions of hybridization to achieve varying degrees of selectivity of probe towards target sequence. For applications requiring high selectivity, one will typically desire to employ relatively stringent conditions to form the hybrids, *e.g.*, one will select relatively low salt and/or high temperature conditions, such as provided by a salt concentration of from about 0.02 M to about 0.15 M salt at temperatures of from about 50°C to about 70°C. Such selective conditions tolerate little, if any, mismatch between the probe and the template or target strand, and would be particularly suitable for isolating related sequences.

Of course, for some applications, for example, where one desires to prepare mutants employing a mutant primer strand hybridized to an underlying template, less stringent (reduced stringency) hybridization conditions will typically be needed in order to allow formation of the heteroduplex. In these circumstances, one may desire to employ salt conditions such as those of from about 0.15 M to about 0.9 M salt, at temperatures ranging from about 20°C to about 55°C. Cross-hybridizing species can thereby be readily identified as positively hybridizing signals with respect to control hybridizations. In any case, it is generally appreciated that conditions can be rendered more stringent by the addition of increasing amounts of formamide, which serves to destabilize the hybrid duplex in the same manner as increased temperature. Thus, hybridization conditions can be readily manipulated, and thus will generally be a method of choice depending on the desired results.

According to another embodiment of the present invention, polynucleotide compositions comprising antisense oligonucleotides are provided. Antisense oligonucleotides have been demonstrated to be effective and targeted inhibitors of protein synthesis, and, consequently, provide a therapeutic approach by which a disease can be treated by inhibiting the synthesis of proteins that contribute to

WO 01/90152

PCT/US01/16776

35

the disease. The efficacy of antisense oligonucleotides for inhibiting protein synthesis is well established. For example, the synthesis of polygalacturonase and the muscarinic type 2 acetylcholine receptor are inhibited by antisense oligonucleotides directed to their respective mRNA sequences (U. S. Patent 5,739,119 and U. S. Patent 5,759,829).

5 Further, examples of antisense inhibition have been demonstrated with the nuclear protein cyclin, the multiple drug resistance gene (MDG1), ICAM-1, E-selectin, STK-1, striatal GABA_A receptor and human EGF (Jaskulski *et al.*, Science, 1988 Jun 10;240(4858):1544-6; Vasanthakumar and Ahmed, Cancer Commun. 1989;1(4):225-32; Peris *et al.*, Brain Res Mol Brain Res. 1998 Jun 15;57(2):310-20; U. S. Patent 10 5,801,154; U.S. Patent 5,789,573; U. S. Patent 5,718,709 and U.S. Patent 5,610,288). Antisense constructs have also been described that inhibit and can be used to treat a variety of abnormal cellular proliferations, *e.g.* cancer (U. S. Patent 5,747,470; U. S. Patent 5,591,317 and U. S. Patent 5,783,683).

Therefore, in certain embodiments, the present invention provides

15 oligonucleotide sequences that comprise all, or a portion of, any sequence that is capable of specifically binding to polynucleotide sequence described herein, or a complement thereof. In one embodiment, the antisense oligonucleotides comprise DNA or derivatives thereof. In another embodiment, the oligonucleotides comprise RNA or derivatives thereof. In a third embodiment, the oligonucleotides are modified DNAs

20 comprising a phosphorothioated modified backbone. In a fourth embodiment, the oligonucleotide sequences comprise peptide nucleic acids or derivatives thereof. In each case, preferred compositions comprise a sequence region that is complementary, and more preferably substantially-complementary, and even more preferably, completely complementary to one or more portions of polynucleotides disclosed herein.

25 Selection of antisense compositions specific for a given gene sequence is based upon analysis of the chosen target sequence and determination of secondary structure, T_m, binding energy, and relative stability. Antisense compositions may be selected based upon their relative inability to form dimers, hairpins, or other secondary structures that would reduce or prohibit specific binding to the target mRNA in a host cell. Highly

30 preferred target regions of the mRNA, are those which are at or near the AUG

translation initiation codon, and those sequences which are substantially complementary to 5' regions of the mRNA. These secondary structure analyses and target site selection considerations can be performed, for example, using v.4 of the OLIGO primer analysis software and/or the BLASTN 2.0.5 algorithm software (Altschul *et al.*, Nucleic Acids Res. 1997, 25(17):3389-402).

The use of an antisense delivery method employing a short peptide vector, termed MPG (27 residues), is also contemplated. The MPG peptide contains a hydrophobic domain derived from the fusion sequence of HIV gp41 and a hydrophilic domain from the nuclear localization sequence of SV40 T-antigen (Morris *et al.*, Nucleic Acids Res. 1997 Jul 15;25(14):2730-6). It has been demonstrated that several molecules of the MPG peptide coat the antisense oligonucleotides and can be delivered into cultured mammalian cells in less than 1 hour with relatively high efficiency (90%). Further, the interaction with MPG strongly increases both the stability of the oligonucleotide to nuclease and the ability to cross the plasma membrane.

According to another embodiment of the invention, the polynucleotide compositions described herein are used in the design and preparation of ribozyme molecules for inhibiting expression of the tumor polypeptides and proteins of the present invention in tumor cells. Ribozymes are RNA-protein complexes that cleave nucleic acids in a site-specific fashion. Ribozymes have specific catalytic domains that possess endonuclease activity (Kim and Cech, Proc Natl Acad Sci U S A. 1987 Dec;84(24):8788-92; Forster and Symons, Cell. 1987 Apr 24;49(2):211-20). For example, a large number of ribozymes accelerate phosphoester transfer reactions with a high degree of specificity, often cleaving only one of several phosphoesters in an oligonucleotide substrate (Cech *et al.*, Cell. 1981 Dec;27(3 Pt 2):487-96; Michel and Westhof, J Mol Biol. 1990 Dec 5;216(3):585-610; Reinhold-Hurek and Shub, Nature. 1992 May 14;357(6374):173-6). This specificity has been attributed to the requirement that the substrate bind via specific base-pairing interactions to the internal guide sequence ("IGS") of the ribozyme prior to chemical reaction.

Six basic varieties of naturally-occurring enzymatic RNAs are known presently. Each can catalyze the hydrolysis of RNA phosphodiester bonds *in trans* (and

WO 01/90152

PCT/US01/16776

37

thus can cleave other RNA molecules) under physiological conditions. In general, enzymatic nucleic acids act by first binding to a target RNA. Such binding occurs through the target binding portion of a enzymatic nucleic acid which is held in close proximity to an enzymatic portion of the molecule that acts to cleave the target RNA.

5 Thus, the enzymatic nucleic acid first recognizes and then binds a target RNA through complementary base-pairing, and once bound to the correct site, acts enzymatically to cut the target RNA. Strategic cleavage of such a target RNA will destroy its ability to direct synthesis of an encoded protein. After an enzymatic nucleic acid has bound and cleaved its RNA target, it is released from that RNA to search for another target and can
10 repeatedly bind and cleave new targets.

The enzymatic nature of a ribozyme is advantageous over many technologies, such as antisense technology (where a nucleic acid molecule simply binds to a nucleic acid target to block its translation) since the concentration of ribozyme necessary to affect a therapeutic treatment is lower than that of an antisense
15 oligonucleotide. This advantage reflects the ability of the ribozyme to act enzymatically. Thus, a single ribozyme molecule is able to cleave many molecules of target RNA. In addition, the ribozyme is a highly specific inhibitor, with the specificity of inhibition depending not only on the base pairing mechanism of binding to the target RNA, but also on the mechanism of target RNA cleavage. Single mismatches, or base-
20 substitutions, near the site of cleavage can completely eliminate catalytic activity of a ribozyme. Similar mismatches in antisense molecules do not prevent their action (Woolf *et al.*, Proc Natl Acad Sci U S A. 1992 Aug 15;89(16):7303-9). Thus, the specificity of action of a ribozyme is greater than that of an antisense oligonucleotide binding the same RNA site.

25 The enzymatic nucleic acid molecule may be formed in a hammerhead, hairpin, a hepatitis δ virus, group I intron or RNaseP RNA (in association with an RNA guide sequence) or Neurospora VS RNA motif. Examples of hammerhead motifs are described by Rossi *et al.* Nucleic Acids Res. 1992 Sep 11;20(17):4559-65. Examples of hairpin motifs are described by Hampel *et al.* (Eur. Pat. Appl. Publ. No. EP 0360257),
30 Hampel and Tritz, Biochemistry 1989 Jun 13;28(12):4929-33; Hampel *et al.*, Nucleic

WO 01/90152

PCT/US01/16776

38

Acids Res. 1990 Jan 25;18(2):299-304 and U. S. Patent 5,631,359. An example of the hepatitis δ virus motif is described by Perrotta and Been, Biochemistry. 1992 Dec 1;31(47):11843-52; an example of the RNaseP motif is described by Guerrier-Takada *et al.*, Cell. 1983 Dec;35(3 Pt 2):849-57; Neurospora VS RNA ribozyme motif is described by Collins (Saville and Collins, Cell. 1990 May 18;61(4):685-96; Saville and Collins, Proc Natl Acad Sci U S A. 1991 Oct 1;88(19):8826-30; Collins and Olive, Biochemistry. 1993 Mar 23;32(11):2795-9); and an example of the Group I intron is described in (U. S. Patent 4,987,071). All that is important in an enzymatic nucleic acid molecule of this invention is that it has a specific substrate binding site which is complementary to one or more of the target gene RNA regions, and that it have nucleotide sequences within or surrounding that substrate binding site which impart an RNA cleaving activity to the molecule. Thus the ribozyme constructs need not be limited to specific motifs mentioned herein.

Ribozymes may be designed as described in Int. Pat. Appl. Publ. No. WO 93/23569 and Int. Pat. Appl. Publ. No. WO 94/02595, each specifically incorporated herein by reference) and synthesized to be tested *in vitro* and *in vivo*, as described. Such ribozymes can also be optimized for delivery. While specific examples are provided, those in the art will recognize that equivalent RNA targets in other species can be utilized when necessary.

Ribozyme activity can be optimized by altering the length of the ribozyme binding arms, or chemically synthesizing ribozymes with modifications that prevent their degradation by serum ribonucleases (see *e.g.*, Int. Pat. Appl. Publ. No. WO 92/07065; Int. Pat. Appl. Publ. No. WO 93/15187; Int. Pat. Appl. Publ. No. WO 91/03162; Eur. Pat. Appl. Publ. No. 92110298.4; U. S. Patent 5,334,711; and Int. Pat. Appl. Publ. No. WO 94/13688, which describe various chemical modifications that can be made to the sugar moieties of enzymatic RNA molecules), modifications which enhance their efficacy in cells, and removal of stem II bases to shorten RNA synthesis times and reduce chemical requirements.

Sullivan *et al.* (Int. Pat. Appl. Publ. No. WO 94/02595) describes the general methods for delivery of enzymatic RNA molecules. Ribozymes may be

WO 01/90152

PCT/US01/16776

39

administered to cells by a variety of methods known to those familiar to the art, including, but not restricted to, encapsulation in liposomes, by iontophoresis, or by incorporation into other vehicles, such as hydrogels, cyclodextrins, biodegradable nanocapsules, and bioadhesive microspheres. For some indications, ribozymes may be directly delivered *ex vivo* to cells or tissues with or without the aforementioned vehicles. Alternatively, the RNA/vehicle combination may be locally delivered by direct inhalation, by direct injection or by use of a catheter, infusion pump or stent. Other routes of delivery include, but are not limited to, intravascular, intramuscular, subcutaneous or joint injection, aerosol inhalation, oral (tablet or pill form), topical, systemic, ocular, intraperitoneal and/or intrathecal delivery. More detailed descriptions of ribozyme delivery and administration are provided in Int. Pat. Appl. Publ. No. WO 94/02595 and Int. Pat. Appl. Publ. No. WO 93/23569, each specifically incorporated herein by reference.

Another means of accumulating high concentrations of a ribozyme(s) within cells is to incorporate the ribozyme-encoding sequences into a DNA expression vector. Transcription of the ribozyme sequences are driven from a promoter for eukaryotic RNA polymerase I (pol I), RNA polymerase II (pol II), or RNA polymerase III (pol III). Transcripts from pol II or pol III promoters will be expressed at high levels in all cells; the levels of a given pol II promoter in a given cell type will depend on the nature of the gene regulatory sequences (enhancers, silencers, *etc.*) present nearby. Prokaryotic RNA polymerase promoters may also be used, providing that the prokaryotic RNA polymerase enzyme is expressed in the appropriate cells. Ribozymes expressed from such promoters have been shown to function in mammalian cells. Such transcription units can be incorporated into a variety of vectors for introduction into mammalian cells, including but not restricted to, plasmid DNA vectors, viral DNA vectors (such as adenovirus or adeno-associated vectors), or viral RNA vectors (such as retroviral, semliki forest virus, sindbis virus vectors).

In another embodiment of the invention, peptide nucleic acids (PNAs) compositions are provided. PNA is a DNA mimic in which the nucleobases are attached to a pseudopeptide backbone (Good and Nielsen, Antisense Nucleic Acid Drug

Dev. 1997 7(4) 431-37). PNA is able to be utilized in a number methods that traditionally have used RNA or DNA. Often PNA sequences perform better in techniques than the corresponding RNA or DNA sequences and have utilities that are not inherent to RNA or DNA. A review of PNA including methods of making, characteristics of, and methods of using, is provided by Corey (*Trends Biotechnol* 1997 Jun;15(6):224-9). As such, in certain embodiments, one may prepare PNA sequences that are complementary to one or more portions of the ACE mRNA sequence, and such PNA compositions may be used to regulate, alter, decrease, or reduce the translation of ACE-specific mRNA, and thereby alter the level of ACE activity in a host cell to which such PNA compositions have been administered.

PNAs have 2-aminoethyl-glycine linkages replacing the normal phosphodiester backbone of DNA (Nielsen *et al.*, *Science* 1991 Dec 6;254(5037):1497-500; Ilanvey *et al.*, *Science*. 1992 Nov 27;258(5087):1481-5; Hyrup and Nielsen, *Bioorg Med Chem*. 1996 Jan;4(1):5-23). This chemistry has three important consequences: firstly, in contrast to DNA or phosphorothioate oligonucleotides, PNAs are neutral molecules; secondly, PNAs are achiral, which avoids the need to develop a stereoselective synthesis; and thirdly, PNA synthesis uses standard Boc or Fmoc protocols for solid-phase peptide synthesis, although other methods, including a modified Merrifield method, have been used.

PNA monomers or ready-made oligomers are commercially available from PerSeptive Biosystems (Framingham, MA). PNA syntheses by either Boc or Fmoc protocols are straightforward using manual or automated protocols (Norton *et al.*, *Bioorg Med Chem*. 1995 Apr;3(4):437-45). The manual protocol lends itself to the production of chemically modified PNAs or the simultaneous synthesis of families of closely related PNAs.

As with peptide synthesis, the success of a particular PNA synthesis will depend on the properties of the chosen sequence. For example, while in theory PNAs can incorporate any combination of nucleotide bases, the presence of adjacent purines can lead to deletions of one or more residues in the product. In expectation of this difficulty, it is suggested that, in producing PNAs with adjacent purines, one should

WO 01/90152

PCT/US01/16776

41

repeat the coupling of residues likely to be added inefficiently. This should be followed by the purification of PNAs by reverse-phase high-pressure liquid chromatography, providing yields and purity of product similar to those observed during the synthesis of peptides.

- 5 Modifications of PNAs for a given application may be accomplished by coupling amino acids during solid-phase synthesis or by attaching compounds that contain a carboxylic acid group to the exposed N-terminal amine. Alternatively, PNAs can be modified after synthesis by coupling to an introduced lysine or cysteine. The ease with which PNAs can be modified facilitates optimization for better solubility or
- 10 for specific functional requirements. Once synthesized, the identity of PNAs and their derivatives can be confirmed by mass spectrometry. Several studies have made and utilized modifications of PNAs (for example, Norton *et al.*, *Bioorg Med Chem.* 1995 Apr;3(4):437-45; Petersen *et al.*, *J Pept Sci.* 1995 May-Jun;1(3):175-83; Orum *et al.*, *Biotechniques.* 1995 Sep;19(3):472-80; Footer *et al.*, *Biochemistry.* 1996 Aug
- 15 20;35(33):10673-9; Griffith *et al.*, *Nucleic Acids Res.* 1995 Aug 11;23(15):3003-8; Partridge *et al.*, *Proc Natl Acad Sci U S A.* 1995 Jun 6;92(12):5592-6; Boffa *et al.*, *Proc Natl Acad Sci U S A.* 1995 Mar 14;92(6):1901-5; Gambacorti-Passerini *et al.*, *Blood.* 1996 Aug 15;88(4):1411-7; Anritage *et al.*, *Proc Natl Acad Sci U S A.* 1997 Nov 11;94(23):12320-5; Seeger *et al.*, *Biotechniques.* 1997 Sep;23(3):512-7). U.S.
- 20 Patent No. 5,700,922 discusses PNA-DNA-PNA chimeric molecules and their uses in diagnostics, modulating protein in organisms, and treatment of conditions susceptible to therapeutics.

- Methods of characterizing the antisense binding properties of PNAs are discussed in Rose (*Anal Chem.* 1993 Dec 15;65(24):3545-9) and Jensen *et al.*
- 25 (*Biochemistry.* 1997 Apr 22;36(16):5072-7). Rose uses capillary gel electrophoresis to determine binding of PNAs to their complementary oligonucleotide, measuring the relative binding kinetics and stoichiometry. Similar types of measurements were made by Jensen *et al.* using BLAcoreTM technology.

- Other applications of PNAs that have been described and will be
- 30 apparent to the skilled artisan include use in DNA strand invasion, antisense inhibition,

mutational analysis, enhancers of transcription, nucleic acid purification, isolation of transcriptionally active genes, blocking of transcription factor binding, genome cleavage, biosensors, *in situ* hybridization, and the like.

Polynucleotide Identification, Characterization and Expression

5 Polynucleotides compositions of the present invention may be identified, prepared and/or manipulated using any of a variety of well established techniques (see generally, Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratories, Cold Spring Harbor, NY, 1989, and other like references). For example, a polynucleotide may be identified by screening a microarray of cDNAs for
10 tumor-associated expression (*i.e.*, expression that is at least two fold greater in a tumor than in normal tissue, as determined using a representative assay provided herein). Such screens may be performed, for example, using the microarray technology of Affymetrix, Inc. (Santa Clara, CA) according to the manufacturer's instructions (and essentially as described by Schena et al., *Proc. Natl. Acad. Sci. USA* 93:10614-10619,
15 1996 and Heller et al., *Proc. Natl. Acad. Sci. USA* 94:2150-2155, 1997). Alternatively, polynucleotides may be amplified from cDNA prepared from cells expressing the proteins described herein, such as tumor cells.

Many template dependent processes are available to amplify a target sequences of interest present in a sample. One of the best known amplification methods
20 is the polymerase chain reaction (PCR™) which is described in detail in U.S. Patent Nos. 4,683,195, 4,683,202 and 4,800,159, each of which is incorporated herein by reference in its entirety. Briefly, in PCR™, two primer sequences are prepared which are complementary to regions on opposite complementary strands of the target sequence. An excess of deoxynucleoside triphosphates is added to a reaction mixture
25 along with a DNA polymerase (*e.g.*, *Taq* polymerase). If the target sequence is present in a sample, the primers will bind to the target and the polymerase will cause the primers to be extended along the target sequence by adding on nucleotides. By raising and lowering the temperature of the reaction mixture, the extended primers will dissociate from the target to form reaction products, excess primers will bind to the

WO 01/90152

PCT/US01/16776

43

target and to the reaction product and the process is repeated. Preferably reverse transcription and PCRTM amplification procedure may be performed in order to quantify the amount of mRNA amplified. Polymerase chain reaction methodologies are well known in the art.

5 Any of a number of other template dependent processes, many of which are variations of the PCRTM amplification technique, are readily known and available in the art. Illustratively, some such methods include the ligase chain reaction (referred to as LCR), described, for example, in Eur. Pat. Appl. Publ. No. 320,308 and U.S. Patent No. 4,883,750; Qbeta Replicase, described in PCT Intl. Pat. Appl. Publ. No. PCT/US87/00880; Strand Displacement Amplification (SDA) and Repair Chain Reaction (RCR). Still other amplification methods are described in Great Britain Pat. Appl. No. 2 202 328, and in PCT Intl. Pat. Appl. Publ. No. PCT/US89/01025. Other
10 nucleic acid amplification procedures include transcription-based amplification systems (TAS) (PCT Intl. Pat. Appl. Publ. No. WO 88/10315), including nucleic acid sequence based amplification (NASBA) and 3SR. Eur. Pat. Appl. Publ. No. 329,822 describes a nucleic acid amplification process involving cyclically synthesizing single-stranded RNA ("ssRNA"), ssDNA, and double-stranded DNA (dsDNA). PCT Intl. Pat. Appl. Publ. No. WO 89/06700 describes a nucleic acid sequence amplification scheme based on the hybridization of a promoter/primer sequence to a target single-stranded DNA
20 ("ssDNA") followed by transcription of many RNA copies of the sequence. Other amplification methods such as "RACE" (Frohman, 1990), and "one-sided PCR" (Ohara, 1989) are also well-known to those of skill in the art.

An amplified portion of a polynucleotide of the present invention may be used to isolate a full length gene from a suitable library (e.g., a tumor cDNA library) using well known techniques. Within such techniques, a library (cDNA or genomic) is
25 screened using one or more polynucleotide probes or primers suitable for amplification. Preferably, a library is size-selected to include larger molecules. Random primed libraries may also be preferred for identifying 5' and upstream regions of genes. Genomic libraries are preferred for obtaining introns and extending 5' sequences.

For hybridization techniques, a partial sequence may be labeled (e.g., by nick-translation or end-labeling with ^{32}P) using well known techniques. A bacterial or bacteriophage library is then generally screened by hybridizing filters containing denatured bacterial colonies (or lawns containing phage plaques) with the labeled probe (see Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratories, Cold Spring Harbor, NY, 1989). Hybridizing colonies or plaques are selected and expanded, and the DNA is isolated for further analysis. cDNA clones may be analyzed to determine the amount of additional sequence by, for example, PCR using a primer from the partial sequence and a primer from the vector. Restriction maps and partial sequences may be generated to identify one or more overlapping clones. The complete sequence may then be determined using standard techniques, which may involve generating a series of deletion clones. The resulting overlapping sequences can then be assembled into a single contiguous sequence. A full length cDNA molecule can be generated by ligating suitable fragments, using well known techniques.

Alternatively, amplification techniques, such as those described above, can be useful for obtaining a full length coding sequence from a partial cDNA sequence. One such amplification technique is inverse PCR (see Triglia et al., *Nucl. Acids Res.* 16:8186, 1988), which uses restriction enzymes to generate a fragment in the known region of the gene. The fragment is then circularized by intramolecular ligation and used as a template for PCR with divergent primers derived from the known region. Within an alternative approach, sequences adjacent to a partial sequence may be retrieved by amplification with a primer to a linker sequence and a primer specific to a known region. The amplified sequences are typically subjected to a second round of amplification with the same linker primer and a second primer specific to the known region. A variation on this procedure, which employs two primers that initiate extension in opposite directions from the known sequence, is described in WO 96/38591. Another such technique is known as "rapid amplification of cDNA ends" or RACE. This technique involves the use of an internal primer and an external primer, which hybridizes to a polyA region or vector sequence, to identify sequences that are 5' and 3' of a known sequence. Additional techniques include capture PCR (Lagerstrom et

WO 01/90152

PCT/US01/16776

45

al., *PCR Methods Applic.* 1:111-19, 1991) and walking PCR (Parker et al., *Nucl. Acids. Res.* 19:3055-60, 1991). Other methods employing amplification may also be employed to obtain a full length cDNA sequence.

5 In certain instances, it is possible to obtain a full length cDNA sequence by analysis of sequences provided in an expressed sequence tag (EST) database, such as that available from GenBank. Searches for overlapping ESTs may generally be performed using well known programs (e.g., NCBI BLAST searches), and such ESTs may be used to generate a contiguous full length sequence. Full length DNA sequences may also be obtained by analysis of genomic fragments.

10 In other embodiments of the invention, polynucleotide sequences or fragments thereof which encode polypeptides of the invention, or fusion proteins or functional equivalents thereof, may be used in recombinant DNA molecules to direct expression of a polypeptide in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences that encode substantially the same or a
15 functionally equivalent amino acid sequence may be produced and these sequences may be used to clone and express a given polypeptide.

As will be understood by those of skill in the art, it may be advantageous in some instances to produce polypeptide-encoding nucleotide sequences possessing non-naturally occurring codons. For example, codons preferred by a particular
20 prokaryotic or eukaryotic host can be selected to increase the rate of protein expression or to produce a recombinant RNA transcript having desirable properties, such as a half-life which is longer than that of a transcript generated from the naturally occurring sequence.

Moreover, the polynucleotide sequences of the present invention can be
25 engineered using methods generally known in the art in order to alter polypeptide encoding sequences for a variety of reasons, including but not limited to, alterations which modify the cloning, processing, and/or expression of the gene product. For example, DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide
30 sequences. In addition, site-directed mutagenesis may be used to insert new restriction

WO 01/90152

PCT/US01/16776

46

sites, alter glycosylation patterns, change codon preference, produce splice variants, or introduce mutations, and so forth.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences may be ligated to a heterologous sequence to encode a fusion protein. For example, to screen peptide libraries for inhibitors of polypeptide activity, it may be useful to encode a chimeric protein that can be recognized by a commercially available antibody. A fusion protein may also be engineered to contain a cleavage site located between the polypeptide-encoding sequence and the heterologous protein sequence, so that the polypeptide may be cleaved and purified away from the heterologous moiety.

Sequences encoding a desired polypeptide may be synthesized, in whole or in part, using chemical methods well known in the art (see Caruthers, M. H. et al. (1980) *Nucl. Acids Res. Symp. Ser.* 215-223, Horn, T. et al. (1980) *Nucl. Acids Res. Symp. Ser.* 225-232). Alternatively, the protein itself may be produced using chemical methods to synthesize the amino acid sequence of a polypeptide, or a portion thereof. For example, peptide synthesis can be performed using various solid-phase techniques (Roberge, J. Y. et al. (1995) *Science* 269:202-204) and automated synthesis may be achieved, for example, using the ABI 431A Peptide Synthesizer (Perkin Elmer, Palo Alto, CA).

A newly synthesized peptide may be substantially purified by preparative high performance liquid chromatography (e.g., Creighton, T. (1983) *Proteins, Structures and Molecular Principles*, WH Freeman and Co., New York, N.Y.) or other comparable techniques available in the art. The composition of the synthetic peptides may be confirmed by amino acid analysis or sequencing (e.g., the Edman degradation procedure). Additionally, the amino acid sequence of a polypeptide, or any part thereof, may be altered during direct synthesis and/or combined using chemical methods with sequences from other proteins, or any part thereof, to produce a variant polypeptide.

In order to express a desired polypeptide, the nucleotide sequences encoding the polypeptide, or functional equivalents, may be inserted into appropriate expression vector, i.e., a vector which contains the necessary elements for the

WO 01/90152

PCT/US01/16776

47

transcription and translation of the inserted coding sequence. Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding a polypeptide of interest and appropriate transcriptional and translational control elements. These methods include *in vitro* recombinant DNA techniques, synthetic techniques, and *in vivo* genetic recombination. Such techniques are described, for example, in Sambrook, J. et al. (1989) *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Press, Plainview, N.Y., and Ausubel, F. M. et al. (1989) *Current Protocols in Molecular Biology*, John Wiley & Sons, New York, N.Y.

10 A variety of expression vector/host systems may be utilized to contain and express polynucleotide sequences. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with virus expression vectors (e.g., baculovirus); plant cell
15 systems transformed with virus expression vectors (e.g., cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems.

The "control elements" or "regulatory sequences" present in an expression vector are those non-translated regions of the vector—enhancers, promoters,
20 5' and 3' untranslated regions—which interact with host cellular proteins to carry out transcription and translation. Such elements may vary in their strength and specificity. Depending on the vector system and host utilized, any number of suitable transcription and translation elements, including constitutive and inducible promoters, may be used. For example, when cloning in bacterial systems, inducible promoters such as the hybrid
25 lacZ promoter of the PBLUESCRIPT phagemid (Stratagene, La Jolla, Calif.) or PSPORT1 plasmid (Gibco BRL, Gaithersburg, MD) and the like may be used. In mammalian cell systems, promoters from mammalian genes or from mammalian viruses are generally preferred. If it is necessary to generate a cell line that contains multiple copies of the sequence encoding a polypeptide, vectors based on SV40 or EBV may be
30 advantageously used with an appropriate selectable marker.

WO 01/90152

PCT/US01/16776

48

In bacterial systems, any of a number of expression vectors may be selected depending upon the use intended for the expressed polypeptide. For example, when large quantities are needed, for example for the induction of antibodies, vectors which direct high level expression of fusion proteins that are readily purified may be used. Such vectors include, but are not limited to, the multifunctional *E. coli* cloning and expression vectors such as BLUESCRIPT (Stratagene), in which the sequence encoding the polypeptide of interest may be ligated into the vector in frame with sequences for the amino-terminal Met and the subsequent 7 residues of beta-galactosidase so that a hybrid protein is produced; pIN vectors (Van Hecke, G. and S. M. Schuster (1989) *J. Biol. Chem.* 264:5503-5509); and the like. pGEX Vectors (Promega, Madison, Wis.) may also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. Proteins made in such systems may be designed to include heparin, thrombin, or factor XA protease cleavage sites so that the cloned polypeptide of interest can be released from the GST moiety at will.

In the yeast, *Saccharomyces cerevisiae*, a number of vectors containing constitutive or inducible promoters such as alpha factor, alcohol oxidase, and PGII may be used. For reviews, see Ausubel et al. (supra) and Grant et al. (1987) *Methods Enzymol.* 153:516-544.

In cases where plant expression vectors are used, the expression of sequences encoding polypeptides may be driven by any of a number of promoters. For example, viral promoters such as the 35S and 19S promoters of CaMV may be used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) *EMBO J.* 6:307-311. Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used (Coruzzi, G. et al. (1984) *EMBO J.* 3:1671-1680; Broglie, R. et al. (1984) *Science* 224:838-843; and Winter, J. et al. (1991) *Results Probl. Cell Differ.* 17:85-105). These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. Such techniques

are described in a number of generally available reviews (see, for example, Hobbs, S. or Murry, L. F. in McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York, N.Y.; pp. 191-196).

An insect system may also be used to express a polypeptide of interest.

5 For example, in one such system, *Autographa californica* nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes in *Spodoptera frugiperda* cells or in *Trichoplusia* larvae. The sequences encoding the polypeptide may be cloned into a non-essential region of the virus, such as the polyhedrin gene, and placed under control of the polyhedrin promoter. Successful insertion of the polypeptide-encoding sequence
10 will render the polyhedrin gene inactive and produce recombinant virus lacking coat protein. The recombinant viruses may then be used to infect, for example, *S. frugiperda* cells or *Trichoplusia* larvae in which the polypeptide of interest may be expressed (Engelhard, E. K. et al. (1994) *Proc. Natl. Acad. Sci. 91*:3224-3227).

In mammalian host cells, a number of viral-based expression systems are
15 generally available. For example, in cases where an adenovirus is used as an expression vector, sequences encoding a polypeptide of interest may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain a viable virus which is capable of expressing the polypeptide in infected host
20 cells (Logan, J. and Shenk, T. (1984) *Proc. Natl. Acad. Sci. 81*:3655-3659). In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells.

Specific initiation signals may also be used to achieve more efficient translation of sequences encoding a polypeptide of interest. Such signals include the
25 ATG initiation codon and adjacent sequences. In cases where sequences encoding the polypeptide, its initiation codon, and upstream sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a portion thereof, is inserted, exogenous translational control signals including the ATG initiation
30 codon should be provided. Furthermore, the initiation codon should be in the correct

reading frame to ensure translation of the entire insert. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers which are appropriate for the particular cell system which is used, such as those described in the literature (Scharf, D. et al. (1994) *Results Probl. Cell Differ.* 20:125-162).

In addition, a host cell strain may be chosen for its ability to modulate the expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" form of the protein may also be used to facilitate correct insertion, folding and/or function. Different host cells such as CHO, COS, HeLa, MDCK, HEK293, and WI38, which have specific cellular machinery and characteristic mechanisms for such post-translational activities, may be chosen to ensure the correct modification and processing of the foreign protein.

For long-term, high-yield production of recombinant proteins, stable expression is generally preferred. For example, cell lines which stably express a polynucleotide of interest may be transformed using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for 1-2 days in an enriched media before they are switched to selective media. The purpose of the selectable marker is to confer resistance to selection, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be proliferated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase (Wigler, M. et al. (1977) *Cell* 11:223-32) and adenine phosphoribosyltransferase (Lowy, I. et al. (1990) *Cell* 22:817-23) genes which can be employed in tk.sup.- or apt.sup.- cells, respectively. Also, antimetabolite, antibiotic or herbicide resistance can be used as the basis for selection; for example, dhfr which confers resistance to

methotrexate (Wigler, M. et al. (1980) *Proc. Natl. Acad. Sci.* 77:3567-70); npt, which confers resistance to the aminoglycosides, neomycin and G-418 (Colbero-Garapin, F. et al (1981) *J. Mol. Biol.* 150:1-14); and ais or pat, which confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively (Murry, *supra*).

5 Additional selectable genes have been described, for example, trpB, which allows cells to utilize indole in place of tryptophan, or hisD, which allows cells to utilize histinol in place of histidine (Hartman, S. C. and R. C. Mulligan (1988) *Proc. Natl. Acad. Sci.* 85:8047-51). The use of visible markers has gained popularity with such markers as anthocyanins, beta-glucuronidase and its substrate GUS, and luciferase and its substrate

10 luciferin, being widely used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system (Rhodes, C. A. et al. (1995) *Methods Mol. Biol.* 55:121-131).

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

15 marker gene sequence, recombinant cells containing sequences can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a polypeptide-encoding sequence under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates

20 expression of the tandem gene as well.

Alternatively, host cells that contain and express a desired polynucleotide sequence may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations and protein bioassay or immunoassay techniques which include,

25 for example, membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein.

A variety of protocols for detecting and measuring the expression of polynucleotide-encoded products, using either polyclonal or monoclonal antibodies specific for the product are known in the art. Examples include enzyme-linked

30 immunosorbent assay (ELISA), radioimmunoassay (RIA), and fluorescence activated

cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on a given polypeptide may be preferred for some applications, but a competitive binding assay may also be employed. These and other assays are described, among other places, in Hampton, R. et al. (1990; Serological Methods, a Laboratory Manual, APS Press, St Paul, Minn.) and Maddox, D. E. et al. (1983; *J. Exp. Med.* 158:1211-1216).

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides include oligolabeling, nick translation, end-labeling or PCR amplification using a labeled nucleotide. Alternatively, the sequences, or any portions thereof may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits. Suitable reporter molecules or labels, which may be used include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with a polynucleotide sequence of interest may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a recombinant cell may be secreted or contained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides of the invention may be designed to contain signal sequences which direct secretion of the encoded polypeptide through a prokaryotic or eukaryotic cell membrane. Other recombinant constructions may be used to join sequences encoding a polypeptide of interest to nucleotide sequence encoding a polypeptide domain which will facilitate purification of soluble proteins. Such purification facilitating domains include, but are not limited to, metal chelating peptides such as histidine-tryptophan modules that allow purification on immobilized metals, protein A domains that allow purification on

immobilized immunoglobulin, and the domain utilized in the FLAGS extension/affinity purification system (Immunex Corp., Seattle, Wash.). The inclusion of cleavable linker sequences such as those specific for Factor XA or enterokinase (Invitrogen, San Diego, Calif.) between the purification domain and the encoded polypeptide may be used to
5 facilitate purification. One such expression vector provides for expression of a fusion protein containing a polypeptide of interest and a nucleic acid encoding 6 histidine residues preceding a thioredoxin or an enterokinase cleavage site. The histidine residues facilitate purification on IMAC (immobilized metal ion affinity chromatography) as described in Porath, J. et al. (1992, *Prot. Exp. Purif.* 3:263-281) while the enterokinase
10 cleavage site provides a means for purifying the desired polypeptide from the fusion protein. A discussion of vectors which contain fusion proteins is provided in Kroll, D. J. et al. (1993; *DNA Cell Biol.* 12:441-453).

In addition to recombinant production methods, polypeptides of the invention, and fragments thereof, may be produced by direct peptide synthesis using
15 solid-phase techniques (Merrifield J. (1963) *J. Am. Chem. Soc.* 85:2149-2154). Protein synthesis may be performed using manual techniques or by automation. Automated synthesis may be achieved, for example, using Applied Biosystems 431A Peptide Synthesizer (Perkin Elmer). Alternatively, various fragments may be chemically synthesized separately and combined using chemical methods to produce the full length
20 molecule.

Antibody Compositions, Fragments Thereof and Other Binding Agents

According to another aspect, the present invention further provides binding agents, such as antibodies and antigen-binding fragments thereof, that exhibit immunological binding to a tumor polypeptide disclosed herein, or to a portion, variant
25 or derivative thereof. An antibody, or antigen-binding fragment thereof, is said to "specifically bind," "immunologically bind," and/or is "immunologically reactive" to a polypeptide of the invention if it reacts at a detectable level (within, for example, an ELISA assay) with the polypeptide, and does not react detectably with unrelated polypeptides under similar conditions.

Immunological binding, as used in this context, generally refers to the non-covalent interactions of the type which occur between an immunoglobulin molecule and an antigen for which the immunoglobulin is specific. The strength, or affinity of immunological binding interactions can be expressed in terms of the dissociation constant (K_d) of the interaction, wherein a smaller K_d represents a greater affinity. Immunological binding properties of selected polypeptides can be quantified using methods well known in the art. One such method entails measuring the rates of antigen-binding site/antigen complex formation and dissociation, wherein those rates depend on the concentrations of the complex partners, the affinity of the interaction, and on geometric parameters that equally influence the rate in both directions. Thus, both the "on rate constant" (K_{on}) and the "off rate constant" (K_{off}) can be determined by calculation of the concentrations and the actual rates of association and dissociation. The ratio of K_{off}/K_{on} enables cancellation of all parameters not related to affinity, and is thus equal to the dissociation constant K_d . See, generally, Davies et al. (1990) Annual Rev. Biochem. 59:439-473.

An "antigen-binding site," or "binding portion" of an antibody refers to the part of the immunoglobulin molecule that participates in antigen binding. The antigen binding site is formed by amino acid residues of the N-terminal variable ("V") regions of the heavy ("H") and light ("L") chains. Three highly divergent stretches within the V regions of the heavy and light chains are referred to as "hypervariable regions" which are interposed between more conserved flanking stretches known as "framework regions," or "FRs". Thus the term "FR" refers to amino acid sequences which are naturally found between and adjacent to hypervariable regions in immunoglobulins. In an antibody molecule, the three hypervariable regions of a light chain and the three hypervariable regions of a heavy chain are disposed relative to each other in three dimensional space to form an antigen-binding surface. The antigen-binding surface is complementary to the three-dimensional surface of a bound antigen, and the three hypervariable regions of each of the heavy and light chains are referred to as "complementarity-determining regions," or "CDRs."

WO 01/90152

PCT/US01/16776

35

Binding agents may be further capable of differentiating between patients with and without a cancer, such as breast cancer, using the representative assays provided herein. For example, antibodies or other binding agents that bind to a tumor protein will preferably generate a signal indicating the presence of a cancer in at least about 20% of patients with the disease, more preferably at least about 30% of patients. Alternatively, or in addition, the antibody will generate a negative signal indicating the absence of the disease in at least about 90% of individuals without the cancer. To determine whether a binding agent satisfies this requirement, biological samples (*e.g.*, blood, sera, sputum, urine and/or tumor biopsies) from patients with and without a cancer (as determined using standard clinical tests) may be assayed as described herein for the presence of polypeptides that bind to the binding agent. Preferably, a statistically significant number of samples with and without the disease will be assayed. Each binding agent should satisfy the above criteria; however, those of ordinary skill in the art will recognize that binding agents may be used in combination to improve sensitivity.

Any agent that satisfies the above requirements may be a binding agent. For example, a binding agent may be a ribosome, with or without a peptide component, an RNA molecule or a polypeptide. In a preferred embodiment, a binding agent is an antibody or an antigen-binding fragment thereof. Antibodies may be prepared by any of a variety of techniques known to those of ordinary skill in the art. *See, e.g.*, Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988. In general, antibodies can be produced by cell culture techniques, including the generation of monoclonal antibodies as described herein, or via transfection of antibody genes into suitable bacterial or mammalian cell hosts, in order to allow for the production of recombinant antibodies. In one technique, an immunogen comprising the polypeptide is initially injected into any of a wide variety of mammals (*e.g.*, mice, rats, rabbits, sheep or goats). In this step, the polypeptides of this invention may serve as the immunogen without modification. Alternatively, particularly for relatively short polypeptides, a superior immune response may be elicited if the polypeptide is joined to a carrier protein, such as bovine serum albumin or keyhole limpet hemocyanin. The immunogen

WO 01/90152

PCT/US01/16776

56

is injected into the animal host, preferably according to a predetermined schedule incorporating one or more booster immunizations, and the animals are bled periodically. Polyclonal antibodies specific for the polypeptide may then be purified from such antisera by, for example, affinity chromatography using the polypeptide coupled to a
5 suitable solid support.

Monoclonal antibodies specific for an antigenic polypeptide of interest may be prepared, for example, using the technique of Kohler and Milstein, *Eur. J. Immunol.* 6:511-519, 1976, and improvements thereto. Briefly, these methods involve the preparation of immortal cell lines capable of producing antibodies having the
10 desired specificity (*i.e.*, reactivity with the polypeptide of interest). Such cell lines may be produced, for example, from spleen cells obtained from an animal immunized as described above. The spleen cells are then immortalized by, for example, fusion with a myeloma cell fusion partner, preferably one that is syngeneic with the immunized animal. A variety of fusion techniques may be employed. For example, the spleen cells
15 and myeloma cells may be combined with a nonionic detergent for a few minutes and then plated at low density on a selective medium that supports the growth of hybrid cells, but not myeloma cells. A preferred selection technique uses HAT (hypoxanthine, aminopterin, thymidine) selection. After a sufficient time, usually about 1 to 2 weeks, colonies of hybrids are observed. Single colonies are selected and their culture
20 supernatants tested for binding activity against the polypeptide. Hybridomas having high reactivity and specificity are preferred.

Monoclonal antibodies may be isolated from the supernatants of growing hybridoma colonies. In addition, various techniques may be employed to enhance the yield, such as injection of the hybridoma cell line into the peritoneal cavity of a suitable
25 vertebrate host, such as a mouse. Monoclonal antibodies may then be harvested from the ascites fluid or the blood. Contaminants may be removed from the antibodies by conventional techniques, such as chromatography, gel filtration, precipitation, and extraction. The polypeptides of this invention may be used in the purification process in, for example, an affinity chromatography step.

WO 01/90152

PCT/US01/16776

57

A number of therapeutically useful molecules are known in the art which comprise antigen-binding sites that are capable of exhibiting immunological binding properties of an antibody molecule. The proteolytic enzyme papain preferentially cleaves IgG molecules to yield several fragments, two of which (the "F(ab)" fragments) each comprise a covalent heterodimer that includes an intact antigen-binding site. The enzyme pepsin is able to cleave IgG molecules to provide several fragments, including the "F(ab)₂" fragment which comprises both antigen-binding sites. An "Fv" fragment can be produced by preferential proteolytic cleavage of an IgM₁ and on rare occasions IgG or IgA immunoglobulin molecule. Fv fragments are, however, more commonly derived using recombinant techniques known in the art. The Fv fragment includes a non-covalent V_H::V_L heterodimer including an antigen-binding site which retains much of the antigen recognition and binding capabilities of the native antibody molecule. Inbar et al. (1972) Proc. Nat. Acad. Sci. USA 69:2659-2662; Hochman et al. (1976) Biochem 15:2706-2710; and Ehrlich et al. (1980) Biochem 19:4091-4096.

A single chain Fv ("sFv") polypeptide is a covalently linked V_H::V_L heterodimer which is expressed from a gene fusion including V_H- and V_L-encoding genes linked by a peptide-encoding linker. Huston et al. (1988) Proc. Nat. Acad. Sci. USA 85(16):5879-5883. A number of methods have been described to discern chemical structures for converting the naturally aggregated--but chemically separated--light and heavy polypeptide chains from an antibody V region into an sFv molecule which will fold into a three dimensional structure substantially similar to the structure of an antigen-binding site. See, e.g., U.S. Pat. Nos. 5,091,513 and 5,132,405, to Huston et al.; and U.S. Pat. No. 4,946,778, to Ladner et al.

Each of the above-described molecules includes a heavy chain and a light chain CDR set, respectively interposed between a heavy chain and a light chain FR set which provide support to the CDRs and define the spatial relationship of the CDRs relative to each other. As used herein, the term "CDR set" refers to the three hypervariable regions of a heavy or light chain V region. Proceeding from the N-terminus of a heavy or light chain, these regions are denoted as "CDR1," "CDR2," and "CDR3" respectively. An antigen-binding site, therefore, includes six CDRs,

WO 01/90152

PCT/US01/16776

58

comprising the CDR set from each of a heavy and a light chain V region. A polypeptide comprising a single CDR, (e.g., a CDR1, CDR2 or CDR3) is referred to herein as a "molecular recognition unit." Crystallographic analysis of a number of antigen-antibody complexes has demonstrated that the amino acid residues of CDRs form extensive contact with bound antigen, wherein the most extensive antigen contact is with the heavy chain CDR3. Thus, the molecular recognition units are primarily responsible for the specificity of an antigen-binding site.

As used herein, the term "FR set" refers to the four flanking amino acid sequences which frame the CDRs of a CDR set of a heavy or light chain V region. Some FR residues may contact bound antigen; however, FRs are primarily responsible for folding the V region into the antigen-binding site, particularly the FR residues directly adjacent to the CDRs. Within FRs, certain amino residues and certain structural features are very highly conserved. In this regard, all V region sequences contain an internal disulfide loop of around 90 amino acid residues. When the V regions fold into a binding-site, the CDRs are displayed as projecting loop motifs which form an antigen-binding surface. It is generally recognized that there are conserved structural regions of FRs which influence the folded shape of the CDR loops into certain "canonical" structures—regardless of the precise CDR amino acid sequence. Further, certain FR residues are known to participate in non-covalent interdomain contacts which stabilize the interaction of the antibody heavy and light chains.

A number of "humanized" antibody molecules comprising an antigen-binding site derived from a non-human immunoglobulin have been described, including chimeric antibodies having rodent V regions and their associated CDRs fused to human constant domains (Winter et al. (1991) *Nature* 349:293-299; Lobuglio et al. (1989) *Proc. Nat. Acad. Sci. USA* 86:4220-4224; Shaw et al. (1987) *J Immunol.* 138:4534-4538; and Brown et al. (1987) *Cancer Res.* 47:3577-3583), rodent CDRs grafted into a human supporting FR prior to fusion with an appropriate human antibody constant domain (Riechmann et al. (1988) *Nature* 332:323-327; Verhoeven et al. (1988) *Science* 239:1534-1536; and Jones et al. (1986) *Nature* 321:522-525), and rodent CDRs supported by recombinantly veneered rodent FRs (European Patent Publication No.

519,596, published Dec. 23, 1992). These "humanized" molecules are designed to minimize unwanted immunological response toward rodent antihuman antibody molecules which limits the duration and effectiveness of therapeutic applications of those moieties in human recipients.

5 As used herein, the terms "veneered FRs" and "recombinantly veneered FRs" refer to the selective replacement of FR residues from, *e.g.*, a rodent heavy or light chain V region, with human FR residues in order to provide a xenogeneic molecule comprising an antigen-binding site which retains substantially all of the native FR polypeptide folding structure. Veneering techniques are based on the understanding that
10 the ligand binding characteristics of an antigen-binding site are determined primarily by the structure and relative disposition of the heavy and light chain CDR sets within the antigen-binding surface. Davies et al. (1990) *Ann. Rev. Biochem.* 59:439-473. Thus, antigen binding specificity can be preserved in a humanized antibody only wherein the CDR structures, their interaction with each other, and their interaction with the rest of
15 the V region domains are carefully maintained. By using veneering techniques, exterior (*e.g.*, solvent-accessible) FR residues which are readily encountered by the immune system are selectively replaced with human residues to provide a hybrid molecule that comprises either a weakly immunogenic, or substantially non-immunogenic veneered surface.

20 The process of veneering makes use of the available sequence data for human antibody variable domains compiled by Kabat et al., in *Sequences of Proteins of Immunological Interest*, 4th ed., (U.S. Dept. of Health and Human Services, U.S. Government Printing Office, 1987), updates to the Kabat database, and other accessible U.S. and foreign databases (both nucleic acid and protein). Solvent accessibilities of V
25 region amino acids can be deduced from the known three-dimensional structure for human and murine antibody fragments. There are two general steps in veneering a murine antigen-binding site. Initially, the FRs of the variable domains of an antibody molecule of interest are compared with corresponding FR sequences of human variable domains obtained from the above-identified sources. The most homologous human V
30 regions are then compared residue by residue to corresponding murine amino acids. The

WO 01/90152

PCT/US01/16776

60

residues in the murine FR which differ from the human counterpart are replaced by the residues present in the human moiety using recombinant techniques well known in the art. Residue switching is only carried out with moieties which are at least partially exposed (solvent accessible), and care is exercised in the replacement of amino acid
5 residues which may have a significant effect on the tertiary structure of V region domains, such as proline, glycine and charged amino acids.

In this manner, the resultant "vencered" murine antigen-binding sites are thus designed to retain the murine CDR residues, the residues substantially adjacent to the CDRs, the residues identified as buried or mostly buried (solvent inaccessible), the
10 residues believed to participate in non-covalent (*e.g.*, electrostatic and hydrophobic) contacts between heavy and light chain domains, and the residues from conserved structural regions of the FRs which are believed to influence the "canonical" tertiary structures of the CDR loops. These design criteria are then used to prepare recombinant nucleotide sequences which combine the CDRs of both the heavy and light chain of a
15 murine antigen-binding site into human-appearing FRs that can be used to transfect mammalian cells for the expression of recombinant human antibodies which exhibit the antigen specificity of the murine antibody molecule.

In another embodiment of the invention, monoclonal antibodies of the present invention may be coupled to one or more therapeutic agents. Suitable agents in
20 this regard include radionuclides, differentiation inducers, drugs, toxins, and derivatives thereof. Preferred radionuclides include ^{90}Y , ^{125}I , ^{125}I , ^{131}I , ^{186}Re , ^{188}Re , ^{211}At , and ^{212}Bi . Preferred drugs include methotrexate, and pyrimidine and purine analogs. Preferred differentiation inducers include phorbol esters and butyric acid. Preferred toxins include ricin, abrin, diphtheria toxin, cholera toxin, gelonin, *Pseudomonas*
25 exotoxin, *Shigella* toxin, and pokeweed antiviral protein.

A therapeutic agent may be coupled (*e.g.*, covalently bonded) to a suitable monoclonal antibody either directly or indirectly (*e.g.*, via a linker group). A direct reaction between an agent and an antibody is possible when each possesses a substituent capable of reacting with the other. For example, a nucleophilic group, such
30 as an amino or sulfhydryl group, on one may be capable of reacting with a carbonyl-

WO 01/90152

PCT/US01/16776

61

containing group, such as an anhydride or an acid halide, or with an alkyl group containing a good leaving group (*e.g.*, a halide) on the other.

Alternatively, it may be desirable to couple a therapeutic agent and an antibody via a linker group. A linker group can function as a spacer to distance an antibody from an agent in order to avoid interference with binding capabilities. A linker group can also serve to increase the chemical reactivity of a substituent on an agent or an antibody, and thus increase the coupling efficiency. An increase in chemical reactivity may also facilitate the use of agents, or functional groups on agents, which otherwise would not be possible.

It will be evident to those skilled in the art that a variety of bifunctional or polyfunctional reagents, both homo- and hetero-functional (such as those described in the catalog of the Pierce Chemical Co., Rockford, IL), may be employed as the linker group. Coupling may be effected, for example, through amino groups, carboxyl groups, sulfhydryl groups or oxidized carbohydrate residues. There are numerous references describing such methodology, *e.g.*, U.S. Patent No. 4,671,958, to Rodwell et al.

Where a therapeutic agent is more potent when free from the antibody portion of the immunoconjugates of the present invention, it may be desirable to use a linker group which is cleavable during or upon internalization into a cell. A number of different cleavable linker groups have been described. The mechanisms for the intracellular release of an agent from these linker groups include cleavage by reduction of a disulfide bond (*e.g.*, U.S. Patent No. 4,489,710, to Spittler), by irradiation of a photolabile bond (*e.g.*, U.S. Patent No. 4,625,014, to Senter et al.), by hydrolysis of derivatized amino acid side chains (*e.g.*, U.S. Patent No. 4,638,045, to Kohn et al.), by serum complement-mediated hydrolysis (*e.g.*, U.S. Patent No. 4,671,958, to Rodwell et al.), and acid-catalyzed hydrolysis (*e.g.*, U.S. Patent No. 4,569,789, to Blattler et al.).

It may be desirable to couple more than one agent to an antibody. In one embodiment, multiple molecules of an agent are coupled to one antibody molecule. In another embodiment, more than one type of agent may be coupled to one antibody. Regardless of the particular embodiment, immunoconjugates with more than one agent may be prepared in a variety of ways. For example, more than one agent may be

WO 01/90152

PCT/US01/16776

62

coupled directly to an antibody molecule, or linkers that provide multiple sites for attachment can be used. Alternatively, a carrier can be used.

A carrier may bear the agents in a variety of ways, including covalent bonding either directly or via a linker group. Suitable carriers include proteins such as albumins (e.g., U.S. Patent No. 4,507,234, to Kato et al.), peptides and polysaccharides such as aminodextran (e.g., U.S. Patent No. 4,699,784, to Shih et al.). A carrier may also bear an agent by noncovalent bonding or by encapsulation, such as within a liposome vesicle (e.g., U.S. Patent Nos. 4,429,008 and 4,873,088). Carriers specific for radionuclide agents include radiohalogenated small molecules and chelating compounds. For example, U.S. Patent No. 4,735,792 discloses representative radiohalogenated small molecules and their synthesis. A radionuclide chelate may be formed from chelating compounds that include those containing nitrogen and sulfur atoms as the donor atoms for binding the metal, or metal oxide, radionuclide. For example, U.S. Patent No. 4,673,562, to Davison et al. discloses representative chelating compounds and their synthesis.

T Cell Compositions

The present invention, in another aspect, provides T cells specific for a tumor polypeptide disclosed herein, or for a variant or derivative thereof. Such cells may generally be prepared *in vitro* or *ex vivo*, using standard procedures. For example, T cells may be isolated from bone marrow, peripheral blood, or a fraction of bone marrow or peripheral blood of a patient, using a commercially available cell separation system, such as the Isolex™ System, available from Nexell Therapeutics, Inc. (Irvine, CA; see also U.S. Patent No. 5,240,856; U.S. Patent No. 5,215,926; WO 89/06280; WO 91/16116 and WO 92/07243). Alternatively, T cells may be derived from related or unrelated humans, non-human mammals, cell lines or cultures.

T cells may be stimulated with a polypeptide, polynucleotide encoding a polypeptide and/or an antigen presenting cell (APC) that expresses such a polypeptide. Such stimulation is performed under conditions and for a time sufficient to permit the generation of T cells that are specific for the polypeptide of interest. Preferably, a tumor

WO 01/90152

PCT/US01/16776

63

polypeptide or polynucleotide of the invention is present within a delivery vehicle, such as a microsphere, to facilitate the generation of specific T cells.

T cells are considered to be specific for a polypeptide of the present invention if the T cells specifically proliferate, secrete cytokines or kill target cells coated with the polypeptide or expressing a gene encoding the polypeptide. T cell specificity may be evaluated using any of a variety of standard techniques. For example, within a chromium release assay or proliferation assay, a stimulation index of more than two fold increase in lysis and/or proliferation, compared to negative controls, indicates T cell specificity. Such assays may be performed, for example, as described in Chien et al., *Cancer Res.* 54:1065-1070, 1994. Alternatively, detection of the proliferation of T cells may be accomplished by a variety of known techniques. For example, T cell proliferation can be detected by measuring an increased rate of DNA synthesis (e.g. by pulse-labeling cultures of T cells with tritiated thymidine and measuring the amount of tritiated thymidine incorporated into DNA). Contact with a tumor polypeptide (100 ng/ml - 100 µg/ml, preferably 200 ng/ml - 25 µg/ml) for 3 - 7 days will typically result in at least a two fold increase in proliferation of the T cells. Contact as described above for 2-3 hours should result in activation of the T cells, as measured using standard cytokine assays in which a two fold increase in the level of cytokine release (e.g., TNF or IFN-γ) is indicative of T cell activation (see Coligan et al., *Current Protocols in Immunology*, vol. 1, Wiley Interscience (Greene 1998)). T cells that have been activated in response to a tumor polypeptide, polynucleotide or polypeptide-expressing APC may be CD4⁺ and/or CD8⁺. Tumor polypeptide-specific T cells may be expanded using standard techniques. Within preferred embodiments, the T cells are derived from a patient, a related donor or an unrelated donor, and are administered to the patient following stimulation and expansion.

For therapeutic purposes, CD4⁺ or CD8⁺ T cells that proliferate in response to a tumor polypeptide, polynucleotide or APC can be expanded in number either *in vitro* or *in vivo*. Proliferation of such T cells *in vitro* may be accomplished in a variety of ways. For example, the T cells can be re-exposed to a tumor polypeptide, or a short peptide corresponding to an immunogenic portion of such a polypeptide, with or

WO 01/90152

PCT/US01/16776

64

without the addition of T cell growth factors, such as interleukin-2, and/or stimulator cells that synthesize a tumor polypeptide. Alternatively, one or more T cells that proliferate in the presence of the tumor polypeptide can be expanded in number by cloning. Methods for cloning cells are well known in the art, and include limiting
5 dilution.

Pharmaceutical Compositions

In additional embodiments, the present invention concerns formulation of one or more of the polynucleotide, polypeptide, T-cell and/or antibody compositions disclosed herein in pharmaceutically-acceptable carriers for administration to a cell or
10 an animal, either alone, or in combination with one or more other modalities of therapy.

It will be understood that, if desired, a composition as disclosed herein may be administered in combination with other agents as well, such as, *e.g.*, other proteins or polypeptides or various pharmaceutically-active agents. In fact, there is virtually no limit to other components that may also be included, given that the
15 additional agents do not cause a significant adverse effect upon contact with the target cells or host tissues. The compositions may thus be delivered along with various other agents as required in the particular instance. Such compositions may be purified from host cells or other biological sources, or alternatively may be chemically synthesized as described herein. Likewise, such compositions may further comprise substituted or
20 derivatized RNA or DNA compositions.

Therefore, in another aspect of the present invention, pharmaceutical compositions are provided comprising one or more of the polynucleotide, polypeptide, antibody, and/or T-cell compositions described herein in combination with a physiologically acceptable carrier. In certain preferred embodiments, the
25 pharmaceutical compositions of the invention comprise immunogenic polynucleotide and/or polypeptide compositions of the invention for use in prophylactic and therapeutic vaccine applications. Vaccine preparation is generally described in, for example, M.F. Powell and M.J. Newman, eds., "Vaccine Design (the subunit and adjuvant approach)," Plenum Press (NY, 1995). Generally, such compositions will comprise one or more

polynucleotide and/or polypeptide compositions of the present invention in combination with one or more immunostimulants.

It will be apparent that any of the pharmaceutical compositions described herein can contain pharmaceutically acceptable salts of the polynucleotides and polypeptides of the invention. Such salts can be prepared, for example, from pharmaceutically acceptable non-toxic bases, including organic bases (e.g., salts of primary, secondary and tertiary amines and basic amino acids) and inorganic bases (e.g., sodium, potassium, lithium, ammonium, calcium and magnesium salts).

In another embodiment, illustrative immunogenic compositions, e.g., vaccine compositions, of the present invention comprise DNA encoding one or more of the polypeptides as described above, such that the polypeptide is generated *in situ*. As noted above, the polynucleotide may be administered within any of a variety of delivery systems known to those of ordinary skill in the art. Indeed, numerous gene delivery techniques are well known in the art, such as those described by Rolland, *Crit. Rev. Therap. Drug Carrier Systems* 15:143-198, 1998, and references cited therein. Appropriate polynucleotide expression systems will, of course, contain the necessary regulatory DNA regulatory sequences for expression in a patient (such as a suitable promoter and terminating signal). Alternatively, bacterial delivery systems may involve the administration of a bacterium (such as *Bacillus-Calmette-Guerrin*) that expresses an immunogenic portion of the polypeptide on its cell surface or secretes such an epitope.

Therefore, in certain embodiments, polynucleotides encoding immunogenic polypeptides described herein are introduced into suitable mammalian host cells for expression using any of a number of known viral-based systems. In one illustrative embodiment, retroviruses provide a convenient and effective platform for gene delivery systems. A selected nucleotide sequence encoding a polypeptide of the present invention can be inserted into a vector and packaged in retroviral particles using techniques known in the art. The recombinant virus can then be isolated and delivered to a subject. A number of illustrative retroviral systems have been described (e.g., U.S. Pat. No. 5,219,740; Miller and Rosman (1989) *BioTechniques* 7:980-990; Miller, A. D. (1990) *Human Gene Therapy* 1:5-14; Scarpa et al. (1991) *Virology* 180:849-852; Burns

WO 01/90152

PCT/US01/16776

66

et al. (1993) *Proc. Natl. Acad. Sci. USA* 90:8033-8037; and Boris-Lawrie and Temin (1993) *Cur. Opin. Genet. Develop.* 3:102-109.

In addition, a number of illustrative adenovirus-based systems have also been described. Unlike retroviruses which integrate into the host genome, adenoviruses
 5 persist extrachromosomally thus minimizing the risks associated with insertional mutagenesis (Haj-Ahmad and Graham (1986) *J. Virol.* 57:267-274; Bell et al. (1993) *J. Virol.* 67:5911-5921; Mittereder et al. (1994) *Human Gene Therapy* 5:717-729; Seth et al. (1994) *J. Virol.* 68:933-940; Barr et al. (1994) *Gene Therapy* 1:51-58; Berkner, K. L. (1988) *BioTechniques* 6:616-629; and Rich et al. (1993) *Human Gene Therapy* 4:461-
 10 476).

Various adeno-associated virus (AAV) vector systems have also been developed for polynucleotide delivery. AAV vectors can be readily constructed using techniques well known in the art. See, e.g., U.S. Pat. Nos. 5,173,414 and 5,139,941; International Publication Nos. WO 92/01070 and WO 93/03769; Lebkowski et al.
 15 (1988) *Molec. Cell. Biol.* 8:3988-3996; Vincent et al. (1990) *Vaccines 90* (Cold Spring Harbor Laboratory Press); Carter, B. J. (1992) *Current Opinion in Biotechnology* 3:533-539; Muzyczka, N. (1992) *Current Topics in Microbiol. and Immunol.* 158:97-129; Kotin, R. M. (1994) *Human Gene Therapy* 5:793-801; Shelling and Smith (1994) *Gene Therapy* 1:165-169; and Zhou et al. (1994) *J. Exp. Med.* 179:1867-1875.

Additional viral vectors useful for delivering the polynucleotides encoding polypeptides of the present invention by gene transfer include those derived from the pox family of viruses, such as vaccinia virus and avian poxvirus. By way of example, vaccinia virus recombinants expressing the novel molecules can be constructed as follows. The DNA encoding a polypeptide is first inserted into an
 25 appropriate vector so that it is adjacent to a vaccinia promoter and flanking vaccinia DNA sequences, such as the sequence encoding thymidine kinase (TK). This vector is then used to transfect cells which are simultaneously infected with vaccinia. Homologous recombination serves to insert the vaccinia promoter plus the gene encoding the polypeptide of interest into the viral genome. The resulting TK.sup.(-)

WO 01/90152

PCT/US01/16776

67

recombinant can be selected by culturing the cells in the presence of 5-bromodeoxyuridine and picking viral plaques resistant thereto.

A vaccinia-based infection/transfection system can be conveniently used to provide for inducible, transient expression or coexpression of one or more polypeptides described herein in host cells of an organism. In this particular system, cells are first infected in vitro with a vaccinia virus recombinant that encodes the bacteriophage T7 RNA polymerase. This polymerase displays exquisite specificity in that it only transcribes templates bearing T7 promoters. Following infection, cells are transfected with the polynucleotide or polynucleotides of interest, driven by a T7 promoter. The polymerase expressed in the cytoplasm from the vaccinia virus recombinant transcribes the transfected DNA into RNA which is then translated into polypeptide by the host translational machinery. The method provides for high level, transient, cytoplasmic production of large quantities of RNA and its translation products. See, *e.g.*, Elroy-Stein and Moss, Proc. Natl. Acad. Sci. USA (1990) 87:6743-6747; Fuerst et al. Proc. Natl. Acad. Sci. USA (1986) 83:8122-8126.

Alternatively, avipoxviruses, such as the fowlpox and canarypox viruses, can also be used to deliver the coding sequences of interest. Recombinant avipox viruses, expressing immunogens from mammalian pathogens, are known to confer protective immunity when administered to non-avian species. The use of an Avipox vector is particularly desirable in human and other mammalian species since members of the Avipox genus can only productively replicate in susceptible avian species and therefore are not infective in mammalian cells. Methods for producing recombinant Avipoxviruses are known in the art and employ genetic recombination, as described above with respect to the production of vaccinia viruses. See, *e.g.*, WO 91/12882; WO 89/03429; and WO 92/03545.

Any of a number of alphavirus vectors can also be used for delivery of polynucleotide compositions of the present invention, such as those vectors described in U.S. Patent Nos. 5,843,723; 6,015,686; 6,008,035 and 6,015,694. Certain vectors based on Venezuelan Equine Encephalitis (VEE) can also be used, illustrative examples of which can be found in U.S. Patent Nos. 5,505,947 and 5,643,576.

WO 01/90152

PCT/US01/16776

68

Moreover, molecular conjugate vectors, such as the adenovirus chimeric vectors described in Michael et al. *J. Biol. Chem.* (1993) 268:6866-6869 and Wagner et al. *Proc. Natl. Acad. Sci. USA* (1992) 89:6099-6103, can also be used for gene delivery under the invention.

5 Additional illustrative information on these and other known viral-based delivery systems can be found, for example, in Fisher-Hoch et al., *Proc. Natl. Acad. Sci. USA* 86:317-321, 1989; Flexner et al., *Ann. N.Y. Acad. Sci.* 569:86-103, 1989; Flexner et al., *Vaccine* 8:17-21, 1990; U.S. Patent Nos. 4,603,112, 4,769,330, and 5,017,487; WO 89/01973; U.S. Patent No. 4,777,127; GB 2,200,651; EP 0,345,242;
10 WO 91/02805; Berkner, *Biotechniques* 6:616-627, 1988; Rosenfeld et al., *Science* 252:431-434, 1991; Kolls et al., *Proc. Natl. Acad. Sci. USA* 91:215-219, 1994; Kass-Eisler et al., *Proc. Natl. Acad. Sci. USA* 90:11498-11502, 1993; Guzman et al., *Circulation* 88:2838-2848, 1993; and Guzman et al., *Cir. Res.* 73:1202-1207, 1993.

In certain embodiments, a polynucleotide may be integrated into the
15 genome of a target cell. This integration may be in the specific location and orientation via homologous recombination (gene replacement) or it may be integrated in a random, non-specific location (gene augmentation). In yet further embodiments, the polynucleotide may be stably maintained in the cell as a separate, episomal segment of DNA. Such polynucleotide segments or "episomes" encode sequences sufficient to
20 permit maintenance and replication independent of or in synchronization with the host cell cycle. The manner in which the expression construct is delivered to a cell and where in the cell the polynucleotide remains is dependent on the type of expression construct employed.

In another embodiment of the invention, a polynucleotide is
25 administered/delivered as "naked" DNA, for example as described in Ulmer et al., *Science* 259:1745-1749, 1993 and reviewed by Cohen, *Science* 259:1691-1692, 1993. The uptake of naked DNA may be increased by coating the DNA onto biodegradable beads, which are efficiently transported into the cells.

In still another embodiment, a composition of the present invention can
30 be delivered via a particle bombardment approach, many of which have been described.

WO 01/90152

PCT/US01/16776

69

In one illustrative example, gas-driven particle acceleration can be achieved with devices such as those manufactured by Powderject Pharmaceuticals PLC (Oxford, UK) and Powderject Vaccines Inc. (Madison, WI), some examples of which are described in U.S. Patent Nos. 5,846,796; 6,010,478; 5,865,796; 5,584,807; and EP Patent No. 0500

5 799. This approach offers a needle-free delivery approach wherein a dry powder formulation of microscopic particles, such as polynucleotide or polypeptide particles, are accelerated to high speed within a helium gas jet generated by a hand held device, propelling the particles into a target tissue of interest.

In a related embodiment, other devices and methods that may be useful for gas-driven needle-less injection of compositions of the present invention include those provided by Bioject, Inc. (Portland, OR), some examples of which are described in U.S. Patent Nos. 4,790,824; 5,064,413; 5,312,335; 5,383,851; 5,399,163; 5,520,639 and 5,993,412.

According to another embodiment, the pharmaceutical compositions described herein will comprise one or more immunostimulants in addition to the immunogenic polynucleotide, polypeptide, antibody, T-cell and/or APC compositions of this invention. An immunostimulant refers to essentially any substance that enhances or potentiates an immune response (antibody and/or cell-mediated) to an exogenous antigen. One preferred type of immunostimulant comprises an adjuvant. Many adjuvants contain a substance designed to protect the antigen from rapid catabolism, such as aluminum hydroxide or mineral oil, and a stimulator of immune responses, such as lipid A, *Bordetella pertussis* or *Mycobacterium tuberculosis* derived proteins. Certain adjuvants are commercially available as, for example, Freund's Incomplete Adjuvant and Complete Adjuvant (Difco Laboratories, Detroit, MI); Merck Adjuvant 65 (Merck and Company, Inc., Rahway, NJ); AS-2 (SmithKline Beecham, Philadelphia, PA); aluminum salts such as aluminum hydroxide gel (alum) or aluminum phosphate; salts of calcium, iron or zinc; an insoluble suspension of acylated tyrosine; acylated sugars; cationically or anionically derivatized polysaccharides; polyphosphazenes; biodegradable microspheres; monophosphoryl lipid A and quil A. Cytokines, such as

WO 01/90152

PCT/US01/16776

70

GM-CSF, interleukin-2, -7, -12, and other like growth factors, may also be used as adjuvants.

Within certain embodiments of the invention, the adjuvant composition is preferably one that induces an immune response predominantly of the Th1 type. High levels of Th1-type cytokines (e.g., IFN- γ , TNF α , IL-2 and IL-12) tend to favor the induction of cell mediated immune responses to an administered antigen. In contrast, high levels of Th2-type cytokines (e.g., IL-4, IL-5, IL-6 and IL-10) tend to favor the induction of humoral immune responses. Following application of a vaccine as provided herein, a patient will support an immune response that includes Th1- and Th2-type responses. Within a preferred embodiment, in which a response is predominantly Th1-type, the level of Th1-type cytokines will increase to a greater extent than the level of Th2-type cytokines. The levels of these cytokines may be readily assessed using standard assays. For a review of the families of cytokines, see Mosmann and Coffman, *Ann. Rev. Immunol.* 7:145-173, 1989.

Certain preferred adjuvants for eliciting a predominantly Th1-type response include, for example, a combination of monophosphoryl lipid A, preferably 3-de-O-acylated monophosphoryl lipid A, together with an aluminum salt. MPL[®] adjuvants are available from Corixa Corporation (Seattle, WA; see, for example, US Patent Nos. 4,436,727; 4,877,611; 4,866,034 and 4,912,094). CpG-containing oligonucleotides (in which the CpG dinucleotide is unmethylated) also induce a predominantly Th1 response. Such oligonucleotides are well known and are described, for example, in WO 96/02555, WO 99/33488 and U.S. Patent Nos. 6,008,200 and 5,856,462. Immunostimulatory DNA sequences are also described, for example, by Sato et al., *Science* 273:352, 1996. Another preferred adjuvant comprises a saponin, such as Quil A, or derivatives thereof, including QS21 and QS7 (Aquila Biopharmaceuticals Inc., Framingham, MA); Escin; Digitonin; or *Gypsophila* or *Chenopodium quinoa* saponins. Other preferred formulations include more than one saponin in the adjuvant combinations of the present invention, for example combinations of at least two of the following group comprising QS21, QS7, Quil A, β -escin, or digitonin.

WO 01/90152

PCT/US01/16776

71

Alternatively the saponin formulations may be combined with vaccine vehicles composed of chitosan or other polycationic polymers, polylactide and polylactide-co-glycolide particles, poly-N-acetyl glucosamine-based polymer matrix, particles composed of polysaccharides or chemically modified polysaccharides, liposomes and lipid-based particles, particles composed of glycerol monoesters, etc. The saponins may also be formulated in the presence of cholesterol to form particulate structures such as liposomes or ISCOMs. Furthermore, the saponins may be formulated together with a polyoxyethylene ether or ester, in either a non-particulate solution or suspension, or in a particulate structure such as a paucilamellar liposome or ISCOM. The saponins may also be formulated with excipients such as Carbopol® to increase viscosity, or may be formulated in a dry powder form with a powder excipient such as lactose.

In one preferred embodiment, the adjuvant system includes the combination of a monophosphoryl lipid A and a saponin derivative, such as the combination of QS21 and 3D-MPL® adjuvant, as described in WO 94/00153, or a less reactogenic composition where the QS21 is quenched with cholesterol, as described in WO 96/33739. Other preferred formulations comprise an oil-in-water emulsion and tocopherol. Another particularly preferred adjuvant formulation employing QS21, 3D-MPL® adjuvant and tocopherol in an oil-in-water emulsion is described in WO 95/17210.

Another enhanced adjuvant system involves the combination of a CpG-containing oligonucleotide and a saponin derivative particularly the combination of CpG and QS21 is disclosed in WO 00/09159. Preferably the formulation additionally comprises an oil in water emulsion and tocopherol.

Additional illustrative adjuvants for use in the pharmaceutical compositions of the invention include Montanide ISA 720 (Seppic, France), SAF (Chiron, California, United States), ISCOMS (CSI), MF-59 (Chiron), the SBAS series of adjuvants (e.g., SBAS-2 or SBAS-4, available from SmithKline Beecham, Rixensart, Belgium), Detox (Enhancyn®) (Corixa, Hamilton, MT), RC-529 (Corixa, Hamilton, MT) and other aminoalkyl glucosaminide 4-phosphates (AGPs), such as those described

WO 01/90152

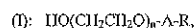
PCT/US01/16776

72

in pending U.S. Patent Application Serial Nos. 08/853,826 and 09/074,720, the disclosures of which are incorporated herein by reference in their entireties, and polyoxyethylene ether adjuvants such as those described in WO 99/52549A1.

Other preferred adjuvants include adjuvant molecules of the general

5 formula



wherein, n is 1-50, A is a bond or $-\text{C}(\text{O})-$, R is C_{1-50} alkyl or Phenyl C_{1-50} alkyl.

One embodiment of the present invention consists of a vaccine formulation comprising a polyoxyethylene ether of general formula (I), wherein n is
 10 between 1 and 50, preferably 4-24, most preferably 9; the R component is C_{1-50} , preferably $\text{C}_4\text{-C}_{20}$ alkyl and most preferably C_{12} alkyl, and A is a bond. The concentration of the polyoxyethylene ethers should be in the range 0.1-20%, preferably from 0.1-10%, and most preferably in the range 0.1-1%. Preferred polyoxyethylene ethers are selected from the following group: polyoxyethylene-9-lauryl ether,
 15 polyoxyethylene-9-stearyl ether, polyoxyethylene-8-stearyl ether, polyoxyethylene-4-lauryl ether, polyoxyethylene-35-lauryl ether, and polyoxyethylene-23-lauryl ether. Polyoxyethylene ethers such as polyoxyethylene lauryl ether are described in the Merck index (12th edition: entry 7717). These adjuvant molecules are described in WO 99/52549.

20 The polyoxyethylene ether according to the general formula (I) above may, if desired, be combined with another adjuvant. For example, a preferred adjuvant combination is preferably with CpG as described in the pending UK patent application GB 9820956.2.

According to another embodiment of this invention, an immunogenic
 25 composition described herein is delivered to a host via antigen presenting cells (APCs), such as dendritic cells, macrophages, B cells, monocytes and other cells that may be engineered to be efficient APCs. Such cells may, but need not, be genetically modified to increase the capacity for presenting the antigen, to improve activation and/or maintenance of the T cell response, to have anti-tumor effects *per se* and/or to be
 30 immunologically compatible with the receiver (*i.e.*, matched HLA haplotype). APCs

WO 01/90152

PCT/US01/16776

73

may generally be isolated from any of a variety of biological fluids and organs, including tumor and peritumoral tissues, and may be autologous, allogeneic, syngeneic or xenogeneic cells.

Certain preferred embodiments of the present invention use dendritic cells or progenitors thereof as antigen-presenting cells. Dendritic cells are highly potent APCs (Banchereau and Steinman, *Nature* 392:245-251, 1998) and have been shown to be effective as a physiological adjuvant for eliciting prophylactic or therapeutic antitumor immunity (see Timmerman and Levy, *Ann. Rev. Med.* 50:507-529, 1999). In general, dendritic cells may be identified based on their typical shape (stellate *in situ*,
10 with marked cytoplasmic processes (dendrites) visible *in vitro*), their ability to take up, process and present antigens with high efficiency and their ability to activate naïve T cell responses. Dendritic cells may, of course, be engineered to express specific cell-surface receptors or ligands that are not commonly found on dendritic cells *in vivo* or *ex vivo*, and such modified dendritic cells are contemplated by the present invention. As
15 an alternative to dendritic cells, secreted vesicles antigen-loaded dendritic cells (called exosomes) may be used within a vaccine (see Zitvogel et al., *Nature Med.* 4:594-600, 1998).

Dendritic cells and progenitors may be obtained from peripheral blood, bone marrow, tumor-infiltrating cells, peritumoral tissues-infiltrating cells, lymph
20 nodes, spleen, skin, umbilical cord blood or any other suitable tissue or fluid. For example, dendritic cells may be differentiated *ex vivo* by adding a combination of cytokines such as GM-CSF, IL-4, IL-13 and/or TNF α to cultures of monocytes harvested from peripheral blood. Alternatively, CD34 positive cells harvested from peripheral blood, umbilical cord blood or bone marrow may be differentiated into
25 dendritic cells by adding to the culture medium combinations of GM-CSF, IL-3, TNF α , CD40 ligand, LPS, flt3 ligand and/or other compound(s) that induce differentiation, maturation and proliferation of dendritic cells.

Dendritic cells are conveniently categorized as "immature" and "mature" cells, which allows a simple way to discriminate between two well characterized
30 phenotypes. However, this nomenclature should not be construed to exclude all

possible intermediate stages of differentiation. Immature dendritic cells are characterized as APC with a high capacity for antigen uptake and processing, which correlates with the high expression of Fcγ receptor and mannose receptor. The mature phenotype is typically characterized by a lower expression of these markers, but a high expression of cell surface molecules responsible for T cell activation such as class I and class II MHC, adhesion molecules (*e.g.*, CD54 and CD11) and costimulatory molecules (*e.g.*, CD40, CD80, CD86 and 4-1BB).

APCs may generally be transfected with a polynucleotide of the invention (or portion or other variant thereof) such that the encoded polypeptide, or an immunogenic portion thereof, is expressed on the cell surface. Such transfection may take place *ex vivo*, and a pharmaceutical composition comprising such transfected cells may then be used for therapeutic purposes, as described herein. Alternatively, a gene delivery vehicle that targets a dendritic or other antigen presenting cell may be administered to a patient, resulting in transfection that occurs *in vivo*. *In vivo* and *ex vivo* transfection of dendritic cells, for example, may generally be performed using any methods known in the art, such as those described in WO 97/24447, or the gene gun approach described by Mahvi et al., *Immunology and cell Biology* 75:456-460, 1997. Antigen loading of dendritic cells may be achieved by incubating dendritic cells or progenitor cells with the tumor polypeptide, DNA (naked or within a plasmid vector) or RNA; or with antigen-expressing recombinant bacterium or viruses (*e.g.*, vaccinia, fowlpox, adenovirus or lentivirus vectors). Prior to loading, the polypeptide may be covalently conjugated to an immunological partner that provides T cell help (*e.g.*, a carrier molecule). Alternatively, a dendritic cell may be pulsed with a non-conjugated immunological partner, separately or in the presence of the polypeptide.

While any suitable carrier known to those of ordinary skill in the art may be employed in the pharmaceutical compositions of this invention, the type of carrier will typically vary depending on the mode of administration. Compositions of the present invention may be formulated for any appropriate manner of administration, including for example, topical, oral, nasal, mucosal, intravenous, intracranial, intraperitoneal, subcutaneous and intramuscular administration.

WO 01/90152

PCT/US01/16776

75

Carriers for use within such pharmaceutical compositions are biocompatible, and may also be biodegradable. In certain embodiments, the formulation preferably provides a relatively constant level of active component release. In other embodiments, however, a more rapid rate of release immediately upon administration may be desired. The formulation of such compositions is well within the level of ordinary skill in the art using known techniques. Illustrative carriers useful in this regard include microparticles of poly(lactide-co-glycolide), polyacrylate, latex, starch, cellulose, dextran and the like. Other illustrative delayed-release carriers include supramolecular biovectors, which comprise a non-liquid hydrophilic core (e.g., a cross-linked polysaccharide or oligosaccharide) and, optionally, an external layer comprising an amphiphilic compound, such as a phospholipid (*see e.g.*, U.S. Patent No. 5,151,254 and PCT applications WO 94/20078, WO/94/23701 and WO 96/06638). The amount of active compound contained within a sustained release formulation depends upon the site of implantation, the rate and expected duration of release and the nature of the condition to be treated or prevented.

In another illustrative embodiment, biodegradable microspheres (e.g., polylactate polyglycolate) are employed as carriers for the compositions of this invention. Suitable biodegradable microspheres are disclosed, for example, in U.S. Patent Nos. 4,897,268; 5,075,109; 5,928,647; 5,811,128; 5,820,883; 5,853,763; 5,814,344, 5,407,609 and 5,942,252. Modified hepatitis B core protein carrier systems, such as described in WO/99 40934, and references cited therein, will also be useful for many applications. Another illustrative carrier/delivery system employs a carrier comprising particulate-protein complexes, such as those described in U.S. Patent No. 5,928,647, which are capable of inducing a class I-restricted cytotoxic T lymphocyte responses in a host.

The pharmaceutical compositions of the invention will often further comprise one or more buffers (e.g., neutral buffered saline or phosphate buffered saline), carbohydrates (e.g., glucose, mannose, sucrose or dextrans), mannitol, proteins, polypeptides or amino acids such as glycine, antioxidants, bacteriostats, chelating agents such as EDTA or glutathione, adjuvants (e.g., aluminum hydroxide), solutes that

WO 01/90152

PCT/US01/16776

76

render the formulation isotonic, hypotonic or weakly hypertonic with the blood of a recipient, suspending agents, thickening agents and/or preservatives. Alternatively, compositions of the present invention may be formulated as a lyophilizate.

The pharmaceutical compositions described herein may be presented in unit-dose or multi-dose containers, such as sealed ampoules or vials. Such containers are typically sealed in such a way to preserve the sterility and stability of the formulation until use. In general, formulations may be stored as suspensions, solutions or emulsions in oily or aqueous vehicles. Alternatively, a pharmaceutical composition may be stored in a freeze-dried condition requiring only the addition of a sterile liquid carrier immediately prior to use.

The development of suitable dosing and treatment regimens for using the particular compositions described herein in a variety of treatment regimens, including e.g., oral, parenteral, intravenous, intranasal, and intramuscular administration and formulation, is well known in the art, some of which are briefly discussed below for general purposes of illustration.

In certain applications, the pharmaceutical compositions disclosed herein may be delivered via oral administration to an animal. As such, these compositions may be formulated with an inert diluent or with an assimilable edible carrier, or they may be enclosed in hard- or soft-shell gelatin capsule, or they may be compressed into tablets, or they may be incorporated directly with the food of the diet.

The active compounds may even be incorporated with excipients and used in the form of ingestible tablets, buccal tablets, troches, capsules, elixirs, suspensions, syrups, wafers, and the like (see, for example, Mathiowitz *et al.*, *Nature* 1997 Mar 27;386(6623):410-4; Hwang *et al.*, *Crit Rev Ther Drug Carrier Syst* 1998;15(3):243-84; U. S. Patent 5,641,515; U. S. Patent 5,580,579 and U. S. Patent 5,792,451). Tablets, troches, pills, capsules and the like may also contain any of a variety of additional components, for example, a binder, such as gum tragacanth, acacia, cornstarch, or gelatin; excipients, such as dicalcium phosphate; a disintegrating agent, such as corn starch, potato starch, alginic acid and the like; a lubricant, such as magnesium stearate; and a sweetening agent, such as sucrose, lactose or saccharin may

WO 01/90152

PCT/US01/16776

77

be added or a flavoring agent, such as peppermint, oil of wintergreen, or cherry flavoring. When the dosage unit form is a capsule, it may contain, in addition to materials of the above type, a liquid carrier. Various other materials may be present as coatings or to otherwise modify the physical form of the dosage unit. For instance, 5 tablets, pills, or capsules may be coated with shellac, sugar, or both. Of course, any material used in preparing any dosage unit form should be pharmaceutically pure and substantially non-toxic in the amounts employed. In addition, the active compounds may be incorporated into sustained-release preparation and formulations.

Typically, these formulations will contain at least about 0.1% of the 10 active compound or more, although the percentage of the active ingredient(s) may, of course, be varied and may conveniently be between about 1 or 2% and about 60% or 70% or more of the weight or volume of the total formulation. Naturally, the amount of active compound(s) in each therapeutically useful composition may be prepared in such a way that a suitable dosage will be obtained in any given unit dose of the compound. 15 Factors such as solubility, bioavailability, biological half-life, route of administration, product shelf life, as well as other pharmacological considerations will be contemplated by one skilled in the art of preparing such pharmaceutical formulations, and as such, a variety of dosages and treatment regimens may be desirable.

For oral administration the compositions of the present invention may 20 alternatively be incorporated with one or more excipients in the form of a mouthwash, dentifrice, buccal tablet, oral spray, or sublingual orally-administered formulation. Alternatively, the active ingredient may be incorporated into an oral solution such as one containing sodium borate, glycerin and potassium bicarbonate, or dispersed in a dentifrice, or added in a therapeutically-effective amount to a composition that may 25 include water, binders, abrasives, flavoring agents, foaming agents, and humectants. Alternatively the compositions may be fashioned into a tablet or solution form that may be placed under the tongue or otherwise dissolved in the mouth.

In certain circumstances it will be desirable to deliver the pharmaceutical compositions disclosed herein parenterally, intravenously, intramuscularly, or even 30 intraperitoneally. Such approaches are well known to the skilled artisan, some of which

WO 01/90152

PCT/US01/16776

78

are further described, for example, in U. S. Patent 5,543,158; U. S. Patent 5,641,515 and U. S. Patent 5,399,363. In certain embodiments, solutions of the active compounds as free base or pharmacologically acceptable salts may be prepared in water suitably mixed with a surfactant, such as hydroxypropylcellulose. Dispersions may also be prepared in glycerol, liquid polyethylene glycols, and mixtures thereof and in oils. Under ordinary conditions of storage and use, these preparations generally will contain a preservative to prevent the growth of microorganisms.

Illustrative pharmaceutical forms suitable for injectable use include sterile aqueous solutions or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersions (for example, see U. S. Patent 5,466,468). In all cases the form must be sterile and must be fluid to the extent that easy syringability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms, such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (e.g., glycerol, propylene glycol, and liquid polyethylene glycol, and the like), suitable mixtures thereof, and/or vegetable oils. Proper fluidity may be maintained, for example, by the use of a coating, such as lecithin, by the maintenance of the required particle size in the case of dispersion and/or by the use of surfactants. The prevention of the action of microorganisms can be facilitated by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, sorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars or sodium chloride. Prolonged absorption of the injectable compositions can be brought about by the use in the compositions of agents delaying absorption, for example, aluminum monostearate and gelatin.

In one embodiment, for parenteral administration in an aqueous solution, the solution should be suitably buffered if necessary and the liquid diluent first rendered isotonic with sufficient saline or glucose. These particular aqueous solutions are especially suitable for intravenous, intramuscular, subcutaneous and intraperitoneal administration. In this connection, a sterile aqueous medium that can be employed will

WO 01/90152

PCT/US01/16776

79

be known to those of skill in the art in light of the present disclosure. For example, one dosage may be dissolved in 1 ml of isotonic NaCl solution and either added to 1000 ml of hypodermoclysis fluid or injected at the proposed site of infusion, (see for example, "Remington's Pharmaceutical Sciences" 15th Edition, pages 1035-1038 and 1570-1580). Some variation in dosage will necessarily occur depending on the condition of the subject being treated. Moreover, for human administration, preparations will of course preferably meet sterility, pyrogenicity, and the general safety and purity standards as required by FDA Office of Biologics standards.

In another embodiment of the invention, the compositions disclosed herein may be formulated in a neutral or salt form. Illustrative pharmaceutically-acceptable salts include the acid addition salts (formed with the free amino groups of the protein) and which are formed with inorganic acids such as, for example, hydrochloric or phosphoric acids, or such organic acids as acetic, oxalic, tartaric, mandelic, and the like. Salts formed with the free carboxyl groups can also be derived from inorganic bases such as, for example, sodium, potassium, ammonium, calcium, or ferric hydroxides, and such organic bases as isopropylamine, trimethylamine, histidine, procaine and the like. Upon formulation, solutions will be administered in a manner compatible with the dosage formulation and in such amount as is therapeutically effective.

The carriers can further comprise any and all solvents, dispersion media, vehicles, coatings, diluents, antibacterial and antifungal agents, isotonic and absorption delaying agents, buffers, carrier solutions, suspensions, colloids, and the like. The use of such media and agents for pharmaceutical active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active ingredient, its use in the therapeutic compositions is contemplated. Supplementary active ingredients can also be incorporated into the compositions. The phrase "pharmaceutically-acceptable" refers to molecular entities and compositions that do not produce an allergic or similar untoward reaction when administered to a human.

In certain embodiments, the pharmaceutical compositions may be delivered by intranasal sprays, inhalation, and/or other aerosol delivery vehicles.

WO 01/90152

PCT/US01/16776

80

Methods for delivering genes, nucleic acids, and peptide compositions directly to the lungs via nasal aerosol sprays has been described, e.g., in U. S. Patent 5,756,353 and U. S. Patent 5,804,212. Likewise, the delivery of drugs using intranasal microparticle resins (Takenaga *et al.*, J Controlled Release 1998 Mar 2;52(1-2):81-7) and lysophosphatidyl-glycerol compounds (U. S. Patent 5,725,871) are also well-known in the pharmaceutical arts. Likewise, illustrative transmucosal drug delivery in the form of a polytetrafluoroethylene support matrix is described in U. S. Patent 5,780,045.

In certain embodiments, liposomes, nanocapsules, microparticles, lipid particles, vesicles, and the like, are used for the introduction of the compositions of the present invention into suitable host cells/organisms. In particular, the compositions of the present invention may be formulated for delivery either encapsulated in a lipid particle, a liposome, a vesicle, a nanosphere, or a nanoparticle or the like. Alternatively, compositions of the present invention can be bound, either covalently or non-covalently, to the surface of such carrier vehicles.

The formation and use of liposome and liposome-like preparations as potential drug carriers is generally known to those of skill in the art (see for example, Jasic, Trends Biotechnol 1998 Jul;16(7):307-21; Takakura, Nippon Rinsho 1998 Mar;56(3):691-5; Chandran *et al.*, Indian J Exp Biol. 1997 Aug;35(8):801-9; Margalit, Crit Rev Ther Drug Carrier Syst. 1995;12(2-3):233-61; U.S. Patent 5,567,434; U.S. Patent 5,352,157; U.S. Patent 5,565,213; U.S. Patent 5,738,868 and U.S. Patent 5,795,587, each specifically incorporated herein by reference in its entirety).

Liposomes have been used successfully with a number of cell types that are normally difficult to transfect by other procedures, including T cell suspensions, primary hepatocyte cultures and PC 12 cells (Remmeisen *et al.*, J Biol Chem. 1990 Sep 25;265(27):16337-42; Muller *et al.*, DNA Cell Biol. 1990 Apr;9(3):221-9). In addition, liposomes are free of the DNA length constraints that are typical of viral-based delivery systems. Liposomes have been used effectively to introduce genes, various drugs, radiotherapeutic agents, enzymes, viruses, transcription factors, allosteric effectors and the like, into a variety of cultured cell lines and animals. Furthermore, the use of

WO 01/90152

PCT/US01/16776

81

liposomes does not appear to be associated with autoimmune responses or unacceptable toxicity after systemic delivery.

In certain embodiments, liposomes are formed from phospholipids that are dispersed in an aqueous medium and spontaneously form multilamellar concentric bilayer vesicles (also termed multilamellar vesicles (MLVs)).

Alternatively, in other embodiments, the invention provides for pharmaceutically-acceptable nanocapsule formulations of the compositions of the present invention. Nanocapsules can generally entrap compounds in a stable and reproducible way (see, for example, Quintanar-Guerrero *et al.*, Drug Dev Ind Pharm. 10 1998 Dec;24(12):1113-28). To avoid side effects due to intracellular polymeric overloading, such ultrafine particles (sized around 0.1 μm) may be designed using polymers able to be degraded *in vivo*. Such particles can be made as described, for example, by Couvreur *et al.*, Crit Rev Ther Drug Carrier Syst. 1988;5(1):1-20; zur Muhlen *et al.*, Eur J Pharm Biopharm. 1998 Mar;45(2):149-55; Zambaux *et al.* J 15 Controlled Release. 1998 Jan 2;50(1-3):31-40; and U. S. Patent 5,145,684.

Cancer Therapeutic Methods

In further aspects of the present invention, the pharmaceutical compositions described herein may be used for the treatment of cancer, particularly for the immunotherapy of breast cancer. Within such methods, the pharmaceutical 20 compositions described herein are administered to a patient, typically a warm-blooded animal, preferably a human. A patient may or may not be afflicted with cancer. Accordingly, the above pharmaceutical compositions may be used to prevent the development of a cancer or to treat a patient afflicted with a cancer. Pharmaceutical compositions and vaccines may be administered either prior to or following surgical 25 removal of primary tumors and/or treatment such as administration of radiotherapy or conventional chemotherapeutic drugs. As discussed above, administration of the pharmaceutical compositions may be by any suitable method, including administration by intravenous, intraperitoneal, intramuscular, subcutaneous, intranasal, intradermal, anal, vaginal, topical and oral routes.

WO 01/90152

PCT/US01/16776

82

Within certain embodiments, immunotherapy may be active immunotherapy, in which treatment relies on the *in vivo* stimulation of the endogenous host immune system to react against tumors with the administration of immune response-modifying agents (such as polypeptides and polynucleotides as provided
5 herein).

Within other embodiments, immunotherapy may be passive immunotherapy, in which treatment involves the delivery of agents with established tumor-immune reactivity (such as effector cells or antibodies) that can directly or indirectly mediate antitumor effects and does not necessarily depend on an intact host
10 immune system. Examples of effector cells include T cells as discussed above, T lymphocytes (such as CD8⁺ cytotoxic T lymphocytes and CD4⁺ T-helper tumor-infiltrating lymphocytes), killer cells (such as Natural Killer cells and lymphokine-activated killer cells), B cells and antigen-presenting cells (such as dendritic cells and macrophages) expressing a polypeptide provided herein. T cell receptors and antibody
15 receptors specific for the polypeptides recited herein may be cloned, expressed and transferred into other vectors or effector cells for adoptive immunotherapy. The polypeptides provided herein may also be used to generate antibodies or anti-idiotypic antibodies (as described above and in U.S. Patent No. 4,918,164) for passive immunotherapy.

Effector cells may generally be obtained in sufficient quantities for adoptive immunotherapy by growth *in vitro*, as described herein. Culture conditions for expanding single antigen-specific effector cells to several billion in number with retention of antigen recognition *in vivo* are well known in the art. Such *in vitro* culture conditions typically use intermittent stimulation with antigen, often in the presence of
25 cytokines (such as IL-2) and non-dividing feeder cells. As noted above, immunoreactive polypeptides as provided herein may be used to rapidly expand antigen-specific T cell cultures in order to generate a sufficient number of cells for immunotherapy. In particular, antigen-presenting cells, such as dendritic, macrophage, monocyte, fibroblast and/or B cells, may be pulsed with immunoreactive polypeptides
30 or transfected with one or more polynucleotides using standard techniques well known

WO 01/90152

PCT/US01/16776

83

in the art. For example, antigen-presenting cells can be transfected with a polynucleotide having a promoter appropriate for increasing expression in a recombinant virus or other expression system. Cultured effector cells for use in therapy must be able to grow and distribute widely, and to survive long term *in vivo*. Studies have shown that cultured effector cells can be induced to grow *in vivo* and to survive long term in substantial numbers by repeated stimulation with antigen supplemented with IL-2 (see, for example, Cheever et al., *Immunological Reviews* 137:177, 1997).

Alternatively, a vector expressing a polypeptide recited herein may be introduced into antigen presenting cells taken from a patient and clonally propagated *ex vivo* for transplant back into the same patient. Transfected cells may be reintroduced into the patient using any means known in the art, preferably in sterile form by intravenous, intracavitary, intraperitoneal or intratumor administration.

Routes and frequency of administration of the therapeutic compositions described herein, as well as dosage, will vary from individual to individual, and may be readily established using standard techniques. In general, the pharmaceutical compositions and vaccines may be administered by injection (e.g., intracutaneous, intramuscular, intravenous or subcutaneous), intranasally (e.g., by aspiration) or orally. Preferably, between 1 and 10 doses may be administered over a 52 week period. Preferably, 6 doses are administered, at intervals of 1 month, and booster vaccinations may be given periodically thereafter. Alternate protocols may be appropriate for individual patients. A suitable dose is an amount of a compound that, when administered as described above, is capable of promoting an anti-tumor immune response, and is at least 10-50% above the basal (*i.e.*, untreated) level. Such response can be monitored by measuring the anti-tumor antibodies in a patient or by vaccine-dependent generation of cytolytic effector cells capable of killing the patient's tumor cells *in vitro*. Such vaccines should also be capable of causing an immune response that leads to an improved clinical outcome (e.g., more frequent remissions, complete or partial or longer disease-free survival) in vaccinated patients as compared to non-vaccinated patients. In general, for pharmaceutical compositions and vaccines comprising one or more polypeptides, the amount of each polypeptide present in a dose

WO 01/90152

PCT/US01/16776

84

ranges from about 25 µg to 5 mg per kg of host. Suitable dose sizes will vary with the size of the patient, but will typically range from about 0.1 mL to about 5 mL.

In general, an appropriate dosage and treatment regimen provides the active compound(s) in an amount sufficient to provide therapeutic and/or prophylactic benefit. Such a response can be monitored by establishing an improved clinical outcome (e.g., more frequent remissions, complete or partial, or longer disease-free survival) in treated patients as compared to non-treated patients. Increases in preexisting immune responses to a tumor protein generally correlate with an improved clinical outcome. Such immune responses may generally be evaluated using standard proliferation, cytotoxicity or cytokine assays, which may be performed using samples obtained from a patient before and after treatment.

Cancer Detection and Diagnostic Compositions, Methods and Kits

In general, a cancer may be detected in a patient based on the presence of one or more breast tumor proteins and/or polynucleotides encoding such proteins in a biological sample (for example, blood, sera, sputum urine and/or tumor biopsies) obtained from the patient. In other words, such proteins may be used as markers to indicate the presence or absence of a cancer such as breast cancer. In addition, such proteins may be useful for the detection of other cancers. The binding agents provided herein generally permit detection of the level of antigen that binds to the agent in the biological sample. Polynucleotide primers and probes may be used to detect the level of mRNA encoding a tumor protein, which is also indicative of the presence or absence of a cancer. In general, a breast tumor sequence should be present at a level that is at least three fold higher in tumor tissue than in normal tissue.

There are a variety of assay formats known to those of ordinary skill in the art for using a binding agent to detect polypeptide markers in a sample. See, e.g., Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988. In general, the presence or absence of a cancer in a patient may be determined by (a) contacting a biological sample obtained from a patient with a binding agent; (b)

WO 01/90152

PCT/US01/16776

85

detecting in the sample a level of polypeptide that binds to the binding agent; and (c) comparing the level of polypeptide with a predetermined cut-off value.

In a preferred embodiment, the assay involves the use of binding agent immobilized on a solid support to bind to and remove the polypeptide from the remainder of the sample. The bound polypeptide may then be detected using a detection reagent that contains a reporter group and specifically binds to the binding agent/polypeptide complex. Such detection reagents may comprise, for example, a binding agent that specifically binds to the polypeptide or an antibody or other agent that specifically binds to the binding agent, such as an anti-immunoglobulin, protein G, protein A or a lectin. Alternatively, a competitive assay may be utilized, in which a polypeptide is labeled with a reporter group and allowed to bind to the immobilized binding agent after incubation of the binding agent with the sample. The extent to which components of the sample inhibit the binding of the labeled polypeptide to the binding agent is indicative of the reactivity of the sample with the immobilized binding agent. Suitable polypeptides for use within such assays include full length breast tumor proteins and polypeptide portions thereof to which the binding agent binds, as described above.

The solid support may be any material known to those of ordinary skill in the art to which the tumor protein may be attached. For example, the solid support may be a test well in a microtiter plate or a nitrocellulose or other suitable membrane. Alternatively, the support may be a bead or disc, such as glass, fiberglass, latex or a plastic material such as polystyrene or polyvinylchloride. The support may also be a magnetic particle or a fiber optic sensor, such as those disclosed, for example, in U.S. Patent No. 5,359,681. The binding agent may be immobilized on the solid support using a variety of techniques known to those of skill in the art, which are amply described in the patent and scientific literature. In the context of the present invention, the term "immobilization" refers to both noncovalent association, such as adsorption, and covalent attachment (which may be a direct linkage between the agent and functional groups on the support or may be a linkage by way of a cross-linking agent). Immobilization by adsorption to a well in a microtiter plate or to a membrane is

WO 01/90152

PCT/US01/16776

86

preferred. In such cases, adsorption may be achieved by contacting the binding agent, in a suitable buffer, with the solid support for a suitable amount of time. The contact time varies with temperature, but is typically between about 1 hour and about 1 day. In general, contacting a well of a plastic microtiter plate (such as polystyrene or polyvinylchloride) with an amount of binding agent ranging from about 10 ng to about 10 μ g, and preferably about 100 ng to about 1 μ g, is sufficient to immobilize an adequate amount of binding agent.

Covalent attachment of binding agent to a solid support may generally be achieved by first reacting the support with a bifunctional reagent that will react with both the support and a functional group, such as a hydroxyl or amino group, on the binding agent. For example, the binding agent may be covalently attached to supports having an appropriate polymer coating using benzoquinone or by condensation of an aldehyde group on the support with an amine and an active hydrogen on the binding partner (*see, e.g.,* Pierce Immunotechnology Catalog and Handbook, 1991, at A12-A13).

In certain embodiments, the assay is a two-antibody sandwich assay. This assay may be performed by first contacting an antibody that has been immobilized on a solid support, commonly the well of a microtiter plate, with the sample, such that polypeptides within the sample are allowed to bind to the immobilized antibody. Unbound sample is then removed from the immobilized polypeptide-antibody complexes and a detection reagent (preferably a second antibody capable of binding to a different site on the polypeptide) containing a reporter group is added. The amount of detection reagent that remains bound to the solid support is then determined using a method appropriate for the specific reporter group.

More specifically, once the antibody is immobilized on the support as described above, the remaining protein binding sites on the support are typically blocked. Any suitable blocking agent known to those of ordinary skill in the art, such as bovine serum albumin or Tween 20TM (Sigma Chemical Co., St. Louis, MO). The immobilized antibody is then incubated with the sample, and polypeptide is allowed to bind to the antibody. The sample may be diluted with a suitable diluent, such as

phosphate-buffered saline (PBS) prior to incubation. In general, an appropriate contact time (*i.e.*, incubation time) is a period of time that is sufficient to detect the presence of polypeptide within a sample obtained from an individual with breast cancer. Preferably, the contact time is sufficient to achieve a level of binding that is at least about 95% of that achieved at equilibrium between bound and unbound polypeptide. Those of ordinary skill in the art will recognize that the time necessary to achieve equilibrium may be readily determined by assaying the level of binding that occurs over a period of time. At room temperature, an incubation time of about 30 minutes is generally sufficient.

10 Unbound sample may then be removed by washing the solid support with an appropriate buffer, such as PBS containing 0.1% Tween 20TM. The second antibody, which contains a reporter group, may then be added to the solid support. Preferred reporter groups include those groups recited above.

The detection reagent is then incubated with the immobilized antibody-polypeptide complex for an amount of time sufficient to detect the bound polypeptide. 15 An appropriate amount of time may generally be determined by assaying the level of binding that occurs over a period of time. Unbound detection reagent is then removed and bound detection reagent is detected using the reporter group. The method employed for detecting the reporter group depends upon the nature of the reporter group. For 20 radioactive groups, scintillation counting or autoradiographic methods are generally appropriate. Spectroscopic methods may be used to detect dyes, luminescent groups and fluorescent groups. Biotin may be detected using avidin, coupled to a different reporter group (commonly a radioactive or fluorescent group or an enzyme). Enzyme reporter groups may generally be detected by the addition of substrate (generally for a 25 specific period of time), followed by spectroscopic or other analysis of the reaction products.

To determine the presence or absence of a cancer, such as breast cancer, the signal detected from the reporter group that remains bound to the solid support is generally compared to a signal that corresponds to a predetermined cut-off value. In 30 one preferred embodiment, the cut-off value for the detection of a cancer is the average

mean signal obtained when the immobilized antibody is incubated with samples from patients without the cancer. In general, a sample generating a signal that is three standard deviations above the predetermined cut-off value is considered positive for the cancer. In an alternate preferred embodiment, the cut-off value is determined using a Receiver Operator Curve, according to the method of Sackett et al., *Clinical Epidemiology: A Basic Science for Clinical Medicine*, Little Brown and Co., 1985, p. 106-7. Briefly, in this embodiment, the cut-off value may be determined from a plot of pairs of true positive rates (*i.e.*, sensitivity) and false positive rates (100% -specificity) that correspond to each possible cut-off value for the diagnostic test result. The cut-off value on the plot that is the closest to the upper left-hand corner (*i.e.*, the value that encloses the largest area) is the most accurate cut-off value, and a sample generating a signal that is higher than the cut-off value determined by this method may be considered positive. Alternatively, the cut-off value may be shifted to the left along the plot, to minimize the false positive rate, or to the right, to minimize the false negative rate. In general, a sample generating a signal that is higher than the cut-off value determined by this method is considered positive for a cancer.

In a related embodiment, the assay is performed in a flow-through or strip test format, wherein the binding agent is immobilized on a membrane, such as nitrocellulose. In the flow-through test, polypeptides within the sample bind to the immobilized binding agent as the sample passes through the membrane. A second, labeled binding agent then binds to the binding agent-polypeptide complex as a solution containing the second binding agent flows through the membrane. The detection of bound second binding agent may then be performed as described above. In the strip test format, one end of the membrane to which binding agent is bound is immersed in a solution containing the sample. The sample migrates along the membrane through a region containing second binding agent and to the area of immobilized binding agent. Concentration of second binding agent at the area of immobilized antibody indicates the presence of a cancer. Typically, the concentration of second binding agent at that site generates a pattern, such as a line, that can be read visually. The absence of such a pattern indicates a negative result. In general, the amount of binding agent immobilized

WO 01/90152

PCT/US01/16776

89

on the membrane is selected to generate a visually discernible pattern when the biological sample contains a level of polypeptide that would be sufficient to generate a positive signal in the two-antibody sandwich assay, in the format discussed above. Preferred binding agents for use in such assays are antibodies and antigen-binding fragments thereof. Preferably, the amount of antibody immobilized on the membrane ranges from about 25 ng to about 1 µg, and more preferably from about 50 ng to about 500 ng. Such tests can typically be performed with a very small amount of biological sample.

Of course, numerous other assay protocols exist that are suitable for use with the tumor proteins or binding agents of the present invention. The above descriptions are intended to be exemplary only. For example, it will be apparent to those of ordinary skill in the art that the above protocols may be readily modified to use tumor polypeptides to detect antibodies that bind to such polypeptides in a biological sample. The detection of such tumor protein specific antibodies may correlate with the presence of a cancer.

A cancer may also, or alternatively, be detected based on the presence of T cells that specifically react with a tumor protein in a biological sample. Within certain methods, a biological sample comprising CD4⁺ and/or CD8⁺ T cells isolated from a patient is incubated with a tumor polypeptide, a polynucleotide encoding such a polypeptide and/or an APC that expresses at least an immunogenic portion of such a polypeptide, and the presence or absence of specific activation of the T cells is detected. Suitable biological samples include, but are not limited to, isolated T cells. For example, T cells may be isolated from a patient by routine techniques (such as by Ficoll/Hypaque density gradient centrifugation of peripheral blood lymphocytes). T cells may be incubated *in vitro* for 2-9 days (typically 4 days) at 37°C with polypeptide (e.g., 5 - 25 µg/ml). It may be desirable to incubate another aliquot of a T cell sample in the absence of tumor polypeptide to serve as a control. For CD4⁺ T cells, activation is preferably detected by evaluating proliferation of the T cells. For CD8⁺ T cells, activation is preferably detected by evaluating cytolytic activity. A level of proliferation

WO 01/90152

PCT/US01/16776

90

that is at least two fold greater and/or a level of cytolytic activity that is at least 20% greater than in disease-free patients indicates the presence of a cancer in the patient.

As noted above, a cancer may also, or alternatively, be detected based on the level of mRNA encoding a tumor protein in a biological sample. For example, at least two oligonucleotide primers may be employed in a polymerase chain reaction (PCR) based assay to amplify a portion of a tumor cDNA derived from a biological sample, wherein at least one of the oligonucleotide primers is specific for (*i.e.*, hybridizes to) a polynucleotide encoding the tumor protein. The amplified cDNA is then separated and detected using techniques well known in the art, such as gel electrophoresis. Similarly, oligonucleotide probes that specifically hybridize to a polynucleotide encoding a tumor protein may be used in a hybridization assay to detect the presence of polynucleotide encoding the tumor protein in a biological sample.

To permit hybridization under assay conditions, oligonucleotide primers and probes should comprise an oligonucleotide sequence that has at least about 60%, preferably at least about 75% and more preferably at least about 90%, identity to a portion of a polynucleotide encoding a tumor protein of the invention that is at least 10 nucleotides, and preferably at least 20 nucleotides, in length. Preferably, oligonucleotide primers and/or probes hybridize to a polynucleotide encoding a polypeptide described herein under moderately stringent conditions, as defined above. Oligonucleotide primers and/or probes which may be usefully employed in the diagnostic methods described herein preferably are at least 10-40 nucleotides in length. In a preferred embodiment, the oligonucleotide primers comprise at least 10 contiguous nucleotides, more preferably at least 15 contiguous nucleotides, of a DNA molecule having a sequence as disclosed herein. Techniques for both PCR based assays and hybridization assays are well known in the art (*see*, for example, Mullis et al., *Cold Spring Harbor Symp. Quant. Biol.*, 51:263, 1987; Erlich ed., *PCR Technology*, Stockton Press, NY, 1989).

One preferred assay employs RT-PCR, in which PCR is applied in conjunction with reverse transcription. Typically, RNA is extracted from a biological sample, such as biopsy tissue, and is reverse transcribed to produce cDNA molecules.

PCR amplification using at least one specific primer generates a cDNA molecule, which may be separated and visualized using, for example, gel electrophoresis. Amplification may be performed on biological samples taken from a test patient and from an individual who is not afflicted with a cancer. The amplification reaction may be
5 performed on several dilutions of cDNA spanning two orders of magnitude. A two-fold or greater increase in expression in several dilutions of the test patient sample as compared to the same dilutions of the non-cancerous sample is typically considered positive.

In another embodiment, the compositions described herein may be used
10 as markers for the progression of cancer. In this embodiment, assays as described above for the diagnosis of a cancer may be performed over time, and the change in the level of reactive polypeptide(s) or polynucleotide(s) evaluated. For example, the assays may be performed every 24-72 hours for a period of 6 months to 1 year, and thereafter performed as needed. In general, a cancer is progressing in those patients in whom the
15 level of polypeptide or polynucleotide detected increases over time. In contrast, the cancer is not progressing when the level of reactive polypeptide or polynucleotide either remains constant or decreases with time.

Certain *in vivo* diagnostic assays may be performed directly on a tumor. One such assay involves contacting tumor cells with a binding agent. The bound
20 binding agent may then be detected directly or indirectly via a reporter group. Such binding agents may also be used in histological applications. Alternatively, polynucleotide probes may be used within such applications.

As noted above, to improve sensitivity, multiple tumor protein markers may be assayed within a given sample. It will be apparent that binding agents specific
25 for different proteins provided herein may be combined within a single assay. Further, multiple primers or probes may be used concurrently. The selection of tumor protein markers may be based on routine experiments to determine combinations that results in optimal sensitivity. In addition, or alternatively, assays for tumor proteins provided herein may be combined with assays for other known tumor antigens.

The present invention further provides kits for use within any of the above diagnostic methods. Such kits typically comprise two or more components necessary for performing a diagnostic assay. Components may be compounds, reagents, containers and/or equipment. For example, one container within a kit may contain a monoclonal antibody or fragment thereof that specifically binds to a tumor protein. Such antibodies or fragments may be provided attached to a support material, as described above. One or more additional containers may enclose elements, such as reagents or buffers, to be used in the assay. Such kits may also, or alternatively, contain a detection reagent as described above that contains a reporter group suitable for direct or indirect detection of antibody binding.

Alternatively, a kit may be designed to detect the level of mRNA encoding a tumor protein in a biological sample. Such kits generally comprise at least one oligonucleotide probe or primer, as described above, that hybridizes to a polynucleotide encoding a tumor protein. Such an oligonucleotide may be used, for example, within a PCR or hybridization assay. Additional components that may be present within such kits include a second oligonucleotide and/or a diagnostic reagent or container to facilitate the detection of a polynucleotide encoding a tumor protein.

The following Examples are offered by way of illustration and not by way of limitation.

EXAMPLE I

PREPARATION OF BREAST TUMOR-SPECIFIC cDNAs USING DIFFERENTIAL DISPLAY RT-PCR

This Example illustrates the preparation of cDNA molecules encoding breast tumor-specific polypeptides using a differential display screen.

A. Preparation of B18Ag1 cDNA and Characterization of mRNA Expression

Tissue samples were prepared from breast tumor and normal tissue of a patient with breast cancer that was confirmed by pathology after removal from the patient. Normal RNA and tumor RNA was extracted from the samples and mRNA was

WO 01/90152

PCT/US01/16776

93

isolated and converted into cDNA using a (dT)₁₂AG (SEQ ID NO:130) anchored 3' primer. Differential display PCR was then executed using a randomly chosen primer (CTTCAACCCTC) (SEQ ID NO:103). Amplification conditions were standard buffer containing 1.5 mM MgCl₂, 20 pmol of primer, 500 pmol dNTP, and 1 unit of *Taq* DNA polymerase (Perkin-Elmer, Branchburg, NJ). Forty cycles of amplification were performed using 94°C denaturation for 30 seconds, 42°C annealing for 1 minute, and 72°C extension for 30 seconds. An RNA fingerprint containing 76 amplified products was obtained. Although the RNA fingerprint of breast tumor tissue was over 98% identical to that of the normal breast tissue, a band was repeatedly observed to be specific to the RNA fingerprint pattern of the tumor. This band was cut out of a silver stained gel, subcloned into the T-vector (Novagen, Madison, WI) and sequenced.

The sequence of the cDNA, referred to as B18Ag1, is provided in SEQ ID NO:1. A database search of GENBANK and EMBL revealed that the B18Ag1 fragment initially cloned is 77% identical to the endogenous human retroviral element S71, which is a truncated retroviral element homologous to the Simian Sarcoma Virus (SSV). S71 contains an incomplete *gag* gene, a portion of the *pol* gene and an LTR-like structure at the 3' terminus (see Werner et al., *Virology* 174:225-238 (1990)). B18Ag1 is also 64% identical to SSV in the region corresponding to the P30 (*gag*) locus. B18Ag1 contains three separate and incomplete reading frames covering a region which shares considerable homology to a wide variety of *gag* proteins of retroviruses which infect mammals. In addition, the homology to S71 is not just within the *gag* gene, but spans several kb of sequence including an LTR.

B18Ag1-specific PCR primers were synthesized using computer analysis guidelines. RT-PCR amplification (94°C, 30 seconds; 60°C → 42°C, 30 seconds; 72°C, 30 seconds for 40 cycles) confirmed that B18Ag1 represents an actual mRNA sequence present at relatively high levels in the patient's breast tumor tissue. The primers used in amplification were B18Ag1-1 (CTG CCT GAG CCA CAA ATG) (SEQ ID NO:128) and B18Ag1-4 (CCG GAG GAG GAA GCT AGA GGA ATA) (SEQ ID NO:129) at a 3.5 mM magnesium concentration and a pH of 8.5, and B18Ag1-2 (ATG GCT ATT TTC GGG GCC TGA CA) (SEQ ID NO:126) and B18Ag1-3 (CCG GTA TCT CCT

WO 01/90152

PCT/US01/16776

94

CGT GGG TAT T) (SEQ ID NO:127) at 2 mM magnesium at pH 9.5. The same experiments showed exceedingly low to nonexistent levels of expression in this patient's normal breast tissue (see Figure 1). RT-PCR experiments were then used to show that B18Ag1 mRNA is present in nine other breast tumor samples (from Brazilian and American patients) but absent in, or at exceedingly low levels in, the normal breast tissue corresponding to each cancer patient. RT-PCR analysis has also shown that the B18Ag1 transcript is not present in various normal tissues (including lymph node, myocardium and liver) and present at relatively low levels in PBMC and lung tissue. The presence of B18Ag1 mRNA in breast tumor samples, and its absence from normal breast tissue, has been confirmed by Northern blot analysis, as shown in Figure 2.

The differential expression of B18Ag1 in breast tumor tissue was also confirmed by RNase protection assays. Figure 3 shows the level of B18Ag1 mRNA in various tissue types as determined in four different RNase protection assays. Lanes 1-12 represent various normal breast tissue samples, lanes 13-25 represent various breast tumor samples; lanes 26-27 represent normal prostate samples; lanes 28-29 represent prostate tumor samples; lanes 30-32 represent colon tumor samples; lane 33 represents normal aorta; lane 34 represents normal small intestine; lane 35 represents normal skin, lane 36 represents normal lymph node; lane 37 represents normal ovary; lane 38 represents normal liver; lane 39 represents normal skeletal muscle; lane 40 represents a first normal stomach sample, lane 41 represents a second normal stomach sample; lane 42 represents a normal lung; lane 43 represents normal kidney; and lane 44 represents normal pancreas. Interexperimental comparison was facilitated by including a positive control RNA of known β -actin message abundance in each assay and normalizing the results of the different assays with respect to this positive control.

RT-PCR and Southern Blot analysis has shown the B18Ag1 locus to be present in human genomic DNA as a single copy endogenous retroviral element. A genomic clone of approximately 12-18 kb was isolated using the initial B18Ag1 sequence as a probe. Four additional subclones were also isolated by XbaI digestion. Additional retroviral sequences obtained from the ends of the XbaI digests of these clones (located as shown in Figure 4) are shown as SEQ ID NO:3 - SEQ ID NO:10,

WO 01/90152

PCT/US01/16776

95

where SEQ ID NO:3 shows the location of the sequence labeled 10 in Figure 4, SEQ ID NO:4 shows the location of the sequence labeled 11-29, SEQ ID NO:5 shows the location of the sequence labeled 3, SEQ ID NO:6 shows the location of the sequence labeled 6, SEQ ID NO:7 shows the location of the sequence labeled 12, SEQ ID NO:8 shows the location of the sequence labeled 13, SEQ ID NO:9 shows the location of the sequence labeled 14 and SEQ ID NO:10 shows the location of the sequence labeled 11-22.

Subsequent studies demonstrated that the 12-18 kb genomic clone contains a retroviral element of about 7.75 kb, as shown in Figures 5A and 5B. The sequence of this retroviral element is shown in SEQ ID NO:141. The numbered line at the top of Figure 5A represents the sense strand sequence of the retroviral genomic clone. The box below this line shows the position of selected restriction sites. The arrows depict the different overlapping clones used to sequence the retroviral element. The direction of the arrow shows whether the single-pass subclone sequence corresponded to the sense or anti-sense strand. Figure 5B is a schematic diagram of the retroviral element containing B18Ag1 depicting the organization of viral genes within the element. The open boxes correspond to predicted reading frames, starting with a methionine, found throughout the element. Each of the six likely reading frames is shown, as indicated to the left of the boxes, with frames 1-3 corresponding to those found on the sense strand.

Using the cDNA of SEQ ID NO:1 as a probe, a longer cDNA was obtained (SEQ ID NO:227) which contains minor nucleotide differences (less than 1%) compared to the genomic sequence shown in SEQ ID NO:141.

B. Preparation of cDNA Molecules Encoding Other Breast Tumor-Specific Polypeptides

Normal RNA and tumor RNA was prepared and mRNA was isolated and converted into cDNA using a (dT)₁₂AG anchored 3' primer, as described above. Differential display PCR was then executed using the randomly chosen primers of SEQ ID NOs:87-125. Amplification conditions were as noted above, and bands observed to be specific to the RNA fingerprint pattern of the tumor were cut out of a silver stained

WO 01/90152

PCT/US01/16776

96

gel, subcloned into either the T-vector (Novagen, Madison, WI) or the pCRII vector (Invitrogen, San Diego, CA) and sequenced. The sequences are provided in SEQ ID NO:11 - SEQ ID NO:86. Of the 79 sequences isolated, 67 were found to be novel (SEQ ID NOs:11-26 and 28-77) (*see also* Figures 6-20).

5 An extended DNA sequence (SEQ ID NO:290) for the antigen B15Ag1 (originally identified partial sequence provided in SEQ ID NO:27) was obtained in further studies. Comparison of the sequence of SEQ ID NO:290 with those in the gene bank as described above, revealed homology to the known human β -A actinin gene. Further studies led to the isolation of the full-length cDNA sequence for the antigen
10 B21GT2 (also referred to as B311D; originally identified partial cDNA sequence provided in SEQ ID NOs:56). The full-length sequence is provided in SEQ ID NO:307, with the corresponding amino acid sequence being provided in SEQ ID NO:308. Further studies led to the isolation of a splice variant of B311D. The B311D clone of SEQ ID NO:316 was sequenced and a XhoI/NotI fragment from this clone was gel
15 purified and 32P-cDTP labeled by random priming for use as a probe for further screening to obtain additional B311D gene sequence. Two fractions of a human breast tumor cDNA bacterial library were screened using standard techniques. One of the clones isolated in this manner yielded additional sequence which includes a poly A+ tail. The determined cDNA sequence of this clone (referred to as B311D_BT1_1A) is
20 provided in SEQ ID NO:317. The sequences of SEQ ID NOs:316 and 317 were found to share identity over a 464 bp region, with the sequences diverging near the poly A+ sequence of SEQ ID NO:317.

Subsequent studies identified an additional 146 sequences (SEQ ID NOs:142-289), of which 115 appeared to be novel (SEQ ID NOs:142, 143, 146-152,
25 154-166, 168-176, 178-192, 194-198, 200-204, 206, 207, 209-214, 216, 218, 219, 221-240, 243-245, 247, 250, 251, 253, 255, 257-266, 268, 269, 271-273, 275, 276, 278, 280, 281, 284, 288 and 291). To the best of the inventors' knowledge none of the previously identified sequences have heretofore been shown to be expressed at a greater level in human breast tumor tissue than in normal breast tissue.

WO 01/90152

PCT/US01/16776

97

In further studies, several different splice forms of the antigen B11Ag1 (also referred to as B305D) were isolated, with each of the various splice forms containing slightly different versions of the B11Ag1 coding frame. Splice junction sequences define individual exons which, in various patterns and arrangements, make up the various splice forms. Primers were designed to examine the expression pattern of each of the exons using RT-PCR as described below. Each exon was found to show the same expression pattern as the original B11Ag1 clone, with expression being breast tumor-, normal prostate- and normal testis-specific. The determined cDNA sequences for the isolated protein coding exons are provided in SEQ ID NOs:292-298, respectively. The predicted amino acid sequences corresponding to the sequences of SEQ ID NOs:292 and 298 are provided in SEQ ID NOs:299 and 300. Additional studies using rapid amplification of cDNA ends (RACE), a 5' specific primer to one of the splice forms of B11Ag1 provided above and a breast adenocarcinoma, led to the isolation of three additional, related, splice forms referred to as isoforms B11C-15, B11C-8 and B11C-9,16. The determined cDNA sequences for these isoforms are provided in SEQ ID NO: 301-303, with the corresponding predicted amino acid sequences being provided in SEQ ID NOs:304-306.

The protein coding region of B11C-15 (SEQ ID NO: 301; also referred to as B305D isoform C) was used as a query sequence in a BLASTN search of the Genbank DNA database. A match was found to a genomic clone from chromosome 21 (Accession no. AP001465). The pairwise alignments provided in the BLASTN output were used to identify the putative exon, or coding, sequence of the chromosome 21 sequence that corresponds to the B305D sequence. Based on the BlastN pairwise alignments, the following pieces of GenBank record AP001465 were put together: base pairs 67978-68499, 72870-72987, 73144-73335, 76085-76206, 77905-78085, 80520-80624, 87602-87633. This sequence was then aligned with the B305D isoform C sequence using the DNA Star Seqman program and excess sequence was deleted in such a way as to maintain the sequence most similar to B305D. The final edited form of the chromosome 21 sequence was 96.5% identical to B305D. This resulting edited sequence from chromosome 21 was then translated and found to contain no stop codons

WO 01/90152

PCT/US01/16776

98

other than the final stop codon in the same position as that for B305D. As with B305D, the chromosome 21 sequence (provided in SEQ ID NO: 325) encoded a protein (SEQ ID NO: 326) with 384 amino acids. An alignment of this protein with the B305D isoform C protein (SEQ ID NO: 304) showed 90% amino acid identity.

5 The cDNA sequence of B305D isoform C (SEQ ID NO: 301) was used to identify homologs by searching the High Throughput Genome Sequencing (HTGS) database (NCBI, National Institutes for Health, Bethesda, MD). Homologs were identified on Chromosome 2 (Clone ID 9838181), Chromosome 10 (Clone ID 10933022), Chromosome 15 (Clone ID 11560284). These homologs shared greater
10 than 90% identity with B305D isoform C at the nucleic acid level. All three of these homologs encode 384 amino acid ORFs that share greater than 90% identity with the amino acid sequence of SEQ ID NO: 304. Further searching of the GenBank database with the sequence of SEQ ID NO: 301 yielded a partial sequence homolog on Chromosome 22 (Clone ID 5931507). cDNA sequences for the Chromosome 2, 10, 15
15 and 22 homologs were constructed based on the homology with B305D isoform C and the conserved sequences at intron-exon junctions. The cDNA sequences for the Chromosome 22, 2, 15 and 10 homologs are provided in SEQ ID NO: 327-330, respectively, with the corresponding amino acid sequences being provided in SEQ ID NO: 331, 334, 333 and 332, respectively.

20 In subsequent studies on B305D isoform A (cDNA sequence provided in SEQ ID NO:292), the cDNA sequence (provided in SEQ ID NO:313) was found to contain an additional guanine residue at position 884, leading to a frameshift in the open reading frame. The determined DNA sequence of this ORF is provided in SEQ ID NO:314. This frameshift generates a protein sequence (provided in SEQ ID NO:315) of
25 293 amino acids that contains the C-terminal domain common to the other isoforms of B305D but that differs in the N-terminal region.

WO 01/90152

PCT/US01/16776

99

EXAMPLE 2

PREPARATION OF B18Ag1 DNA FROM HUMAN GENOMIC DNA

This Example illustrates the preparation of B18Ag1 DNA by
 5 amplification from human genomic DNA.

B18Ag1 DNA may be prepared from 250 ng human genomic DNA using
 20 pmol of B18Ag1 specific primers, 500 pmol dNTPS and 1 unit of *Taq* DNA
 polymerase (Perkin Elmer, Branchburg, NJ) using the following amplification
 parameters: 94°C for 30 seconds denaturing, 30 seconds 60°C to 42°C touchdown
 10 annealing in 2°C increments every two cycles and 72°C extension for 30 seconds. The
 last increment (a 42°C annealing temperature) should cycle 25 times. Primers were
 selected using computer analysis. Primers synthesized were B18Ag1-1, B18Ag1-2,
 B18Ag1-3, and B18Ag1-4. Primer pairs that may be used are 1+3, 1+4, 2+3, and 2+4.

Following gel electrophoresis, the band corresponding to B18Ag1 DNA
 15 may be excised and cloned into a suitable vector.

EXAMPLE 3

PREPARATION OF B18Ag1 DNA FROM BREAST TUMOR cDNA

This Example illustrates the preparation of B18Ag1 DNA by
 20 amplification from human breast tumor cDNA.

First strand cDNA is synthesized from RNA prepared from human breast
 tumor tissue in a reaction mixture containing 500 ng poly A+ RNA, 200 pmol of the
 primer (T)₁₂AG (*i.e.*, TTT TTT TTT TTT AG) (SEQ ID NO:130), 1X first strand
 25 reverse transcriptase buffer, 6.7 mM DTT, 500 nmol dNTPs, and 1 unit AMV or
 MMLV reverse transcriptase (from any supplier, such as Gibco-BRL (Grand Island,
 NY)) in a final volume of 30 µl. After first strand synthesis, the cDNA is diluted
 approximately 25 fold and 1 µl is used for amplification as described in Example 2.

While some primer pairs can result in a heterogeneous population of transcripts, the
 30 primers B18Ag1-2 (5'ATG GCT ATT TTC GGG GGC TGA CA) (SEQ ID NO:126)

WO 01/90152

PCT/US01/16776

100

and B18Ag1-3 (5'CCG GTA TCT CCT CGT GGG TAT T) (SEQ ID NO:127) yield a single 151 bp amplification product.

EXAMPLE 4

5 IDENTIFICATION OF B-CELL AND T-CELL EPITOPES OF B18Ag1

This Example illustrates the identification of B18Ag1 epitopes.

The B18Ag1 sequence can be screened using a variety of computer algorithms. To determine B-cell epitopes, the sequence can be screened for
10 hydrophobicity and hydrophilicity values using the method of Hopp, *Prog. Clin. Biol. Res.* 172B:367-77 (1985) or, alternatively, Cease et al., *J. Exp. Med.* 164:1779-84 (1986) or Spouge et al., *J. Immunol.* 138:204-12 (1987). Additional Class II MHC (antibody or B-cell) epitopes can be predicted using programs such as AMPHI (e.g., Margalit et al., *J. Immunol.* 138:2213 (1987)) or the methods of Rothbard and Taylor
15 (e.g., *EMBO J.* 7:93 (1988)).

Once peptides (15-20 amino acids long) are identified using these techniques, individual peptides can be synthesized using automated peptide synthesis equipment (available from manufacturers such as Perkin Elmer/Applied Biosystems Division, Foster City, CA) and techniques such as Merrifield synthesis. Following
20 synthesis, the peptides can used to screen sera harvested from either normal or breast cancer patients to determine whether patients with breast cancer possess antibodies reactive with the peptides. Presence of such antibodies in breast cancer patient would confirm the immunogenicity of the specific B-cell epitope in question. The peptides can also be tested for their ability to generate a serologic or humoral immune in animals
25 (mice, rats, rabbits, chimps etc.) following immunization *in vivo*. Generation of a peptide-specific antiserum following such immunization further confirms the immunogenicity of the specific B-cell epitope in question.

To identify T-cell epitopes, the B18Ag1 sequence can be screened using different computer algorithms which are useful in identifying 8-10 amino acid motifs
30 within the B18Ag1 sequence which are capable of binding to HLA Class I MHC

WO 01/90152

PCT/US01/16776

101

molecules. (see, e.g., Rammensee et al., *Immunogenetics* 41:178-228 (1995)). Following synthesis such peptides can be tested for their ability to bind to class I MHC using standard binding assays (e.g., Sette et al., *J. Immunol.* 153:5586-92 (1994)) and more importantly can be tested for their ability to generate antigen reactive cytotoxic T-cells following *in vitro* stimulation of patient or normal peripheral mononuclear cells using, for example, the methods of Bakker et al., *Cancer Res.* 55:5330-34 (1995); Visseren et al., *J. Immunol.* 154:3991-98 (1995); Kawakami et al., *J. Immunol.* 154:3961-68 (1995); and Kast et al., *J. Immunol.* 152:3904-12 (1994). Successful *in vitro* generation of T-cells capable of killing autologous (bearing the same Class I MHC molecules) tumor cells following *in vitro* peptide stimulation further confirms the immunogenicity of the B18Ag1 antigen. Furthermore, such peptides may be used to generate murine peptide and B18Ag1 reactive cytotoxic T-cells following *in vivo* immunization in mice rendered transgenic for expression of a particular human MHC Class I haplotype (Vitiello et al., *J. Exp. Med.* 173:1007-15 (1991)).

15 A representative list of predicted B18Ag1 B-cell and T-cell epitopes, broken down according to predicted HLA Class I MHC binding antigen, is shown below:

Predicted Th Motifs (B-cell epitopes) (SEQ ID NOS.: 131-133)

SSGGRITDDFIIRYLLVGI
 20 QGAAQKPINLSKXIEVVQGHDE
 SPGVFLEHLQEAYRIYTPFDLSA

Predicted HLA A2.1 Motifs (T-cell epitopes) (SEQ ID NOS.: 134-140)

YLLVGIQGA
 25 GAAQKPINL
 NLSKXIEVV
 EVVQGHDES
 HLQEAYRIY
 NIAFVAQAA
 30 FVAQAAPDS

WO 01/90152

PCT/US01/16776

102

EXAMPLE 5

IDENTIFICATION OF T-CELL EPITOPES OF B11Ag1

This Example illustrates the identification of B11Ag1 (also referred to as B305D) epitopes. Four peptides, referred to as B11-8, B11-1, B11-5 and B11-12 (SEQ ID NOs:309-312, respectively) were derived from the B11Ag1 gene.

Human CD8 T cells were primed *in vitro* to the peptide B11-8 using dendritic cells according to the protocol of Van Tsai et al. (*Critical Reviews in Immunology* 18:65-75, 1998). The resulting CD8 T cell cultures were tested for their ability to recognize the B11-8 peptide or a negative control peptide, presented by the B-LCL line, JY. Briefly, T cells were incubated with autologous monocytes in the presence of 10 ug/ml peptide, 10 ng/ml IL-7 and 10 ug/ml IL-2, and assayed for their ability to specifically lyse target cells in a standard 51-Cr release assay. As shown in Fig. 22, the bulk culture line demonstrated strong recognition of the B11-8 peptide with weaker recognition of the peptide B11-1.

A clone from this CTL line was isolated following rapid expansion using the monoclonal antibody OKT3 and human IL-2. As shown in Fig. 23, this clone (referred to as A1), in addition to being able to recognize specific peptide, recognized JY LCL transduced with the B11Ag1 gene. This data demonstrates that B11-8 is a naturally processed epitope of the B11Ag1 gene. In addition these T cells were further found to recognize and lyse, in an HLA-A2 restricted manner, an established tumor cell line naturally expressing B11Ag1 (Fig. 24). The T cells strongly recognize a lung adenocarcinoma (LT-140-22) naturally expressing B11Ag1 transduced with HLA-A2, as well as an A2+ breast carcinoma (CAMA-1) transduced with B11Ag1, but not untransduced lines or another negative tumor line (SW620).

These data clearly demonstrate that these human T cells recognize not only B11-specific peptides but also transduced cells, as well as naturally expressing tumor lines.

CTL lines raised against the antigens B11-5 and B11-12, using the procedures described above, were found to recognize corresponding peptide-coated targets.

WO 01/90152

PCT/US01/16776

103

EXAMPLE 6

CHARACTERIZATION OF BREAST TUMOR GENES DISCOVERED BY
DIFFERENTIAL DISPLAY PCR

5 The specificity and sensitivity of the breast tumor genes discovered by differential display PCR were determined using RT-PCR. This procedure enabled the rapid evaluation of breast tumor gene mRNA expression semiquantitatively without using large amounts of RNA. Using gene specific primers, mRNA expression levels in a variety of tissues were examined, including 8 breast tumors, 5 normal breasts, 2
10 prostate tumors, 2 colon tumors, 1 lung tumor, and 14 other normal adult human tissues, including normal prostate, colon, kidney, liver, lung, ovary, pancreas, skeletal muscle, skin, stomach and testes.

To ensure the semiquantitative nature of the RT-PCR, β -actin was used as internal control for each of the tissues examined. Serial dilutions of the first strand
15 cDNAs were prepared and RT-PCR assays performed using β -actin specific primers. A dilution was then selected that enabled the linear range amplification of β -actin template, and which was sensitive enough to reflect the difference in the initial copy number. Using this condition, the β -actin levels were determined for each reverse transcription reaction from each tissue. DNA contamination was minimized by DNase
20 treatment and by assuring a negative result when using first strand cDNA that was prepared without adding reverse transcriptase.

Using gene specific primers, the mRNA expression levels were determined in a variety of tissues. To date, 38 genes have been successfully examined by RT-PCR, five of which exhibit good specificity and sensitivity for breast tumors
25 (B15AG-1, B31GA1b, B38GA2a, B11A1a and B18AG1a). Figures 21A and 21B depict the results for three of these genes: B15AG-1 (SEQ ID NO:27), B31GA1b (SEQ ID NO:148) and B38GA2a (SEQ ID NO:157). Table I summarizes the expression level of all the genes tested in normal breast tissue and breast tumors, and also in other tissues.

WO 01/90152

PCT/US01/16776

104

TABLE I

Percentage of Breast Cancer Antigens that are Expressed in Various Tissues

5	Breast Tissues	Over-expressed in Breast Tumors	84%
		Equally Expressed in Normals and Tumor	16%
10	Other Tissues	Over-expressed in Breast Tumors but not in any Normal Tissues	9%
		Over-expressed in Breast Tumors but Expressed in Some Normal Tissues	30%
15		Over-expressed in Breast Tumors but Equally Expressed in All Other Tissues	61%

20

EXAMPLE 7

PREPARATION AND CHARACTERIZATION OF ANTIBODIES AGAINST BREAST TUMOR

POLYPEPTIDES

Polyclonal antibodies against the breast tumor antigen B305D were prepared as follows.

The breast tumor antigen expressed in an *E. coli* recombinant expression system was grown overnight in LB broth with the appropriate antibiotics at 37 °C in a shaking incubator. The next morning, 10 ml of the overnight culture was added to 500 ml to 2x YT plus appropriate antibiotics in a 2L-baffled Erlenmeyer flask. When the Optical Density (at 560 nm) of the culture reached 0.4-0.6, the cells were induced with IPTG (1 mM). Four hours after induction with IPTG, the cells were harvested by centrifugation. The cells were then washed with phosphate buffered saline and centrifuged again. The supernatant was discarded and the cells were either frozen for future use or immediately processed. Twenty ml of lysis buffer was added to the cell pellets and vortexed. To break open the *E. coli* cells, this mixture was then run through the French Press at a pressure of 16,000 psi. The cells were then centrifuged again and

WO 01/90152

PCT/US01/16776

105

the supernatant and pellet were checked by SDS-PAGE for the partitioning of the recombinant protein. For proteins that localized to the cell pellet, the pellet was resuspended in 10 mM Tris pH 8.0, 1% CHAPS and the inclusion body pellet was washed and centrifuged again. This procedure was repeated twice more. The washed inclusion body pellet was solubilized with either 8 M urea or 6 M guanidine HCl containing 10 mM Tris pH 8.0 plus 10 mM imidazole. The solubilized protein was added to 5 ml of nickel-chelate resin (Qiagen) and incubated for 45 min to 1 hour at room temperature with continuous agitation. After incubation, the resin and protein mixture were poured through a disposable column and the flow through was collected. The column was then washed with 10-20 column volumes of the solubilization buffer. The antigen was then eluted from the column using 8M urea, 10 mM Tris pH 8.0 and 300 mM imidazole and collected in 3 ml fractions. A SDS-PAGE gel was run to determine which fractions to pool for further purification.

As a final purification step, a strong anion exchange resin such as HiPrepQ (Biorad) was equilibrated with the appropriate buffer and the pooled fractions from above were loaded onto the column. Antigen was eluted off the column with a increasing salt gradient. Fractions were collected as the column was run and another SDS-PAGE gel was run to determine which fractions from the column to pool. The pooled fractions were dialyzed against 10 mM Tris pH 8.0. The protein was then vialled after filtration through a 0.22 micron filter and the antigens were frozen until needed for immunization.

Four hundred micrograms of B305D antigen was combined with 100 micrograms of muramyl dipeptide (MDP). Every four weeks rabbits were boosted with 100 micrograms mixed with an equal volume of Incomplete Freund's Adjuvant (IFA). Seven days following each boost, the animal was bled. Sera was generated by incubating the blood at 4 °C for 12-24 hours followed by centrifugation.

Ninety-six well plates were coated with B305D antigen by incubating with 50 microliters (typically 1 microgram) of recombinant protein at 4 °C for 20 hours. 250 microliters of BSA blocking buffer was added to the wells and incubated at room temperature for 2 hours. Plates were washed 6 times with PBS/0.01% Tween. Rabbit

WO 01/90152

PCT/US01/16776

106

sera was diluted in PBS. Fifty microliters of diluted sera was added to each well and incubated at room temperature for 30 min. Plates were washed as described above before 50 microliters of goat anti-rabbit horse radish peroxidase (HRP) at a 1:10000 dilution was added and incubated at room temperature for 30 min. Plates were again washed as described above and 100 microliters of TMB microwell peroxidase substrate was added to each well. Following a 15 min incubation in the dark at room temperature, the colorimetric reaction was stopped with 100 microliters of 1N H₂SO₄ and read immediately at 450 nm. The polyclonal antibodies showed immunoreactivity to B305D.

10 Immunohistochemical (IHC) analysis of B305D expression in breast cancer and normal breast specimens was performed as follows. Paraffin-embedded formal fixed tissue was sliced into 8 micron sections. Steam heat induced epitope retrieval (SILIER) in 0.1 M sodium citrate buffer (pH 6.0) was used for optimal staining conditions. Sections were incubated with 10% serum/PBS for 5 minutes. Primary
15 antibody was added to each section for 25 min at indicated concentrations followed by a 25 min incubation with either an anti-rabbit or anti-mouse biotinylated antibody. Endogenous peroxidase activity was blocked by three 1.5 min incubations with hydrogen peroxide. The avidin biotin complex/horseradish peroxidase (ABC/HRP) systems was used along with DAB chromagen to visualize antigen expression. Slides
20 were counterstained with hematoxylin. B305D expression was detected in both breast tumor and normal breast tissue. However, the intensity of staining was much less in normal samples than in tumor samples and surface expression of B305D was observed only in breast tumor tissues.

A summary of real-time PCR and immunohistochemical analysis of
25 B305D expression in an extensive panel of normal tissues is presented in Table II below. These results demonstrate minimal expression of B305D in testis, inconclusive results in gall bladder, and no detection in all other tissues tested.

WO 01/90152

PCT/US01/16776

107

TABLE II

mRNA	IHC staining	Tissue type	Summary
Moderately positive	Positive	Testis	Nuclear staining of small minority of spermatids; spermatozoa negative; seminoma negative
Negative	Negative	Thymus	No expression
N/A	Negative	Artery	No expression
Negative	Negative	Skeletal muscle	No expression
Negative	Positive (weak staining)	Small bowel	No expression
Negative	Positive (weak staining)	Ovary	No expression
Negative		Pituitary	No expression
Negative	Positive (weak staining)	Stomach	No expression
Negative	Negative	Spinal cord	No expression
Negative	Negative	Spleen	No expression
Negative	Negative	Ureter	No expression
N/A	Negative	Gall bladder	Inconclusive
N/A	Negative	Placenta	No expression
Negative	Negative	Thyroid	No expression
Negative	Negative	Heart	No expression
Negative	Negative	Kidney	No expression
Negative	Negative	Liver	No expression
Negative	Negative	Brain-cerebellum	No expression
Negative	Negative	Colon	No expression
Negative	Negative	Skin	No expression
Negative	Negative	Bone marrow	No expression
N/A	Negative	Parathyroid	No expression
Negative	Negative	Lung	No expression
Negative	Negative	Esophagus	No expression
Negative	Positive (weak staining)	Uterus	No expression
Negative	Negative	Adrenal	No expression
Negative	Negative	Pancreas	No expression
N/A	Negative	Lymph node	No expression
Negative	Negative	Brain-cortex	No expression
N/A	Negative	Fallopian tube	No expression
Negative	Positive (weak staining)	Bladder	No expression
Negative	N/A	Bone	No expression
Negative	N/A	Salivary gland	No expression

WO 01/90152

PCT/US01/16776

108

Negative	N/A	Activated PBMC	No expression
Negative	N/A	Resting PBMC	No expression
Negative	N/A	Trachea	No expression
Negative	N/A	Vena cava	No expression
Negative	N/A	Retina	No expression
Negative	N/A	Cartilage	No expression

EXAMPLE 8

PROTEIN EXPRESSION OF BREAST TUMOR ANTIGENS

5 This example describes the expression and purification of the breast tumor antigen B305D in *E. coli* and in mammalian cells.

Expression of B305D isoform C-15 (SEQ ID NO:301; translated to 384 amino acids) in *E. coli* was achieved by cloning the open reading frame of B305D isoform C-15 downstream of the first 30 amino acids of the *M. tuberculosis* antigen Ra12 (SEQ ID NO:318) in pET17b. First, the internal EcoRI site in the B305D ORF was mutated without changing the protein sequence so that the gene could be cloned at the EcoRI site with Ra12. The PCR primers used for site-directed mutagenesis are shown in SEQ ID NO:319 (referred to as AW012) and SEQ ID NO:320 (referred to as AW013). The ORF of EcoRI site-modified B305D was then amplified by PCR using 15 the primers AW014 (SEQ ID NO:321) and AW015 (SEQ ID NO:322). The PCR product was digested with EcoRI and ligated to the Ra12/pET17b vector at the EcoRI site. The sequence of the resulting fusion construct (referred to as Ra12mB11C) was confirmed by DNA sequencing. The determined cDNA sequence for the fusion construct is provided in SEQ ID NO:323, with the amino acid sequence being provided 20 in SEQ ID NO:324.

The fusion construct was transformed into BL21(DE3)CodonPlus-RIL *E. coli* (Stratagene) and grown overnight in LB broth with kanamycin. The resulting culture was induced with IPTG. Protein was transferred to PVDF membrane and blocked with 5% non-fat milk (in PBS-Tween buffer), washed three times and 25 incubated with mouse anti-His tag antibody (Clontech) for 1 hour. The membrane was washed 3 times and probed with HRP-Protein A (Zymed) for 30 min. Finally, the

WO 01/90152

PCT/US01/16776

109

membrane was washed 3 times and developed with ECL (Amersham). Expression was detected by Western blot.

For recombinant expression in mammalian cells, B305D isoform C-15 (SEQ ID NO:301; translated to 384 amino acids) was subcloned into the mammalian expression vectors pCEP4 and pcDNA3.1 (Invitrogen). These constructs were transfected into HEK293 cells (ATCC) using Fugene 6 reagent (Roche). Briefly, the HEK cells were plated at a density of 100,000 cells/ml in DMEM (Gibco) containing 10% FBS (Hyclone) and grown overnight. The following day, 2 ul of Fugene 6 was added to 100 ul of DMEM containing no FBS and incubated for 15 minutes at room temperature. The Fugene 6/DMEM mixture was added to 1 ug of B305D/pCEP4 or B305D/pcDNA plasmid DNA and incubated for 15 minutes at room temperature. The Fugene/DNA mix was then added to the HEK293 cells and incubated for 48-72 hours at 37 °C with 7% CO₂. Cells were rinsed with PBS, the collected and pelleted by centrifugation.

For Western blot analysis, whole cell lysates were generated by incubating the cells in Triton-X100 containing lysis buffer for 30 minutes on ice. Lysates were then cleared by centrifugation at 10,000 rpm for 5 minutes at 4 °C. Samples were diluted with SDS_PAGE loading buffer containing beta-mercaptoethanol, and boiled for 10 minutes prior to loading the SDS_PAGE gel. Proteins were transferred to nitrocellulose and probed using Protein A purified anti-B305D rabbit polyclonal sera (prepared as described above) at a concentration of 1 ug/ml. The blot was revealed with a goat anti-rabbit Ig coupled to HRP followed by incubation in ECL substrate. Expression of B305D was detected in the HEK293 lysates transfected with B305D, but not in control HEK293 cells transfected with vector alone.

For FACS analysis, cells were washed further with ice cold staining buffer and then incubated with a 1:100 dilution of a goat anti-rabbit Ig (H+L)-FITC reagent (Southern Biotechnology) for 30 minutes on ice. Following 3 washes, the cells were resuspended in staining buffer containing Propidium Iodide (PI), a vital stain that

WO 01/90152

PCT/US01/16776

110

allows for identification of permeable cells, and then analyzed by FACS. The FACS analysis showed surface expression of B305D protein.

From the foregoing it will be appreciated that, although specific
5 embodiments of the invention have been described herein for purposes of illustration, various modifications may be made without deviating from the spirit and scope of the invention. Accordingly, the invention is not limited except as by the appended claims.

WO 01/90152

PCT/US01/16776

111

CLAIMS

What is Claimed:

1. An isolated polynucleotide comprising a sequence selected from the group consisting of:

(a) sequences provided in SEQ ID NO: 1, 3-86, 142-298, 301-303, 307, 313, 314, 316, 317, 325 and 327-330;

(b) complements of the sequences provided in SEQ ID NO: 1, 3-86, 142-298, 301-303, 307, 313, 314, 316, 317, 325 and 327-330;

(c) sequences consisting of at least 20 contiguous residues of a sequence provided in SEQ ID NO: 1, 3-86, 142-298, 301-303, 307, 313, 314, 316, 317, 325 and 327-330;

(d) sequences that hybridize to a sequence provided in SEQ ID NO: 1, 3-86, 142-298, 301-303, 307, 313, 314, 316, 317, 325 and 327-330, under moderately stringent conditions;

(e) sequences having at least 75% identity to a sequence of SEQ ID NO: 1, 3-86, 142-298, 301-303, 307, 313, 314, 316, 317, 325 and 327-330;

(f) sequences having at least 90% identity to a sequence of SEQ ID NO: 1, 3-86, 142-298, 301-303, 307, 313, 314, 316, 317, 325 and 327-330; and

(g) degenerate variants of a sequence provided in SEQ ID NO: 1, 3-86, 142-298, 301-303, 307, 313, 314, 316, 317, 325 and 327-330.

2. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of:

(a) SEQ ID NO: 299, 300, 304-306, 308-312, 314, 326 and 331-334;

(b) sequences encoded by a polynucleotide of claim 1;

(c) sequences having at least 70% identity to a sequence encoded by a polynucleotide of claim 1; and

WO 01/90152

PCT/US01/16776

112

(d) sequences having at least 90% identity to a sequence encoded by a polynucleotide of claim 1.

3. An expression vector comprising a polynucleotide of claim 1 operably linked to an expression control sequence.

4. A host cell transformed or transfected with an expression vector according to claim 3.

5. An isolated antibody, or antigen-binding fragment thereof, that specifically binds to a polypeptide of claim 2.

6. A method for detecting the presence of a cancer in a patient, comprising the steps of:

- (a) obtaining a biological sample from the patient;
- (b) contacting the biological sample with a binding agent that binds to a polypeptide of claim 2;
- (c) detecting in the sample an amount of polypeptide that binds to the binding agent; and
- (d) comparing the amount of polypeptide to a predetermined cut-off value and therefrom determining the presence of a cancer in the patient.

7. A fusion protein comprising at least one polypeptide according to claim 2.

8. An oligonucleotide that hybridizes to a sequence recited in SEQ ID NO: 1, 3-86, 142-298, 301-303, 307, 313, 314, 316, 317, 325 and 327-330 under moderately stringent conditions.

WO 01/90152

PCT/US01/16776

113

9. A method for stimulating and/or expanding T cells specific for a tumor protein, comprising contacting T cells with at least one component selected from the group consisting of:

- (a) polypeptides according to claim 2;
- (b) polynucleotides according to claim 1; and
- (c) antigen-presenting cells that express a polypeptide according to claim 2,

under conditions and for a time sufficient to permit the stimulation and/or expansion of T cells.

10. An isolated T cell population, comprising T cells prepared according to the method of claim 9.

11. A composition comprising a first component selected from the group consisting of physiologically acceptable carriers and immunostimulants, and a second component selected from the group consisting of:

- (a) polypeptides according to claim 2;
- (b) polynucleotides according to claim 1;
- (c) antibodies according to claim 5;
- (d) fusion proteins according to claim 7;
- (e) T cell populations according to claim 10; and
- (f) antigen presenting cells that express a polypeptide according to claim 2.

12. A method for stimulating an immune response in a patient, comprising administering to the patient a composition of claim 11.

13. A method for the treatment of a cancer in a patient, comprising administering to the patient a composition of claim 11.

14. A method for determining the presence of a cancer in a patient, comprising the steps of:

- (a) obtaining a biological sample from the patient;
- (b) contacting the biological sample with an oligonucleotide according to claim 8;
- (c) detecting in the sample an amount of a polynucleotide that hybridizes to the oligonucleotide; and
- (d) compare the amount of polynucleotide that hybridizes to the oligonucleotide to a predetermined cut-off value, and therefrom determining the presence of the cancer in the patient.

15. A diagnostic kit comprising at least one oligonucleotide according to claim 8.

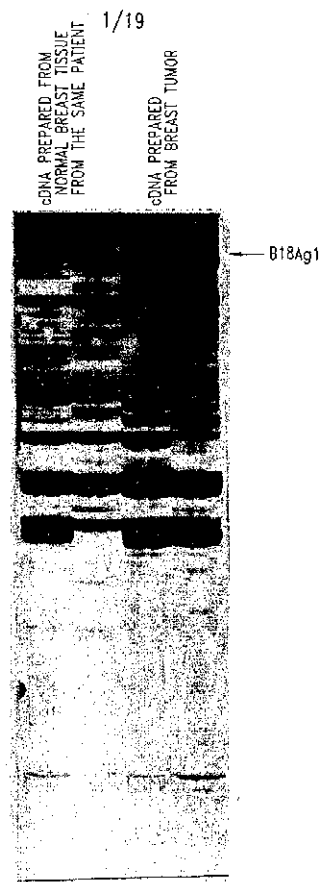
16. A diagnostic kit comprising at least one antibody according to claim 5 and a detection reagent, wherein the detection reagent comprises a reporter group.

17. A method for inhibiting the development of a cancer in a patient, comprising the steps of:

- (a) incubating CD4+ and/or CD8+ T cells isolated from a patient with at least one component selected from the group consisting of: (i) polypeptides according to claim 2; (ii) polynucleotides according to claim 1; and (iii) antigen presenting cells that express a polypeptide of claim 2, such that T cell proliferate;
 - (b) administering to the patient an effective amount of the proliferated T cells,
- and thereby inhibiting the development of a cancer in the patient.

WO 01/90152

PCT/US01/16776

*Fig. 1*

SUBSTITUTE SHEET (RULE 26)

WO 01/90152

PCT/US01/16776

2/19

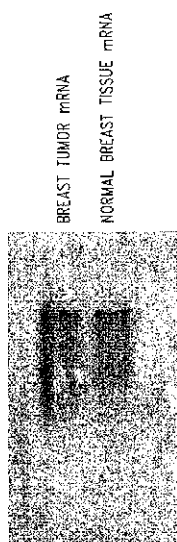
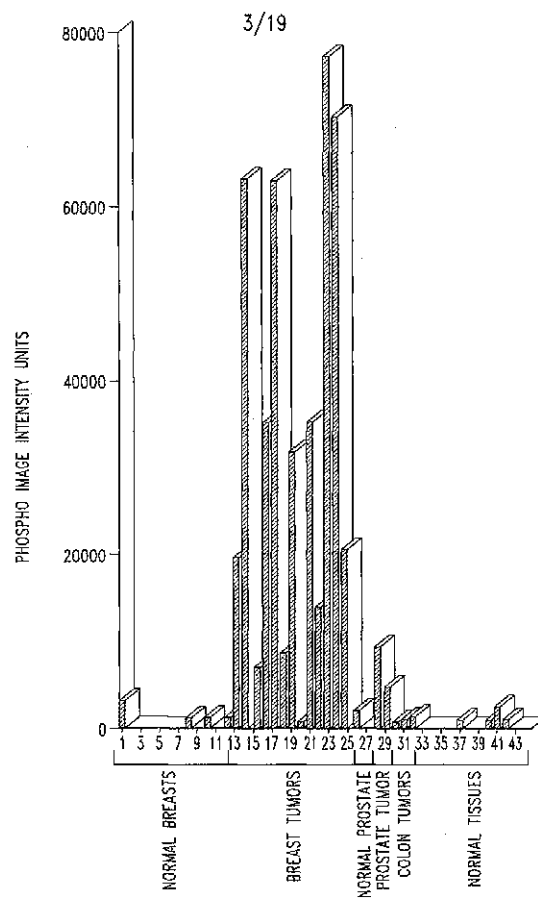


Fig. 2

WO 01/90152

PCT/US01/16776

*Fig. 3*

SUBSTITUTE SHEET (RULE 26)

WO 01/90152

PCT/US01/16776

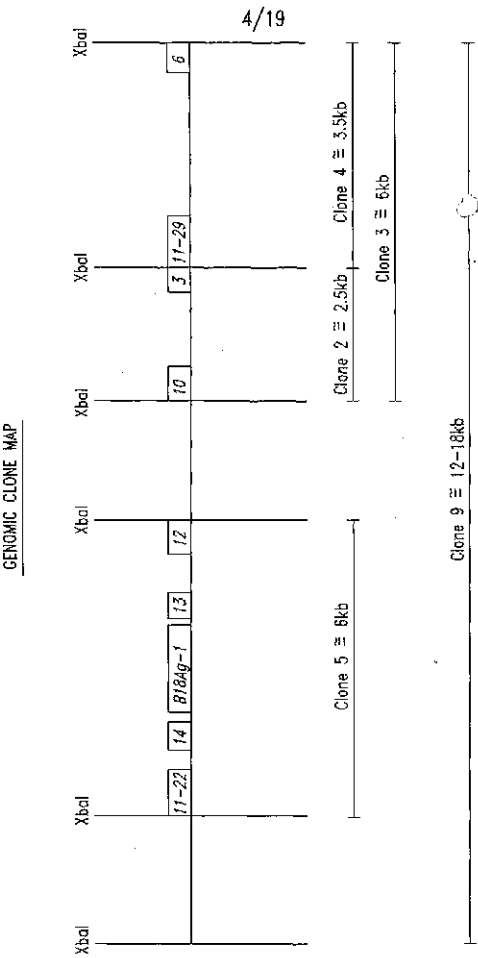


Fig. 4

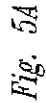


Fig. 5B

WO 01/90152

PCT/US01/16776

6/19

NUCLEOTIDE SEQUENCE OF THE REPRESENTATIVE
BREAST-TUMOR SPECIFIC cDNA B18Ag1

TTA GAG ACC CAA TTG CGA CCT AAT TGG GAC CCA AAT TTC TCA AGT GGA 48
Leu Glu Thr Gln Leu Gly Pro Asn Trp Asp Pro Asn Phe Ser Ser Gly
1 5 10 15

GGG AGA ACT TTT GAC GAT TTC CAC CGG TAT CTC CTC GTG GGT ATT CAG 96
Gly Arg Thr Phe Asp Asp Phe His Arg Tyr Leu Leu Val Gly Ile Gln
20 25 30

GGA GCT GCC CAG AAA CCT ATA AAC TTG TCT AAG GCG ATT GAA GTC GTC 144
Gly Ala Ala Gln Lys Pro Ile Asn Leu Ser Lys Ala Ile Glu Val Val
35 40 45

CAG GGG CAT GAT GAG TCA CCA GGA GTG TTT TTA GAG CAC CTC CAG GAG 192
Gln Gly His Asp Glu Ser Pro Gly Val Phe Leu Glu His Leu Gln Glu
50 55 60

GCT TAT CGG ATT TAC ACC CCT TTT GAC CTG GCA GCC CCC GAA AAT AGC 240
Ala Tyr Arg Ile Tyr Thr Pro Phe Asp Leu Ala Ala Pro Glu Asn Ser
65 70 75 80

CAT GCT CTT AAT TTG GCA TTT GTG GCT CAG GCA GCC CCA GAT AGT AAA 288
His Ala Leu Asn Leu Ala Phe Val Ala Gln Ala Ala Pro Asp Ser Lys
85 90 95

AGG AAA CTC CAA AAA CTA GAG GGA TTT TGC TGG AAT GAA TAC CAG TCA 336
Arg Lys Leu Gln Lys Leu Glu Gly Phe Cys Trp Asn Glu Tyr Gln Ser
100 105 110

GCT TTT AGA GAT AGC CTA AAA GGT TTT 363
Ala Phe Arg Asp Ser Leu Lys Gly Phe
115 120

Fig. 6

NUCLEOTIDE SEQUENCE OF THE REPRESENTATIVE
BREAST-TUMOR SPECIFIC cDNA B17Ag1

GC TGGGCACAGT GGCTCATACC TGTAATCCTG ACCGTTTCAG AGGCTCAGGT 60

CG CTGAGCCCA AGATTTCAG ACTAGTCTGG GTAACATAGT GAGACCTAT 120

AA AAATAAAAAA ATGAGCCTGG TGTAGTGGCA CACACCAGCT GAGGAGGGAG 180

CT AGGAGA 196

Fig. 7

WO 01/90152

PCT/US01/16776

7/19

NUCLEOTIDE SEQUENCE OF THE REPRESENTATIVE
BREAST-TUMOR SPECIFIC cDNA B17Ag2

```

GC TTGGGGGCTC TGACTAGAAA TTCAAGGAAC CTGGGATTCA AGTCCAAC TG   60
AC TTACACTGTG GNETCCAATA AACTGCTTCT TTCCTATTCC CTCTCTATTA   120
AA GGAAAACGAT GTCTGTGTAT AGCCAAGTCA GNTATECTAA AAGGAGATAC   180
AT TAAATATCAG AATGTAAAAA CTGGGAACCA GGTTCCTCAG CTGGGATTAA   240
CA AGAAGACTGA ACAGTACTAC TGTGAAAAGC CCGAAGNGGC AATATGTTCA   300
TT GAAGGATGGC TGGGAGAATG AATGCTCTGT CCECCAGTCC CAAGTCACT   360
CT CCTTTATAGC CTAGGAGA                                     388

```

Fig. 8

NUCLEOTIDE SEQUENCE OF THE REPRESENTATIVE
BREAST-TUMOR SPECIFIC cDNA B13Ag2a

```

GC CTATAATCAT GTTCTCAT TTTTCACAT TTTATTAACC AATTCTGT   60
AA AATATGAGGG AATATATGA AACAGGGAGG CAATGTTTCA ATAATTGATC   120
TG ATTTCTACAT CAGATGCTCT TTCCTTCTCT GTTTATTTCC TTTTATTTC   180
GG TCGAATGTAA TAGCTTTGTT TCAAGAGAGA GTTTTGGCAG TTTCTGTAGC   240
CT GCTCATGTCT CCAGGCATCT ATTTGCACTT TAGGAGGTGT CGTGGGAGAC   300
CT ATTTTITCCA TATTTGGGCA ACTACTA                               337

```

Fig. 9

SUBSTITUTE SHEET (RULE 26)

WO 01/90152

PCT/US01/16776

8/19

NUCLEOTIDE SEQUENCE OF THE REPRESENTATIVE
BREAST-TUMOR SPECIFIC cDNA B13Ag1b

```
GC EATACAGTC ETTTCATT ATTAAACCC CACCTGAAC GCATAAAGT   60
GC TGGTGTTTT TACTGTAAC AATAAGGAGA CTTGCTCTT CATTAAACC  120
AT TTCATATTTT ACGTCGAGG GTTTTACCG GTTCCTTTT ACACCTCTTA  180
TT TAAGTCGTTT GGAAACAAGT ATTTTTCCT TCCTGGCAGC TTTAAACATT  240
TT TGTGCTGGG GGAAGCTGG TCACGTTTC TCACAGTTGC AAATCAAGGC  300
CC AAGAAAAAAA AATTTTTTGG TTTTATTGA AACTGGACCG GATAAACGGT  360
CG GCTGCTGTAT ATAGTTTTAA ATGGTTTATT GCACCTCTT AAGTGCAC  420
GG GGGGNTTTG NATAGAAAGT NTTANTTAC ANAGTCACG GGAATTTTNT  480
NA CTGAGCTAAA AAGGGCTGNT TTTGGGTGG GGGCAGATGA AGGCTCACAG  540
TC TCTTAGAGGG GGAAGTCT A                               571
```

Fig. 10

WO 01/90152

PCT/US01/16776

9/19

NUCLEOTIDE SEQUENCE OF THE REPRESENTATIVE
BREAST-TUMOR SPECIFIC cDNA B13Ag1a

```
TA ATAACTTAAA TATATTTTGA TCACCCACTG GGGTGATAAG ACAATAGATA   60
TT TCCAAAAAGC ATAAACCAA AGTATCATA CAAACCAAT TCATACTGCT   120
CC GCCTGAAAC TTCACCTTCT AACTGCTAC CTAACCAAT TCTACCCTTC   180
GG TGGTGCTCA CTACTCTTTT TTTTCTTTT TTTNTTTGG AGATGGAGTC   240
CA GCCAGGGGT GGAGTACAAT GGCACAACCT CAGTCACTG NAACCTCCGC   300
TT EATGAGATT TCTGNTTCA GCTTCCAG TAGCTGGGAC TACAGTGTG   360
TG CCTGNTAAT CTTTTTNGT TTTNGGTAG AGATGGGGT TTTACATGTT   420
TG GTNTGAACT CCGACCTCA AGTGATCCAC CCACCTCAGG CTCCCAAAGT   480
TA CAGACATGAG CCACTGNGCC CAGNCCTGGT GCATGCTCAC TTCTCTAGGC   540
```

Fig. 11

SUBSTITUTE SHEET (RULE 26)

WO 01/90152

PCT/US01/16776

10/19

NUCLEOTIDE SEQUENCE OF THE REPRESENTATIVE
BREAST-TUMOR SPECIFIC cDNA B11Ag1

```
TG CACATGCAGA ATATTCTATC GGTACTTCAG CTATTACTCA TTTTGATGGC   60
AG CCTATECTCA AGATGAGTAT TTAGAAAGAA TTGATTAGC GATAGACCAA   120
GC ACTCTGACTA CACGAAATTG TTCAGATGTG ATGGATTAT GACAGTTGAT   180
GA GATTATTAAG TGATTATTTT AAAGGGAATC CATTAAATCC AGAATATCTT   240
TC AAGATGATAT AGAAATAGAA CAGAAAGAGA CTACAATGA AGATGTATCA   300
TA TTGAAGAGCC TATAGTAGAA AATGAATTAG CTGCATTTAT TAGCETTACA   360
TT TTCTGATGA ATCTTATATT CAGCCATCGA CATAGCATT CTTGATGGGC   420
GA ATAATAGAAA CTGGGTCCGG GGCTATTGAT GAATTCATCC NCAGTAAATT   480
AC AAAATATAAC TCGATTGCAT TTGGATGATG GAATACTAAA TCTGGCAAAA   540
GG AGCTACTAGT AACCTCTCTT TTTGAGATGC AAAATTTTCT TTTAGGGTTT   600
CT ACTTTACGGA TATTGGAGCA TAACGGGA                               638
```

Fig. 12

SUBSTITUTE SHEET (RULE 26)

WO 01/90152

PCT/US01/16776

11/19

NUCLEOTIDE SEQUENCE OF THE REPRESENTATIVE
BREAST-TUMOR SPECIFIC cDNA B3CA3c

```

ACTGATGGAT GTCGCCGAG GCGAGGGGCC TTATCTGATG CTCGGCTGCC TGTCGTGAT   60
GTGCGCGGCG ATTGGGCTGT TTATCTCAA CACCGCCACG GCGGTGCTGA TGGCGCCTAT  120
TGCCTTAGCG CCGGCGAAGT CAATGGGCGT CTCACCCAT CCTTTGCCA TGGTGGTGGC  180
GATGGCGGCT TCGGCGGCGT TTATGACCC GGTCTCCTCG CCGGTAACA CCCTGGTGGT  240
TGGCCCTGGC AAGTACTCAT TTAGCGATTT TGTCAAATA GCGGTG   286

```

Fig. 13

NUCLEOTIDE SEQUENCE OF THE REPRESENTATIVE
BREAST-TUMOR SPECIFIC cDNA B9CG1

```

AG CAGCCCTTC TTCTCAATT CATCTGTCAC TACCTGGTG TAGTATCTCA   60
CA TTTTATAGC CTCCTCCCTG GTCTGTCTTT TGATTTTCTT GCTGTAAATC  120
AC ATAAC TGCAA GTAACATTT CTAAAGTGTG GTTATGCTCA TGCACTCCT  180
AA ATAGTTTCCA TTACGTCCTT AATAAAATTC GGATTGTGTC TTTNCTATTN  240
CA CCTATGACCG AA   262

```

Fig. 14

WO 01/90152

PCT/US01/16776

12/19

NUCLEOTIDE SEQUENCE OF THE REPRESENTATIVE
BREAST-TUMOR SPECIFIC cDNA B9CG3

AG CAAAGCCAGT GGTTCGAGCT CTCTACTGTG TAAACTCCTA AACCAAGGCC 60
TA AATGGTGGCA GGATTTTAT TATAACATG TACCATGCA AATTTCCTAT 120
GA TATATTCTTC TACATTAAA CAATAAAAAT AATCTATTT TAAAAGCTA 180
AG TTAGGTAGA GTGTTAATG AGAGGTATA AGGTATAAT CACCAGTCAA 240
TG CCTATGACCG A 261

*Fig. 15*NUCLEOTIDE SEQUENCE OF THE REPRESENTATIVE
BREAST-TUMOR SPECIFIC cDNA B2CA2

CGACGTGGT AAAATGGAC ATGAAGCCAC CGCTGGTCTT TTCGTCCGAG CGATAGGCGC 60
CGGCCAGCCA GCGGAACGGT TGCCCGGATG GCGAAGCGAG CCGGAGTTCT TCGGACTGAG 120
TATGAATCTT GTTGTA AAA TACTGCGGC CTTCGTTGGA CGACGTCCG TCGAAATCTT 180
AATCATGGTT GAGCCGGATG CTGCCCCGA AGCCCT 276

Fig. 16

SUBSTITUTE SHEET (RULE 26)

WO 01/90152

PCT/US01/16776

13/19

NUCLEOTIDE SEQUENCE OF THE REPRESENTATIVE
BREAST-TUMOR SPECIFIC cDNA B3CA1

```

CCCAGGTCAA ECAGGCTGCA ACAGGAGGT CCTTGGATTG GGCACGAAGC AGCGCTTCGC   60
TGTTTTCAG GATTTTCAAC CAGTCGGCTT GGCCTTCTC ATGGAGCGAG AGCGCTTGC   120
CCAGCTCATT TTCAGGCGC TGTATTGCG TGGAAAAAG CACATCTCA CCGCAAGA   180
CATCTTTGA AATCGGCTGT TCCGCGATT CCAGATANTG CGAGGAGAGC TTGCTCGAAT   240
AGGTCATCT AACCTTCAA TGCACCAT GTGCGCAAT GAATATCTA ACAATTCAAC   300
TAGTTGGCAT AANAACGAA CGAAAATCC AATAGTCTGA AGAGCTCTT TG   352

```

*Fig. 17*NUCLEOTIDE SEQUENCE OF THE REPRESENTATIVE
BREAST-TUMOR SPECIFIC cDNA B3CA2

```

CTGATGTGC ACGGCTGGA TTTACGGGTG GTGCGGTTT ACCCTGGCA GCTGGCGGTC   60
TTCCCGACCA GCGCCAGCAG GATGTGTGG GCAAGGATAA CCGCGTGCBC ATCGCTCGA   120
CTATATGCC TACTGGCAAG GCGGAGCCCG TGGAGGCGG ATTCAGGTT ANCGGTCCT   180
GGAGCTTTTC CACCGCTCC ATGCATTGTG ACAGGCTGT TCTAGGCGGT CTGTTGCCA   240
AGCGTGATGG TACGTCTGCC CTGGAGCATG TGACTTCTG   280

```

Fig. 18

SUBSTITUTE SHEET (RULE 26)

WO 01/90152

PCT/US01/16776

14/19

NUCLEOTIDE SEQUENCE OF THE REPRESENTATIVE
BREAST-TUMOR SPECIFIC cDNA B3CA3

AG GGAGCAAGGA GAAGGCATGG AGAGGCTEAG GCTGGTCTTG GCCTACGACT 60
CT GTGCGCGGG ATGGTGGAGA ACTGAAGCGG GACCTCTCTG AGGTCTCTCG 120
TC NCCGTCCAGG AGGAGGGTCT TTECGTGGTE TNGGAGGAGC GGGGGGAGAA 180
TC ATGGTCACCA TCCC 204

*Fig. 19*NUCLEOTIDE SEQUENCE OF THE REPRESENTATIVE
BREAST-TUMOR SPECIFIC cDNA B4CA1

TC AGGAGCGGGT AGAGTGGCAC CATTGAGGGG ATATTCAAAA ATATTATTTT 60
TG ATAGTTCCTG AGTTTTCTT TGACCATGA GTTATATTGG AGTTTATTTT 120
CC AATCGCATGG ACATGTTAGA CTTATTTTCT GTTAATGATT NCTATTTTAA 180
GA TTTGAGAAAT TGGTNTTAT TATATCAATT TTTGGTATTT GTTGAGTTTG 240
GC TTAGTATGTG ACCA 264

Fig. 20

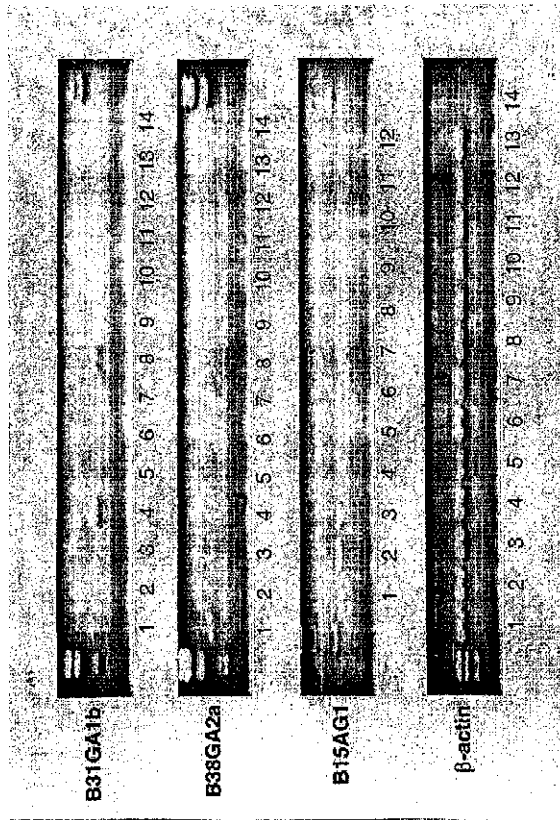


Fig. 21A

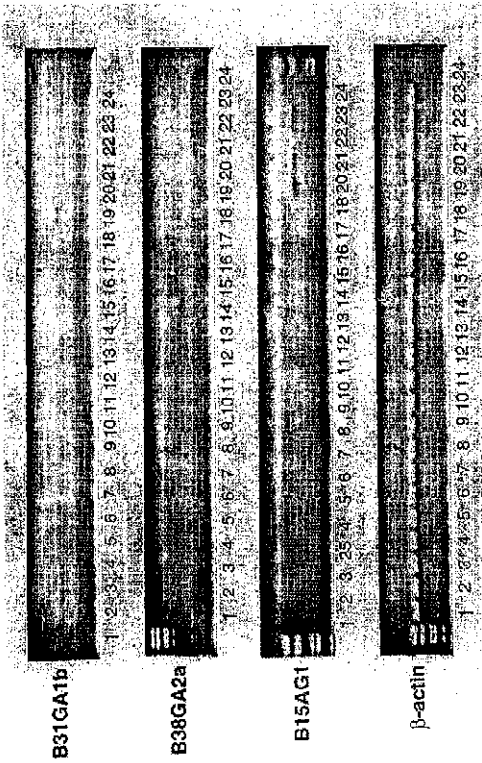
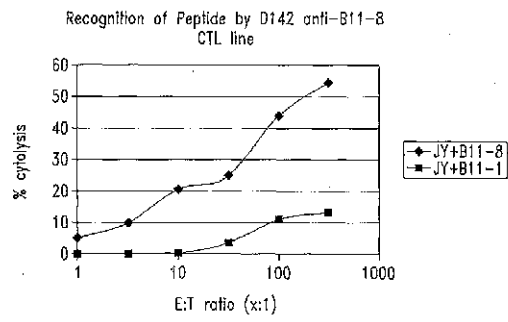


Fig. 21B

WO 01/90152

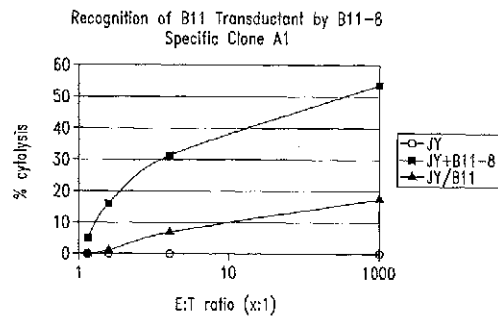
PCT/US01/16776

17/19

*Fig. 22*

SUBSTITUTE SHEET (RULE 26)

18/19

*Fig. 23*

WO 01/90152

PCT/US01/16776

19/19

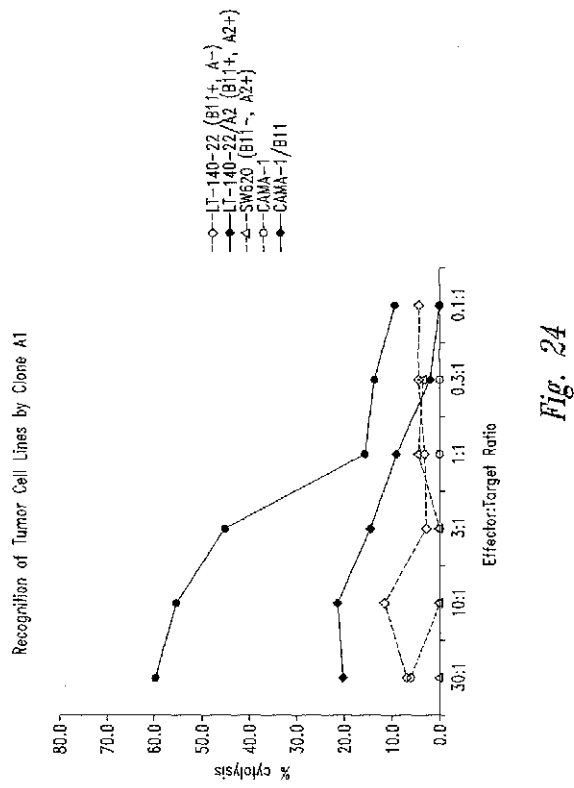


Fig. 24

SUBSTITUTE SHEET (RULE 26)

WO 01/90152

PCT/US01/16776

1

SEQUENCE LISTING

<110> Corixa Corporation
 Frudakis, Tony M.
 Reed, Steven G.
 Smith, John M.
 Misher, Linda E.
 Dillon, David C.
 Retter, Marc W.
 Wang, Aijun
 Skeiky, Yasir A.W.
 Harlocker, Susan L.
 Day, Craig H.

<120> COMPOSITIONS AND METHODS FOR THE
 THERAPY AND DIAGNOSIS OF BREAST CANCER

<130> 210121.41930FC

<140> PCT
 <141> 2001-05-22

<160> 334

<170> FastSEQ for Windows Version 3.0

<210> 1
 <211> 363
 <212> DNA
 <213> Homo sapien

<400> 1
 ttatggagcc aattgggacc taattgggac ccaaatctct chaetggag gagabctttt 60
 gacgatttcc aacggatctc cctcgtgggt attcaggag ctgccagaa acntataaac 120
 ttgtcttaag cgattgaagt cgtccgggg catgatgagt caccaggagt gtttttagag 180
 caccctcagg aggtctatcg gatttacac ccttttgacc tggcagcccc gaaaatagc 240
 catgctctta atttggcatt tctggctcag gcagccccag atagtaaaag gaaactccaa 300
 aaactagagg gattttgctg gaatgaatac cagtcagctt tttagagtag cctaaaaggt 360
 ttt 363

<210> 2
 <211> 121
 <212> PRT
 <213> Homo sapien

<400> 2
 Leu Glu Thr Gln Leu Gly Pro Asn Trp Asp Pro Asn Phe Ser Ser Gly
 1 5 10 15
 Gly Arg Thr Phe Asp Asp Phe His Arg Tyr Leu Leu Val Gly Ile Gln
 20 25 30
 Gly Ala Ala Gln Lys Pro Ile Asn Leu Ser Lys Ala Ile Glu Val Val
 35 40 45
 Gln Gly His Asp Glu Ser Pro Gly Val Phe Leu Glu His Leu Gln Glu
 50 55 60

WO 01/90152

PCT/US01/16776

2

Ala Tyr Arg Ile Tyr Thr Pro Phe Asp Leu Ala Ala Pro Glu Asn Ser
 65 70 75 80
 His Ala Leu Asn Leu Ala Phe Val Ala Gln Ala Ala Pro Asp Ser Lys
 85 90 95
 Arg Lys Leu Gln Lys Leu Glu Gly Phe Cys Trp Asn Gln Tyr Gln Ser
 100 105 110
 Ala Phe Arg Asp Ser Leu Lys Gly Phe
 115 120

<210> 3
 <211> 1080
 <212> DNA
 <213> Homo sapien
 <220>
 <221> misc_feature
 <222> (1)...(1080)
 <223> n = A,T,C or G

<400> 3
 tcttagaattc ttcctacccc gaactcttgg gaaaactttaa atcagtcacc tacagtctac 60
 caaccatttta ggaggagcaa agctaacctca gctcctccgg agcgttttta agatccccc 120
 tcttcaaaagc ctaacagatc aagcagctct cccgtgcacc aacctgcgcc aggtaaatgc 180
 caaaaagggt cctaaaccaa gcccaaggcc ccgtctccaa gaaactctac caggagaaaa 240
 gtgggaattt gactttacag agtaaaacc acacgggtt gggtacaaat acctctcagt 300
 actggttagac acctctctcg gatggactga agcatttgcct accaaaaacg aaactgtcaa 360
 tatggttagtt aagtttttac tcaatgaatt catccctcga cgtgggtgc ctgttgcatt 420
 aggtctctgat aatggaacgg ccttcgcatt gtctatagtt taateagtaa gtaagtgctt 480
 aaacattcaa tggaaagctc attgtgccta tcgacccaga gctctggcca agtagaacgc 540
 atgaactgca cctataaaaa acactcttac aaaaataalc ttaaaaaacc gtgttaattg 600
 tgttagtctc cttcccttag cctacttag agttaagggt caacctctac tgggctgggt 660
 tctttacatt ttgaattcat ntttnggaag gggctgccta tctttcttia actaaaaaan 720
 gccatttgg caaaacttcc ncaactaatt tntactncc tacytctccc caacaggta 780
 aaaaactctc tgccttttc aaggaaacct ccactcatt cctnaacaa aggcctgctc 840
 tcttccccc agttaactna tttttnttaa aattcccaaa aaangaaccc cctgctgaa 900
 aaaccccc ctcacaccc cggccnaaga ggaaggttcc cttagaacc ccacccctca 960
 aagcccccga accttaaan tagttccagg gggtnagccc taagaagccc atttgtaaa 1020
 cctanaaatt tttctctttn taasacacac nntttnttt tcttaaaaa aaacctntt 1080

<210> 4
 <211> 1087
 <212> DNA
 <213> Homo sapien
 <220>
 <221> misc_feature
 <222> (1)...(1087)
 <223> n = A,T,C or G

<400> 4
 tctgagactg cgcctggatc ccgcacagct gaggagacct gaagaccaga gaaaacacag 60
 caagttagcc ctttaaaacta ctacacctgt ttgtctctta atttattctg ttttattktg 120
 tttccatcat ttttaagggt taataaacta ttgttcagac ctacagcatat aaaaatgacc 180
 atctgttagac ctacagctcc aacactaccc caagagttgt ctggttttgt taaaattact 240
 gccagggttc agctgcagat atccctggaa ggaatattcc agattccctg agtagtttcc 300
 aggttaaaat cctataggct tctctgttt ttgaggaagag ttccctgtcag agaaaacat 360
 gattttggat ttttaacttt aatgcttgtt aaacgtata aaaaaaattt tctaccccta 420
 gctttaaagt actgttagtg agaaattaan attcctcag gaggattaaa ctgccatttc 480

WO 01/00152

PCT/US01/16776

3

```

agttacccta attccaaatg ttttgggggt tagaatotto tttatgttc ttgaagaaet 540
gttttatatt ttcccatona gataaattct ctccccccct onttttntnt ctnttttttt 560
aaacgggani cttgctcogt tgtccanget gggaaatttt ttttggccaa tctongatno 560
cttgcaanaa tctgontccc caaaallacc ncccttttcc cactccacc connggaatt 720
acctggaatt anagggcccc ncccccccc cggctaattt gtttttgttt ttagtaaaaa 780
acgggtttcc tgttttgttt aggatggccc amntctgacc cctnatact cccctctngc 840
ctctnaatnt tnggontang gottaccccc cccngnnngt tttctccal tnaattttto 900
tntggantot tgaatnccgg gttttccctt ttaaacnat ttttttttt nnnccccan 960
tttttccccc cccntrnta anggggggtt cccaaacgg gtccncccc angtcccaa 1020
tttttccccc cccctctctt tttcttttno cccaaaanto ctatcttttc cttnaaatat 1080
cnaantnt 1087

```

```

<210> 5
<211> 1010
<212> DNA
<213> Homo sapien:

```

```

<220>
<221> misc_feature
<222> (1) ... (1010)
<223> n = A,T,C or G

```

```

<400> 5
tctagaccua gaaatgggag gattttagag tgaatgatga tttctctatc atctgcagtt 60
agtaaacatt ctccacagtt tctgosaaaa gtacaaaac cactgcagat gtaaaacact 120
agglaacaaa ctaataatct cccaaatacc tcccccaag ctcaacaatt ttaactgttl 180
aggatcactg gctetaatca ccatgacatg aggtccacc caaaccatca agngctaaac 240
agacagaatg tttccactcc tgatccactg tgtgggaaga agacccgaac ttaccactg 300
ggggggcctg ntcanaanaa sagcccatgc cccgggtnt ncccttnaac cggacagaa 360
naaccaccca tccccacano tctctgttct ntggggccct catctgtgt cctctnttnc 420
tttgggggan acntggggaa ggtaccccat tctnttgacc cctcnanaaa acccngtgg 480
ccctttgccc tgattcnont gggccttttc tcttttccct tttgggttgt ttaatttccc 540
aatgtccccc gaacccctct ctnctgccc aaaaactacc taaatttctc nctangnntt 600
ttcttggtgt tcttttttca aggttaacott nccgttccan nccnaccnaa aatttnttcc 660
nactnctggn cccnbaaaaa mnaatonccc cnaattgccc gaattgtgtt ggttttctct 720
nctgggggaa accctttaa tttccccctt ggcgggacc ctttttttcc ccccttnga 780
agggcggggg ttcttccaga aactccaat ncaaaagccc tgcctctgtn tgaacccctt 840
ttcttaaat taaaaaatan ccggttnnga nnggcctctt tccctccng gngggngmg 900
aaantcttta cccnbaaaa ggttgcttag cccnngtcc ccaatccccc nggaadaata 960
aacctttttn aaaaaggaa tataattttn ccactccttn gttctcttcc 1010

```

```

<210> 6
<211> 950
<212> DNA
<213> Homo sapien:

```

```

<220>
<221> misc_feature
<222> (1) ... (950)
<223> n = A,T,C or G

```

```

<400> 6
tctagagcto ggggcggcga gctetaatc gactcaatat agggcgtaga ctgatctca 60
gctcaotgca atctctgccc cgggggtcat gcgtctctcc tgcctcaguc ttccaaqtat 120
ctgggattac agggctgcga caccacccc ggcataattt gttattitaa tagagatggg 180
gttttccctt gttggccann atggtctena accctgccc tanngtgat ccccncccn 240
ngantcnaa ctgctgggga tnnccgnnn nnnctccca nnnnnnnnn nnnctnccc 300
tntctcttnc tcnnnnnnn cctncttccc ncttctccc annntntnt cnnccnccc 360

```


WO 01/90152

PCT/US01/16776

4

```

cnnnccnnct nccnnccnnt tcnctnnnch tctcnnnnn nctnnnnnnn gnnnnnnnnn 420
cnnctacnct nccnnccnnt cctctctnn cctcnnnnn cctcnnnnn tntctctctn 480
ntnnnnnnct cccnnnnntet cctcnnnnn tctcnnnnn nccnnccncc nccnnccncc 540
ctnnnnnnn cccnnnnntet cctcnnnnn nctcnnnnn cctcnnnnn nctcnnnnn 600
cnnccnnct cctcnnnnn nctcnnnnn cctcnnnnn nctcnnnnn nctcnnnnn 660
tctcnnnnn cctcnnnnn cctcnnnnn cctcnnnnn nctcnnnnn nctcnnnnn 720
tctcnnnnn cctcnnnnn cctcnnnnn cctcnnnnn nctcnnnnn nctcnnnnn 780
tctcnnnnn cctcnnnnn cctcnnnnn cctcnnnnn nctcnnnnn nctcnnnnn 840
cctcnnnnn cctcnnnnn cctcnnnnn cctcnnnnn nctcnnnnn nctcnnnnn 900
nccnnccncc cctcnnnnn cctcnnnnn cctcnnnnn nctcnnnnn nctcnnnnn 960

```

```

<210> 7
<211> 1086
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(1086)
<223> n = A,T,C or G

```

```

<400> 7
tctagagctc gggcgccgga gtcctattaa cctcctactaa agggagtgga ctgcctcaga 60
ctgtctctgt gtctatgttg aaagaagtag nctaaagaga tctcattttg tctctactaa 120
agaaataatt cctcgccttg agatgctatt aatctgtaac cctagcccca acctgtctct 180
cctagagaca tctgtctgtg tgaatcagg ttaattgat ttagggctat gctttgktaa 240
aaaagtcttt gaagataata tctgtgttaa agtctctac cctctctaa tctcagtaa 300
cctagagaca aatcctctgc ggaagccgc aggtacctct gtctaggaat gccaggtatt 360
gtctcagatt tctcctctgt tgaatcagg agatctgccc tctaggggag ggtcagacct 420
gactgtctct cagcccgaca tctcctctgc cagctctccc cagcccgaca cctcagaaag 480
gtctgtctgt aggaagatta nctaaagagg aaggtctctt gctttgaggt aagaagaagg 540
ctctgtctct tctcctctgc tgggcaataa aatgtctctg tgttaaaccc gaatgtatgt 600
tctcctctct gagaatagga gaaacatcc ttaaggtctg aggtcagaca cctcggcgcc 660
atactctctt ttaatgcacg agatgtttgt nctaatgaaa cctcagggca nctcctctct 720
tctcctctct tctcagacaa aactttgttc nctttctctg cgaacctctc cctctcttan 780
cctatggccc tctcctctct cctcctctct ggtgaaaaa tctcctctaa tctcagggga 840
tctcctctct tctcctctct gctcctctct cctcctctct cctcctctct tctcctctct 900
aactgtctct ggtcctctct tctcctctct cctcctctct aactcctctc gctcagggga 960
gctcctctct aactcctct tctcctctct aactcctct aactcctct gctcagggga 1020
cgtcctctct tgaagtcctc cctcctctct aactcctct cgtcctctct tctcctctct 1086
nctcctct

```

```

<210> 8
<211> 1177
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(1177)
<223> n = A,T,C or G

```

```

<400> 8
nctcctctct tctcagacaa nctcctctct agctcagga gctcagacaa gccctcagta 60
aactcctctc gctcagacaa gctcctctct agctcagga gctcagacaa gccctcagta 120
atgtgtctta aactcagcta cactcctctt tgaatcagaa tctcctctct cgtcctctct 180
ctgtcaggaat ctgttggaat ctactcagaa tggcagacat gatctctcaa atgtcctctc 240
aggaaggtgt gatcctcag tctcagacaa cactcctctc tctcctctct aggtcagga 300

```

WO 01/90152

PCT/US01/16776

5

```

aggggcccgt gcanctgtta ccaaggagac tnatgtgtt egggctcagg ctttaaccanc 360
aaaacacctca ncnncnaagg ctgaactigat cgcctctact caggctctcg gatggggtaa 420
gggatatataa cgttaacact gacgcaggt acgctttgc tactgtgout gtaegtggag 480
ccatctacca ggaagctggg ctactactc ggcaagtgcc tgnatccac tgaagaagga 540
catcaaaagg aaaaamnggc tgttgcctgt ggttaaccna aactgatcn ncagctonaa 600
gatgctgtgt tgaactttac tcnctctct taaacttgot gccacacnc tctttccca 660
accagatctg ctgacaatc ccatactca aaaaaaaan aacttgcc cgaacccna 720
accaataaaa aoggggagg tngtnganc nnoctgacc aaaaatgat gatccccgg 780
gctgcaggaa tcaaatcan cttatcnat accccaacn ngggggggg ggcngtncc 840
cttnaccctc ncttctctc ttttncccc ccccgccnt ccttttttaa ctgtgaaa 900
ggaaaacctg ncttaacaan tctctcctg gacntccc ttoocggn gtttanaaa 960
aaaaagccno antccctcc naaatttga ongaagana aggaattiaa cttttattt 1020
ttntoctt antttgtann cccctttta ccagcgcaa cngccatent ttaanaaaa 1080
aaaagaang tttattttt ctngaacca tccaatana aancacccg nggggaagg 1140
ggngggaggc cnotcaccoc cttntgtng gngggnc 1177

```

<210> 9
 <211> 1146
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc feature
 <222> (1) .. (1146)
 <223> n = A,T,C or G

```

<400> 9
nccnntntt gatgttgtct ttttgcctc tctttggata ctttccctct cttcagagt 60
gaaaagggtc aaaaaggact gttgacagt atccaggty ggcacatgt tccagagta 120
agactccatc agtgaagta aagcctggg cttttcagag aaggaggat tatgggtttt 180
ccaattatcc agtgaagaa tagaagaag gacataaac caggagggg gtggagcact 240
catcccccag agggacttgt gccctctca gtgtagtag aggggctact tctccccc 300
aoggttgcaa caaggaggca atgggtgatg agcctacag ggacatanc gaggagaaat 360
gggatgacc taaggaggta ggcgtgtttt aaggcgttg gactgggtga gggaaactot 420
cctctctctc agagagaagc agtacaaggc gactgaacc ggtgaagyt gaggggaaa 480
aaaggtctgt gttcaggaag accttggaag taaaattatg atgtgtgat gaatggagcc 540
atggaagggtg tgcctctgac caaacacag catgtatca tgttaggaa acgatcag 600
gaagccggga atttcaita caacccccon cncagctta acaltgttg gttcagtac 660
ccttaagggt gccactccac tcaaatitt gccattctac ttgonaaat ttccaaact 720
tccttttita aggcgaatc cntatccct naaaacnaa aaaaatctg cncctattct 780
ggaaaaggcc cacccttacc caggtggaa gaatttttnc cttttttttt tttttgaag 840
cmtttntta attgaacctn attcncccc ccaaaaaaa aacccnccng gggggcggat 900
ttccaaaaac naattccctt accaaaaaa aaaaaacmc cttnttccc ttcnccctn 960
ttcttttaat tagggagaga tnaagocccc caatttcnng gnotngatnn gtttcccc 1020
cccccatlth ccaaaacttt ttccanona ggaancncc cttttttng gtcngatna 1080
naaaccllcc aaaccatttt tccnaaaaa ntittgtng ngggaaaaan aactntttt 1140
atagan

```

<210> 10
 <211> 545
 <212> DNA
 <213> Homo sapien

```

<400> 10
cttcattggg taaggggccc ctagaggtcg acggtatcga taagcttgat atogaattcc 60
tgcagccggg gggalccact agttctagag tcagggaaga ccaaccaact tcttgatttt 120
tattggctct gattctcgag gccagtttct ttottctgtt gagtatgag gatgtcaag 180
cagatctggc tgtggaaaag agactgtgg cagcaagttt agaggcgtga ctgaagta 240

```

WO 01/90152

PCT/US01/16776

6

```

cactgcactc tgcgctgcty aatcagcttt ctggtaacca cgggcaacgg cagtgttttc 300
cttttgatgt colltaacgt ggattacaag cactgctga gttgaatagc ccacgctcct 360
gtagatggc tccacgtaca tgcacagtac caaaggcgtc cctactgtca gtgttaacgt 420
taatatcctt acccactcgg agagcctgag tgaaggcgat caattcagcc cttttgtgct 480
gagggtgttg ctggttaagc cctgaaccca caacacatct gtctccatgg taacagctgc 540
aaccgg 545

<210> 11
<211> 196
<212> DNA
<213> Homo sapien

<400> 11'
tctctaggc tgggcaacgt ggctcatacc tctaactctg accgttccag aggcctcagg 60
ggggggatcg cttagacca agatttcaag actagtctgg gtaacatagt gagacctat 120
ctctacgaaa aaataaaaa atgagcctgg tgtagtggca caacacagct gaggagggag 180
aatcgagcct aggaga 196

<210> 12
<211> 388
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(388)
<223> n = A,T,C or G

<400> 12
tctctaggc ttgggggctc tgaactagaa ttcaaggaa cttgggaitca agtcccaactg 60
tgacaccaa ttacactgtg gntccaata aactgottct ttctattcc ctctctatta 120
aatasaataa ggaacaacgt gtctgtgtat agccaagtc gntatcttaa aaggagatac 180
taagtgcact taatatcag aatgtaaa cttgggaaca ggttccnagc ctgggattaa 240
actgacnca agaaagactga acgttaacta tctgaaaagc ccgaagmggc aatatgttca 300
ctctaccgtt gaaggatggc tggggaatg aatgctctgt cccccagtc caagctcaact 360
tactataact ccttatagc ctaggaga 388

<210> 13
<211> 337
<212> DNA
<213> Homo sapien

<400> 13
tagtagttgc ctataatcat gttttcaat atttccat tttattaacc aattctgtt 60
taccctgaaa aatatgaggg aaatatatga aacagggagg caatgttcc ataattgac 120
acagatagt attctacat cagatgctct tctcttctt gtttatttcc tttttatttc 180
ggttgagggt tgaatgtaa lagctttgtt tcaagagaga gttttggcag tttctgtac 240
ttctgacact gctcatgtct ccaagcatct atttgcactt taggaggtgt cgtgggagac 300
tgagggtct atttttcca tatttgggca actacta 337

<210> 14
<211> 571
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(571)

```

WO 01/90152

PCT/US01/16776

7

```

<223> n = A,T,C or G

<400> 14
tagtagttgc atacagtgcc atttccattt atttaacccc caccatgaag gcataaactg    60
agtggttcagc tgggtgttttt tactgtcaac aataaggaga ctttgccttt catttaaucc    120
aaaatcatat ttcatatttt acgctcgagg gttttttacc gttccttttt acactcotta    180
aaacagtttt taagtctgttt ggaacaagat attttttctt tccctggcagc ttttaacakt    240
atagcaaatc tctgtctggg ggaatgctgg tcaatgtttc tcaagttgc aataaaggc    300
atttgcaacc aagaaaaaaa aatttttttg tttttattga aactggacgc gataaacggt    360
gtttggagcg gctgctgtat atagttttta atggttttat gcacctcctt aagttgcact    420
tatgtggggg ggggnttttg natagaaagt ntttantcac anagtcacag ggaactttat    480
cttttgggna ctgagctaaa aagggcctgt tttcgggtgg gggcaqatga aggtcacag    540
gaggccttcc tcttagaggg ggggaactnct a                                571

<210> 15
<211> 548
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1) ... (548)
<223> n = A,T,C or G

<400> 15
tatatactta ataaccttaa tatcttttga tcaaccactg gggtgataag acaatagata    60
taaaagttat tccaaaaagc ataaaaacca agtatcatcc caaaocaaa tcaatctgct    120
tcccacccc gccctgaaac ttcaactttc aactgtctac cttaacaaat tctacacttc    180
aagtcctttg tgcctgctca ctactctttt tttttttttt ttttttttgg agatggagtc    240
tggctgtgca gcccaagggt ggaatacaat ggcacaacct cagctcaact naactccgcg    300
ctcccaggtc catgagatcc tctgnttca gccttccagc tagctgggac tcaagtgctg    360
catcacatgc cctggntaat ctttttttng tttngggtag agatgggggt tttaaatgtt    420
ggccaggntg gtttcgaact cctgacctca agtgatccc ccacctcagg ctcccaaggt    480
gctaggatta cagacatgag ccaactgngc cagncctggg gcctgctcac ttctctaggg    540
aactacta                                548

<210> 16
<211> 638
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1) ... (638)
<223> n = A,T,C or G

<400> 16
ttccgttatg cactgcaga atattcctac ggtacttcag ctattactca ttttgatggo    60
gcaatccagc cctatccctc agatgagtat ttgaaagaa ttgatttagc gatagaccaa    120
gctggttaag cctctgacta caggaattg ttcaagatg aggtatttat gacagttagt    180
ctttggaaga gtttttaag tgaattattt aaagggactc cattaatccc agaatatctt    240
ggttiagctc aagatgatat ggaatagaa caaagaaga ctacaantga agtgtatcaa    300
ccactgata tgaagagcc tatagtaga aatgaattag ctgcatttat tagactaac    360
catagcgttt tcaatgatga atcttatatt cagcactaga caatgaatta cctgatgggc    420
aaccttaaga ataatageaa ctgggtgagg ggcatttgat gaattcatcc ncagtaaat    480
tggatatnac aaatataac tccgttgcat ttggatgatg gaattactaa tctggcaaa    540
gttaactttg agctactagt aacctctctt ttggagatgc aaatatttct ttttagggtt    600
cttattctct actttacgga tattgagca taacggga                                638

```

WO 01/90152

PCT/US01/16776

8

```

<210> 17
<211> 286
<212> DNA
<213> Homo sapien

<400> 17
actgatggat gtgcgcggag gggagggggc ttatctgatg ctgcgcctgc tggctgtgat    60
gtgcgcggcg attgggctgt ttatctcaaa cgcgcgccag ggggtgctga tggcgccctat    120
tgcccttagc gggcggaagt caatgggcgt ctccacctat ccttttgcca tgggtgtggc    180
gatggcggtt tggcggcgtt ttatgacccc ggtctctctg ccggttaaca cactgtgtgt    240
tggccctggc aagtaactat ttatggattt tgcataata ggcgtg
286

<210> 19
<211> 262
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(262)
<223> n = A,T,C or G

<400> 18
tgggtctatg cagccncttc ttctcaattt caictgtcac taccctggty tagtatctna    60
tagccttaca tttttatagc ctctccctg gtctgtcttt tgaatttctt gcctgtatc    120
catactacac ataacctgca gttaacattt cttaagtgty gttatgctca tgtcactcct    180
gtgnaagaa atagtttcca ttaccgtctt atataaatc ggaattgttc ttlnctattn    240
taactcttca cctatgacgg aa
262

<210> 19
<211> 261
<212> DNA
<213> Homo sapien

<400> 19
tgggtctatg caagccaggt ggtttgagct ctctactgtg taaactctta aaccasggcc    60
attttatgta atgtgtggca ggaattttat tataaacatg taccactgca aatttctctat    120
aatcttgaga tttatctctt tacattttaa caataaaaat aatctatttt taaaagccta    180
atttgccttg ttaggtaaga gtgtttaaag agaggggtata aggtatataat caccagtcac    240
cgtttctctg cctatgaacc a
261

<210> 20
<211> 294
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(294)
<223> n = A,T,C or G

<400> 20
tacaacgagg cgacgtcggg aaaaatggao atgaagccac cgtctgtctt ttcgtccgag    60
cgatagcgcc cggccagcca ggggaacggg tgcccgatg ggaagcgag ccggagttct    120
tcggactgag tatgaatctt gttgtgaaa tactcgccgc ctccgttcca cgaactcgag    180
tcgaatctt cgaactctct acgatcgaag tcttctgtgg cgaagatcgc ggtcagttcc    240
gcccaaccga aatcatgggt gaggcggatg ctgncccga agnctcgtt tgn
294

```

WO 01/90152

PCT/US01/16776

9

```
<210> 21
<211> 208
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1) ... (208)
<223> n = A,T,C or G

<400> 21
ttggttaagg gcatggagcg agacgcctga cgtttgctg aaaattcttc attgattcgt 60
atcaatgaat aggaatttc ccaagagggg aatgtcctgt tgcctgcnag tttttn'gtt 120
gttctcatgg anaaggcaan gagctcttca gaattattgn attntcgttc ggtctctctg 180
caactagtcg ncttgcnaag atcttcat 208

<210> 22
<211> 287
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1) ... (287)
<223> n = A,T,C or G

<400> 22
nccnttgagc tgaagtgttg agatnlytaa tggttgtaag ggtgattcag ggggattcgg 60
gtggcgggtc acccggcagt gggctctccc acagccacgc aggtatttgg gcaggtacgg 120
ngtcgcgatk gctcgactat atgctatggc aggcagcccg tggaaaggng atcaggtcac 180
ggcgctggag ctttccacgg tccatgnatt gngatggctg ttctaggcgg ctgltgcca 240
gcytgatggt acgctggctg gagcattgat ttctgggtgc aaggtgg 287

<210> 23
<211> 204
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1) ... (204)
<223> n = A,T,C or G

<400> 23
ttgggtaaag ggagcaagga gaaggcatgg agaggutcan gctggtcctg gactaagaat 60
gggcgaagct gtggccgggg atgggtgaga actgaagcgg gaactcctcg aggtcctcgg 120
ngttaaactc nccgtccagg aggaaggtct ttccgtggtc tgggaggagc ggggggagaa 180
gctctccttc atggttcaaa taac 204

<210> 24
<211> 264
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1) ... (264)
```

WO 01/90152

PCT/US01/16776

10

```

<223> n = A,T,C or G

<400> 24
tggatttggtc aggagcgggt agagtggcac cattgagggg atattcaaaa atattatttt 60
gtctaaatg atagttgctg agtttttctt tgacctatga gttatatlgy agttatatt 120
tlaactttcc aatcgcatgg acatgtttaga cttattttct gttaatgatt notattttta 180
ttanattgga tttagaast tggttnttat tatatcaatt ttgggtattt gttgagtttg 240
acattatagc ttagtatqgy aaca 264

<210> 25
<211> 376
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(376)
<223> n = A,T,C or G

<400> 25
ttacaacgag gggaaactcc gtctctacaa aaattaaaa attagccagg tgtgtgtgtg 60
tgcacccgca atcccageta cttygggggt tgagacacaa gantcaccta natgtgggag 120
gtcaaggttg catggttcct gattgtgcca ctgcactcca gcttggttga cagaacgaga 180
cctgctccca aagaaanaag aataggaagt tcaagaatca tggntgtggn gaccagcaat 240
ctgcactcat acaacccctg caggcaange tgatgcagcc tangttcaag agctgtgttt 300
tctggaggca gcagttnggy ctccactcca gtatcaaggg cacactcgca onagccatct 360
gtctctccgtg tgtaac 376

<210> 26
<211> 372
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(372)
<223> n = A,T,C or G

<400> 26
ttacaacgag gggaaactcc gtctctacaa aaattaaaa attagccagg tgtgtgtgtg 60
tgcacctgta atcccageta cttygggggt tgagacacaa gantcaccta natgtgggag 120
gtcaaggttg catggttcct gattgtgcca ctgcactcca gcttggttga cagaacgaga 180
accctgcttc aaaaagaaaa gaataggaag ttcaagaacc ctggtgtgtg ngcccagcaa 240
tctgcattta aacatccctt gcaggcaatg ctgatgaagg cttaagttaa gagctgtgtt 300
tctggaggca gnatgaagg ctccactcca gtatcaaggg caacactgca aagacccctg 360
tctctgttgg ta 372

<210> 27
<211> 477
<212> DNA
<213> Homo sapien

<400> 27
ttctgtccac atctacaagt ttattttatt ttgtgggttt tcagggtgac taagtttttc 60
notacattga aagagaagt tgcataaagg tgcacaggaa atcatttttt taagtgaata 120
tgataatag ggtccgtgct taatacaact gagaatatt tgtttctctg ttttttagag 180
tcacctctta agtccatc coacatggt gaaaaaaaa tagaagtat ttgttctacc 240
tttaaggaga ctgcagggat tctctctgaa aacggagtat ggaatonate ttaataaant 300

```

WO 01/90152

PCT/US01/16776

11

```

atgaattgg ttggtttttt gggataagaa attcccaact cagtgtgtgt aaattcaact 360
gaattttttt gggcaaaaat agtcaaaaat gtcaatttgg tcaataaaat acatgttact 420
attaaaagat attaaagac aaattttttt agagctctaa gatgtgtgtg geosgae 477

<210> 28
<211> 438
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(438)
<223> n = A,T,C or G

<400> 28
tctncaact cttgattgtc aaaaaccttn taggctatct ctaaaagctg actggtatto 60
attccagcaa aatccctcta gtttttggag ttccctttta ctatctgggg ctgcttgagc 120
cacaaatgcc aaattaagag catggctatt ttgggggggt gacaggtcaa aagggtgtga 180
aatccgatac gctctctgga ggtgctctaa aaacactoot ggtgactcat catgcccttg 240
gaagacttca atcgacttag acaagtttat aggtttcttg gcagctccct gaataacac 300
gaggagatac cgttggaatc cgtcaaaagt tctccctcca ctgagaaat ttgggttaca 360
attaggtcac aattgggtct ctaactacta tctctctaga ttctctctcc ggnctatttg 420
ttgatgtggt gttgaaga 438

<210> 29
<211> 620
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(620)
<223> n = A,T,C or G

<400> 29
aagaggttac cagcccaagc cttgacaac ttcaataggg tgtcaagcct gtgggtgcaa 60
agaagtcaaa aattgagttt tgggactcto agctatgatt tcaagagata taaagaaaca 120
cctaaccact agatattcag acaaaagttt actacgggga tgaagcttct acggaaaccc 180
tctactagga aagtacagaa gagaatgttg ggtttggagg ccccaaacag aatccctct 240
agaacactgc ctatgaacac tgtgagaaga tggccactgt catccagaca ccagaatgat 300
agancacaca aaaaattatg caatattgou tataaacctt acagacacto aatgcaagcc 360
caatgaaaaa aaaaatgaga agaagactgt nccctaaact gccacgggag cagaactgcc 420
ccaggccatg gaagccacagc tottatatca atgtgaacct gatgttgaga catggaatcc 480
naggaaatcn tttaaact tccacggttn aatgactgcc ctattanatt cngaacttan 540
atcnggcoct gtgactcttt tgcattggcc attcccoctt ttggaatgg ctnttttttt 600
ccatgctctg tncctcttta 620

<210> 30
<211> 100
<212> DNA
<213> Homo sapien

<400> 30
ttacacagag ggggtcaatg tcaataatgt cacaataaaa caatctcttc tttttttttt 60
tttttttttt tttttttttt tttttttttt tttttttttt 100

<210> 31
<211> 762

```


WO 01/90152

PCT/US01/16776

12

```

<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1) ... (762)
<223> n = A,T,C or G

<400> 31
tagctctatgc gccggacaga gcagaattaa attggaagtt gccctoaagg ottttctaoce 60
acctctctcc tgaagagaga aagaagaggg gcaggaaaga ggttaggatt tcaattttcaa 120
gaatnagota attaggagag cagagtttag acagcagtag ggcctccatg atacaaucca 180
tggacaaagt ccttggttag taactggcag acctgatct gctaaagttt tgaattctct 240
ctgcccataa aagatggaga gcaggagtgc atccscatc aacacgtgtc caagaagaag 300
tctcagggag acaggggtat caaaaaacaa gattcttaatt gggaaggaa tcaaaccaaa 360
aaattagatt ttctctatcc tatatataat atacagatat ttaacacatt attccagagg 420
tggctccagt ccttggggct tgagagatgg tgaaaacttt tgttccacat taettcttgc 480
tctcaaatcc tgaagtatat cagaatggga caggcaatgt ttgtctccac actggggcac 540
agacccaaat gtttctgtag ccgaagaaga gaagccgaaa agacatgaag gatgcttaag 600
gggggttggg aagocaaat tggtaatac ttttccctct gctgtgttc cngaagtctc 660
cnetgaagga abcttaaaa cctttgtga gaaatgccc ccttaccatg acaantgttc 720
ccattgcttt tagggngatg gaaacacca gggttttgat cc 762

<310> 32
<211> 276
<212> DNA
<213> Homo sapien

<400> 32
tagctctatgc gtttattaac ctccctccc taagtaacaa ccaagaagga aggagctgtt 60
altaccacac cccttttaca gatgcataa taatgaaga gaagtgaagt gaatttgcca 120
cacacacagt aaattggcag agtcagattt gaatccatgg agtctggtct gcactttcaa 180
tcaccgaata ccctttctaa gaaagctgtg ctgaatgagt gaatggataa atcagtgctc 240
actcaacatc ttgtctctaga tatcccgcat agaat 276

<210> 33
<211> 477
<212> DNA
<213> Homo sapien

<400> 33
tagtagttgc caastatttg aaattttacc cagaagtgat tgaaaacttt tlggaacaaa 60
aaacaaataa agccaaaagg taastaataa atatctttgc actctcgttc ttacctatcc 120
ataacttttt caacgttaagc tctctgctgt gtagtgtag tctggttata ttgaactttt 180
tagttatatt tttttattca cttttccact agaaagtcac tattgattta gcaacacatg 240
tgatctcatt tcaatttttc tttttatagg caaaatttga tgcctatgaa caaaaatact 300
caagcccaatt atcttttttc ccccgaatc ctgaaaattg cagggggcag aggyaagtta 360
tcccattaaa aaattgtaaa tatgttcagt ttatgtttac aatgcacaa aacataagaa 420
aatgtgtatt acttgagctg ctgattgtaa gcagttttat ctccagggga actacta 477

<210> 34
<211> 631
<212> DNA
<213> Homo sapien

<400> 34
tagtagttgc caattcagat gatcagaaat gctgctttcc taagcattgt ctgttatccc 60
cgcatgcatc ttggaacttt ggcagtgaag agccaaaagg aagaggtgaa tgacatatat 120

```

WO 01/90152

PCT/US01/16776

13

```

atatatatat attcaatgaa agtaaatgt atagctaat atactttota gttatocaea 180
tgagttatango tttatgocat tgggcgctg catattttaa tcaagaagata aaggaasatc 240
tgggcalttt tagaagtga tacatgtttt tttaaaactg ttaaatatta ttccgataat 300
tgtutaagaa cgggaatgtt cttaaaaatt actaaaacag tattgtttga ggaagagaaa 360
actgtactgt ttgccattat tacagtctga caagtgcacg tcaagtcacc caactctcca 420
ggcatcagta tccacctcat agctttacac attttgaagg ggaatattgc agcatctcca 480
ggcctgacat ctgggaagg ctcagatcca cctactgcgc ctgtctgctt gatttgtttt 540
aaatatttgt gcttggtgtc actttaagc cacagccctg cctaaaagcc agcagagaa 600
agaaaccgca ccattctata gcaactact a 631

```

<210> 35

<211> 578

<212> DNA

<213> Homo sapien

<400> 35

```

tagtagttgc catcccatat tacagaagcc tctgtatata tgactttatt ggaagtgtac 60
tgttttctct ccaaacccat ttatcgtlaa ttcaaccagc ttggatcaat ctgtgtttcc 120
actgatacca tgaacctac ttggagcaga cattgcacag ttttctgtgg taataactaa 180
aggttttatt gctaaagctgt catcttatgc ttagtatttt tttttacag tggggaattg 240
ctgagattac attttgttat tcatagata ctttgggata actgacact gtctcttttt 300
ttcgcctttt aattgctatc atcatgcttt tgaacacaga scacattagt cctcaagtal 360
tacataagct tgcctgtcac gccctgtggt ttaagagct atctttgggc taaggtcac 420
aagaatgggg aaagtgttcc ctatgtttct gtaagttcca ataaagatt gccaggggcc 480
gggtactgtg gctcgcactg taatccacag actttgggaa gctgaggtg gcgcatcatg 540
ttagggcagg tgttcgaac cagctgggc aactacta 578

```

<210> 36

<211> 583

<212> DNA

<213> Homo sapien

<400> 36

```

tagtagttgc ctgtatccc agcaactcag gaggctgggg caggagatc agttgaacct 60
gggagggcaga agttgtactt agcaagatc gccacattgc acttcagcct gggcaacaa 120
gtgagatctt cttctcaaaa acaaaaaaaa gaanaagaaa agnaaaggaa aaacgtata 180
aaacagcca aaacaaatg atcattcttt taatacgaaa gactaattta atgtgtttat 240
taatacaaa cagttgaatc ttctgagtta ttggtgaaaa taacctgtga gttaattag 300
ggtcttactt ggggtgaacg ttigtgtttt acaggttata aaatggttaa caaggaat 360
ggtcctaaa gaactttata aactactaaa aataaataaa atataaatgg ataggtgcta 420
tggatgagat ttttgttata tttaaatctt tgaagtcatt ttggatgctc atttgtgtc 480
tggtaatttc cattaggaaa aggttatgat atggggaac tgtttctgaa aattgcggaa 540
tgtttctaat ctgtaaatg ctgtatctc agggcaacta cta 583

```

<210> 37

<211> 716

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> [1]...[716]

<223> n = A,T,C or G

<400> 37

```

gatctactag tcatntggat totatccatg gcagctaago ctttctgaat ggattctact 60
gctttcttgt tctttaatcc agacatttat atatgtttat gttcacagga agggcaatgt 120
ttagtgaasa caattctaaa ttttttattt tgcattttcc tgctaatttc cgtcacactc 180

```

WO 01/90152

PCT/US01/16776

14

```

cagcaggett cctggggagaa taaggagaaa tacagctaaa gaaattgtcc cigtctactt 240
acagcctaact ggtatgcata accacttcaa taaggtanaca ggaanagtaa taaccaggta 300
gaatggacca aaactgatac agaaaaatca gaggaagaga ggaacaaata ttactcgagt 360
cctagaatgt acaaggcttt ttaattacat attttacgta aggcctgcata aaaaacagggt 420
agtaactcaac atttgtccca ttttacctat aaggaacatg aagcttaaat tgaataattt 480
antgcataga ttttatagtt agcccatgtt caggtcccta tgttatactt actagctgta 540
tgaatatgag aaaaataatt ttttatitlc ttggcatcag tattttcata tgcnaaataa 600
agctaaagtt atttagcaaa cagtcagcat agtgcctgat acatagtagg tgcocaaac 660
atgattacac tantattngg tattcnanaa atccaatata ggcntggata aaaccg 716

<210> 38
<211> 688
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(688)
<223> n = A,T,C or G

<400> 38
ttctgtccac atatcatccc actttaattg ttaatcagca aaactttcaa tgaanaaica 60
tcaattttaa ccaggatcac accaggaaac tgaagtggtt ttttttttta ctttaaaaaa 120
aaaaaaanaa accaaacaaa ccaaaaacga ttaacagcaa aagattttta aaaaattaca 180
tttctcttac aactgtcaat cagagaaaca tagttcttaa gtctgttaa tcttgcaatt 240
aacagagaaa cttgatgaan agttgtactt ggaattttgt ggaatttttt ttttgtctaa 300
tctccacta tttttttgca aacagttaatt taagtttgtg tgaacatcc ccgtagttag 360
agtgtaaaca atgtatagga aggaactatg gataaagatg tgcatacat atgcattaca 420
tgtagggaac ttcacaaactt catgcactca gaaaanaatg ttgaagagga ggaagggaag 480
gcccagggtc accatccagg tgccttgagg acagagaagt cagaagtggc actgttgaaa 540
tttagaagac catgtgtgaa tggtttcagg cctgggatgt ttgccacaa gaagtgcctc 600
cagaaaattt ctttccattt tgaataacag ggtggttga tgggtacggt ggttgaccua 660
acgaagaaaa tgaattcttg ccccttcc

<210> 39
<211> 585
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(585)
<223> n = A,T,C or G

<400> 39
tagtagttgc cgcnaacctt aaanttggaa agcatgatgt ctaggaaaca tantaaata 60
gggtalgcct atgtgtctaa gagagatgtt agaatitaaa gtccatantt ttatgtattt 120
tgaacatgca attatccctt aaactccaaa actgattccg aagctattac aattaaaaag 180
tttgcacagg cgtgylgggc ggtgggtgac gctgtaatc ccagcacttt gggaggtcga 240
ggcaccggga tccagaggtc ggyagttcaa gacctctctg gctaacacgg tgaagtcca 300
tctctactaa aaatacgaaa aaattccccc ggcgtggtgg cgggagccta taatccacag 360
tactccggag gctgaggcag gagaatggcg tgaacccagg acacggaggt tgcagtgtgc 420
caacatcccg tcaatggccc ccagcctggg ggaacaggaac agantcccg tctcnaaaa 480
agaaaaatcc taactnatant ttcnacttta ttttaantta cacagaactn cctcttggtt 540
cccccttacc attcatctca cccacctcct atagggaacn nctaa 585

<210> 40
<211> 475

```

WO 01/00152

PCT/US01/16776

15

```

<212> DNA
<213> Homo sapien

<400> 40
tcgtccaca ccaatcttag aagctotgaa aagaatttgt ctttaaatat cttttaatag 60
taacatgtat ttatgggacc aatttgacat ttctgactgt tttttccaaa aagtcaggt 120
gattttcagc acactgagtt ggggaatttt tatcccaaaa gaccaccaa ttcatatatt 180
atttaagatt gattccatcc tccgttttca agggaaatcc ctgcagctcc ctttaaggta 240
gaacaaatcc ttcttatttt tttttccaaa ttgtgggatt ggcattttaa aggtgactct 300
aaaaaaacag aaaaacaata tgtctnagtt gtttaagca cgaacccata ttatcatatt 360
cacttaaaaa atgatttccc tgtgaaactt ttggcaactt cctttttcaa tgtaggyaaa 420
aacttagtca aactgaaaac caacaaaata aataaaactt gtatgtgtgg acaga 475

<210> 41
<211> 423
<212> DNA
<213> Homo sapien

<400> 41
taagagggtg catcgggtga aacgttaggc acatctagag cttagagaag tctggggtag 60
gaaaaaaato taagtattta taagggtata ggttaacattt aaaagtaggg ctagctgaca 120
ttattllaga ggaacacata cggggagata agggcaagg actaagacca gaggaacact 180
aatatttagt gctccctccc atctctggta aaatagtaa cttttaagll agcttcaagg 240
aagatttttg gcatgattta gtgtgaaaa gtaggtctcc ttgggtttat attactaatt 300
ttgttttaag atccttgtaa gtgctttaat aagtcattgt tatatcaaac gctctanaac 360
attgtagcat gttaaatgta acatataact taacctttgt tgtatatgga tgtacccctc 420
cta 423

<210> 42
<211> 527
<212> DNA
<213> Homo sapien

<220>
<221> misc feature
<222> (1) ... (527)
<223> n = A, T, C or G

<400> 42
tctctaggc taatgtgtgt gttctgttaa aagttaaaa ttaaaatttt taaaaataga 60
aaaaagotta tagantaaga atatatagaa agaaatatt ttgtacatt tgcacatga 120
gtttatgttt taagctaagt gttattacaa aagagccaaa aaggttttaa aaattaaaac 180
gtttgtaaag ttacagtacc attatgttaa ttatctcttg aagaaagaaa aacttttttt 240
tataaatgta gtgtagacta agcatcagtt atttataaag tctggcagtg ttcaataant 300
tcttaggact tccattccac tccatgactc accagagaca acttccagtc ctgtaagctc 360
cattcgttgg aagtgcccta tacaggtgca caatttatt tacagtatt ttactgtacc 420
ttctctatgt ttccatatgt ttcatatata aatatccact ggttactatn ggcacacagg 480
taattccagt aacacggcct gtatcgtctt ggtanoccta gngaaga 527

<210> 43
<211> 331
<212> DNA
<213> Homo sapien

<400> 43
tcttcaacct cgtaggacaa ctctaatatg cctgggcaact atttttaggt tactaccttg 60
gctgcctctc tttaagaaaa aaaaaagaag aaaaaaagac ttttccaaa gttctctctc 120
ctctagttgg aaatttagag aaatcaagtt tttaatttgg tgttatttca gatcanaaat 180

```

WO 01/90152

PCT/US01/16776

16

```

tcaaacacatt gtaaacatta agcttttgtt caatcccttg ggaagaggat tcattctgat 240
atttacgggtt caaaagaagt tgtaatttgg tgccttggaac acagagaacc agttattaac 300
ttcctactac tattatataa taaataataa c 331

```

```

<210> 44
<211> 592
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(592)
<223> n = A,T,C or G

```

```

<400> 44
ggcttagtag ttgccaggca aaatarcgtt gattctctc aggaagcaac cccaacaccc 60
cigtgttgctt ctgaccttat acctagacta aagtcocago agacccctag aggtgagggtt 120
cagagtgacc cttgaggaga tgtgtctaac tagaasaagaa ctgcttgagt ttctctaattt 180
atataagcag aactctggag agaggtcata ggaatggata ttaagggtgt gagataatgg 240
cggaaaggaat atagagttgg atcaggctgg acttattgat ttgaaccac taagttagaga 300
ttctgctttt gatgttgcag ctocaggagt taasaaaggt tttaattggtt ctacttggtt 360
ttttgcttgg ttagctgaaa tatggataaa agatggccca ctgtgagcaa gctgggaatg 420
cctgactctct ctccgittaa ttttagaggaa ggaatccaaa agtttaggga gacttggatg 480
ctggraktgg attggtcact ttggaacta cccwccccag ctgggaagggt ccagaagata 540
ccccccttgc ccccgctttg cgaatggat ttgtgatggc ggcactact aa 592

```

```

<210> 45
<211> 567
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(567)
<223> n = A,T,C or G

```

```

<400> 45
ggcttagtag ttgccattgc gactgcttgc tcaacgagcg ttgaacatgg cggatttgtt 60
agattcaacg gatttgagtt ttaccagcaa agcgaaccaa ggcgggcccc gagaattatg 120
ggtttggttg ctttgaanaa atggaatacc tgtaggonta gtcagaasaq ccttcttgca 180
gaaoagttgg ttctggggcg aacgtctatc agatgcccc ttggaaaggg tagcgttgtat 240
ttggagagcg ctgatagcgt gctctctgat gctgttctg ctggacaggt gacaaagat 300
atgcaaaagca agtcgaact agacgtcaag ctctgtgagc aaattatgt agactctac 360
ttatacttgt aggaatgata gccaaaggtg gggactttaa gactaagggt gtttgtactt 420
ggcgcagatg tcccaagcag aaagancga tgcctagttt tatacgggca actactaagc 480
cgaatkoawg caaacctggc gacttaata attggatccg anctoggtac cagcttgatg 540
catascitga gttwtctata ntgtcno 567

```

```

<210> 46
<211> 908
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(908)
<223> n = A,T,C or G

```

WO 01/90152

PCT/US01/16776

17

```

<400> 46
gagcgaaaga cggaggcgag ngntatngag cghngaagcg gagaggcgca aaaaagcuacc 60
gctttccccc ggggggtccc atttattaaag gcagggtggag gacagggttc cggatggaag 120
ggcgaggggg cggaaagcaat taatgtgaat aggcatttca ttacgacccc ggcttaacat 180
ttaagcttgc ggttggatag tgggtgggaat tggtagcgga taacaatttc acacaggaaa 240
cagctatgac catgattaac ccaggtatatt taggtgacat tatagaataa ctcaagttat 300
gcataaagct tggtaaccgag ttccggtacca ctagttaacg ccgccaatgt gtggaattcg 360
gcttagtagt tgcgacccat ggaagtgtac ctaggctaga atacctgagy tctcccttag 420
ctccactcac attaaattgt atcttttcta cattagatgt cctcagcgcc ttatttctgc 480
tggacwatag ataaattaat cctgatagga tgatagcagc agattaatta ctgagagtat 540
gttaattgtt cttccctcct atataacgta ttgcaatttt aatggggcas ttctggagat 600
aatccctgaa ggcgaaggaw tgaatcttga ggttgagaaa ggcgaatca gtctccagct 660
gcagtttttg gagaaggtga tttatgttat gtctcagaa tgacaccata tggcacaata 720
ctaagcccgta attccagcac actggcgggc gttactaatg gatccgagct cgttaccag 780
cttgatgcat agcttgagta tctataggtt cactaaatag cctggcggtta tcatggctat 840
agctgtttcc tgttggaant tgttatccgc tcccaattcc ccccaacata cgaagcggaa 900
czttaagat 960

```

<210> 47

<211> 480

<212> DNA

<213> Homo sapien

<220>

<221> misc feature

<222> (1)..(480)

<223> n = A,T,C or G

```

<400> 47
tgccacaag gaaagtitta aatttccctc tgaggattct tggtagcat caaattcagt 60
ggttttttaag gttgtttttot gtouaataac tctaacttta agccaacacg tatatggaag 120
cacagataxa atattacaca gataacagag gaggtagatc aaagtaraga tagttggggg 180
atttaatttc tggaaactag gtctcccaat cttctctgt gtgaggaac ttcttggaag 240
gggggattot aaagtctctt ggaagacagt ttgaaacca ccatgttgtt ctcaatcct 300
ttatttttaa aaagttagtg aacattttga gagagaaaag ggcttgggtt agatgaagtc 360
ccccccccc ctttttttt tttagctga aatagatacc ctatgttnaa xgaagagatt 420
attatttacc atgcctatx scacatgctc tttagtgggc nctccctac cttccttaag 480

```

<210> 48

<211> 591

<212> DNA

<213> Homo sapien

```

<400> 48
aaggaggtag cgaagtgaat ttccgcttca ctagtctggt gtggctagtc ggtttctgg 60
tggccaacat taccgaattc caactcaacc gttcttggag gttcaagcgg gaggtaaccg 120
gaggatggtg gctgaatttc tggcctttct ttgctgtggg atcggtagcc gccataatcg 180
gtatgtttat caagatcttc tttaactaac cgaactctcc gatitaaact cccagagcgt 240
ggttkaagga ggggaggggg attcagtcac gggagtactg gtcccagatc ttccgcatcg 300
tcgtgcaaat gctatcaaac ttctgtctca ataagtttgc gaccttccga accgtgaagc 360
actccgaane gttccgggtg ctgctgtgct gtagctccca aaatcttgat aacacraagg 420
taaccgaatc gogtcaagga accccggcat ctccgttaat ctgcattatgc gtacccotta 480
agccgaattc cagacacatg gggccgttca ctaattggat ccgaactccg taaccnagcc 540
tgatgcgtaa cttgatttat tctataggtt ccttaaata acotggcggtt n 591

```

<210> 49

<211> 454

<212> DNA

WO 01/90152

PCT/US01/16776

18

```

<213> Homo sapien

<400> 49
aagagggtac ctgccttga attaaatgt otaaggaaar tggagatga ttaagattg 60
gigigycyta gtccaccaa aatgtattta ttacatcctg ctcttttcta gttgaacaga 120
aagaaagctg ctgtgggaa aggagggata aatactgaag ggatttacta aacaaatgto 180
catcacagag ttttctcttt ttttttttg agaaagagto ttgctotgto aaccaggctg 240
gaatgaagwg gtatgaldto agttgaatgc aacctctacc tctaggttcc aagcgattct 300
catgctcag cctctgagc agctgggact ataggcgact gctaccatgc caggctcaatt 360
tttatattt tattagagac ggggtgttgc catgtggcc aggcaggtct cgaactcctg 420
ggctctongat gatctgccc accgtaccct ctta 454

<210> 50
<211> 463
<212> DNA
<213> Homo sapien

<400> 50
aagagggtac caaaaaaag aaaaaggaaa aaaaagaaaa caacttgtat aaggctttct 60
gctgcataac gctttttttt ttaaaataaa tggcgcaaac aaatgttttt gcattcacac 120
caatigctgg ttctgaatcc gtactcttca aaggtatttg tgcagatcaa tccaatagtg 180
atgcccctga ggttttctgg actgcccag ttgttaact ttctatgag gagccattga 240
gagactgttt ggaacatgct gtgttcaatg agccgtgatg tccgggggac gtglacatca 300
tgtttacgtg ggggtgggtc tgccttggct gctgggcata tggctgggtg cccatcatgc 360
ccatctgcat ctgcataagg tattggggcg ttgatccat atagccatga ttgctgtgtt 420
agccactgtt catcattggc tgggacatgc tttacccttc tta 463

<210> 51
<211> 399
<212> DNA
<213> Homo sapien

<400> 51
cttcaacctc ccaaaagtgt gggattacag gactgagcca ccaagctcag cctaagcttc 60
ttttcacta cctctaaagc gatctaccac agtgaagagg ggtcaagag cagtgcattt 120
tgattcaatt aatggaactt agatttatta altacaactt ttctcttagc atgttggthc 180
ctaatatttt aagagatagg acttacttag aatgaagctt tcaattttaag aatttctct 240
ttacccttct ctattagctt gacagatag acaatatacg tatltttatt aactaacota 300
ccttgagata ttaacttita aaggcotata taactgaatg tgtatttgta actgttaagc 360
cccaacagtat ttaattatat catgaatgct ttgaggttg 399

<210> 52
<211> 392
<212> DNA
<213> Homo sapien

<400> 52
cttcaacctc aatcaacctt ggtaattgat aaatcatca cttaactttc tgatataatg 60
gcaataatta tctgagaaaa aaaaagtgtg aagatttaaa cttgcatttc tctcagaato 120
tgtgaagata ttggaataat tcaaaagcgg aatcagtagt atcagccgaa gaaactcaat 180
tagctagaa cgttgaccca tggatctaaq tccctgcctt tccactaacc agctgatttg 240
ttttgttaa aactcctaca cgtctgggct tggctgccta atttgcaaa gtaaaagctg 300
aaataggaa ataatgaac gtgtctttt ggtctctttt ccatcoatta ctctgatttt 360
acaaagaggc ctgtattccc ctggtgaggt tg 392

<210> 53
<211> 179
<212> DNA

```

WO 01/90152

PCT/US01/16776

19

```

<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(179)
<223> n = A,T,C or G

<400> 53
ttcgggtgat gccctctcag gctacagtga agactggatt acagaaaggc gccagcgaga 60
tttcagattc ctgttaacct cttaagaaaa ggagtcgcgc ctcaactga' gtaqaaatga 120
ctagttcagg atcagagaa acmtctgact ccgattctag aggaactgagt gaactgcan 179

<210> 54
<211> 112
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(112)
<223> n = A,T,C or G

<400> 54
ttcgggtgat gccctctcag gctacatcat natagaagca aagtagaana atcnngtttg 60
tgcattttcc cacaacaaa attcaaatga ntggaaagaa ttgganagt at 112

<210> 55
<211> 225
<212> DNA
<213> Homo sapien

<400> 55
tgagcttcag cttctgacaa ctcaatagat aatcaaaagga caactttaac agggattcac 60
aaaggagtat atccaaatgc caataaacct aiaaaaagga attcagcttc atcatcatca 120
gaagwatgca aattaaaaac ataattgaga accactatgt cccactagaa tagataaant 180
cttaaaagac tggtaaaacc aagtgttggc saggaagag gagca 225

<210> 56
<211> 175
<212> DNA
<213> Homo sapien

<400> 56
gctctctttg ccttaacaa acattctcaa aaacctgcta gactcctaag cattctcttg 60
ttagttatgg gattttaccc ctgtctctata aagatgttat gtacaaaaa tgaagtggag 120
ggcctatccc tgaaggaggg gagggatctc tagtgttgtc agaagcggaa gctca 175

<210> 57
<211> 223
<212> DNA
<213> Homo sapien

<400> 57
agccatttac caccatgga tgaatggatt ttgtaattct agctgttgta ttttgtgat 60
ttgttaattt tgttgttttt ctgtgaaca catcattgg atatggaggg taaaggagtg 120
tcccagttgc tccctggtca tcnctttata gccattactg tttgtttct tgttaactag 180
gttaggtttt ggtctctctt gctccactgc aaaaaaaaa aaa 223

```


WO 01/90152

PCT/US01/16776

20

```

<210> 58
<211> 211
<212> DNA
<213> Homo sapien

<400> 58
gttcgaagg taaactgttag gtacgggac tcacaactgg ggaactgtca aagacgaatt 60
aaactgacttg gatcaactcaa atgtgactga ggaacacact gaaggtgaag aacatcatcc 120
agtggcagac actgaaacta aggaagatga agttgaagag gtaaaagagg aggttccaaa 180
agagatgact ttggatgggt gyttaaatggc t 211

<210> 59
<211> 208
<212> DNA
<213> Homo sapien

<400> 59
gctatctcttg cattaccacac ttgcaacca tcataacca tgtggccagg ttgcagccc 60
aggotgaaca taaggggact gcttgcaat acttcatgct gttgtgctg actgatggtg 120
ctgtgacgga tggggaagcc aacgtgaag ctgtgtgctg tgcctcgaac ctgcccatgt 180
cagtgatcat tatgggtggt aaatggct 208

<210> 60
<211> 171
<212> DNA
<213> Homo sapien

<400> 60
agccatttc caccatctact aaattctagt tcacaactca acttcttcca taaaacatct 60
aacctctgac aacagttagc atagcttct tccttcttta acctcttaga gtatttatgg 120
tcactgccc acattctgc aactgaataa agttggttaag gcaagaggag c 171

<210> 61
<211> 134
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(134)
<223> n = A,T,C or G

<400> 61
cgggtgatgc ctctcaggc ttgtgtgtgt caactonact cactggcctc ttctccagca 60
actggtgaan atgtctctca gaaanctcc acacgcatgt cagggtgggg tgggaancat 120
cenaactctc nggc 134

<210> 62
<211> 145
<212> DNA
<213> Homo sapien

<400> 62
agagggtaaa tatgaacag tatataaagg aagaagtga ctgagaggaa ctctctcaag 60
gccatttaac caataagtga tagagtcaag gtcaccca gggtgacgg attocaggtc 120
ccaagctcat tactggtacc ctctt 145

<210> 63

```

WO 01/90152

PCT/US01/16776

21

```

<211> 297
<212> DNA
<213> Homo sapien

<400> 63
tgcaatgaga ggaattcaga gggtttatgc caaagaacaa acaagtccctc tgcagcctaa 60
ctaatltgtt ttgtggctgc gaagccatgt agagggcgat caggaagtag atgtccctc 120
ccacagtcag cgcacatggtg gtccggtaaa gcatttggtc aggcaggcct cgtttcaggt 180
agacgggac acatcagctt tntggaaaa ctlttgtgac tctggagctt tgttttccc 240
agcataatca tcaactgtty aatggagggt cagtttagtt ggttaaggaa gaggaga 297

<210> 64
<211> 300
<212> DNA
<213> Homo sapien

<400> 64
gcactgagag gaaattccaa tactatgttg aataggagtg gtgagagagg gcaccccttg 60
cttgtggagg ttltcaaggg gaatgcttcc agcttttggc cattcagtat aatattaaag 120
aatgttttac cattttctgt ctgacgtgtt ttctgtgttt ttgtgtgttc tcttcattct 180
ccatttttag gcttttacct gttaggaaata tatctctttt aatgatactt cacttttggt 240
ctctttttgt agactctact catagtgtga taagcactgg ttgtgtgagg caagaggagc 300

<210> 65
<211> 203
<212> DNA
<213> Homo sapien

<400> 65
gctctctctg ccttaccacac tcaccacagta tgtcagcaat ttatctcgtt ttacctacga 60
aacagccctgt atccaaacac ttaacacact caactgaaaa gtccaggcaa caatgccttt 120
ctcatgggtc tctctgctcc agttctgac ctltctcttt tcttagaaca tgcatttarg 180
togatagagg tctctctcag tgc 203

<210> 66
<211> 344
<212> DNA
<213> Homo sapien

<400> 66
taagggggac catgcatlga gaaagcgaga ctcaacttga agctgaaatg ctgttgccct 60
tgcagtgctg gtagcaggag ttctgtgctt tctgggctaa ggtctctgga tgaccctcga 120
catggagaaag gcagagttgt gtgccccttc tcatggcctc gtcaaggaat catggactgc 180
cacacacaaa atgcctgttt tattaacgac atgaattga aggaagagac acaattcact 240
gatgtggctc gtaaccatgg atatggtcac atacagaggt gtgattatgt aaaggttant 300
tcccccacc tcatgtggaa actagcctca atgcagggtt caca 344

<210> 67
<211> 157
<212> DNA
<213> Homo sapien

<400> 67
gcactgagag gaacttcgta gggaggttga aotggctgct gaggaggggg aacacaggg 60
taaccagact gatagcattt ggttggataa tatgtgtgtt gaggagggac actacttata 120
gcagaggggt gtgtatagcc tggggaggga tcaacccg 157

<210> 68

```

WO 01/90152

PCT/US01/16776

22

```

<211> 137
<212> DNA
<213> Homo sapien

<400> 68
gcaactgagag gaactttctag aaagtgaag ttatagacata aataaata aaactttaa 60
actaaggaga gacagccag cagggtggt cagcctgta atccagaaac ttggggagcc 120
tgaggaygca tcacccg 137

<210> 69
<211> 137
<212> DNA
<213> Homo sapien

<400> 69
cgggtgatgc ctctcagga tttattttga agactataga ctggacttct tatcaactga 60
agntccgtt aaaaatacca gtgtattat ttctacctgt caaatccat ttcaaatgtt 120
gaagtctctc tcaatgc 137

<210> 70
<211> 220
<212> DNA
<213> Homo sapien

<220>
<221> misc feature
<222> (1)...(220)
<223> n = A, T, C or G

<400> 70
agcatgttga gccagacac gcaatctgaa tgagtgtgaa cctaaagtaa atgtctacac 60
gctgctggt ctgacatggc acacatcnc gtggaggga casctctgct cngcctacaa 120
cgagggcact ctcatagaca ggttcaccc accaaactgc aagaggtcga nnaagtaatr 180
ccagggtaaa aggaacagg tgggalyoa ycaacatct 220

<210> 71
<211> 353
<212> DNA
<213> Homo sapien

<220>
<221> misc feature
<222> (1)...(353)
<223> n = A, T, C or G

<400> 71
cgttaggttc tctatcaact gtaaacct acacttgggt aaacaggga cattaacat 60
tccacaaas atatgcaag tgacttcaa tgtttatct aaagtgtcc aaacgaaac 120
tgattttctc cctaaacct gtgatgtgg galgaitaan cctgagtggt ctacagcaag 180
ttaagtgcac ggtgcctaat gaagtgaac tgaatacag catctacaag gaagtacctc 240
tcaacaaag gaaactttgc ttctcnaag gcatttagca gtgtctgaag taatttctgt 300
attacaactc acggggcggg ggggtgaat ctantggana gnagacccta acg 353

<210> 72
<211> 343
<212> DNA
<213> Homo sapien

```

WO 01/90152

PCT/US01/16776

23

```

<400> 72
gcactgagag gaacttcaa tacyatkac agagtgaaca tgcacccyac aqaacaggag 60
aaaatgttgg caatctctcc atctgacaaa agcctaatat ccagawtota awaggacott 120
aaacaaattt atgagaaaag aacacacaaa ctcaacaaaa agtgggtgaa ggaawtgcts 180
aaargagagc atgtattcag ccagttaaca yatgaaaaaa aggtcctasa tcaactgawca 240
ttagagaaat gcaaatcaaa accacatga gatccatctt yayrocagtt agaaygggtga 300
tcattaaaz stcaggaac aacagatgct ggacaaggtg tca 343

<210> 73
<211> 321
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(321)
<223> n = A,T,C or G

<400> 73
gcactgagag gaacttcaa gagagagaga gaggtaaac ctgtacttgg ggagagaaac 60
agaaygtgag aaagtcttgg gtctggaagc agcttctaag atctttctat ttgttcaatt 120
tcaaaagttc catgttgaca aagtgccatc ctgtggggtg ctgtttctcg agtccacgtg 180
ataactcatt tatcacaggg agataccccc aaaaaagtg agcaaatctt aaaaaggttg 240
cttgagttca gacttaataa ccatcttga atgacacaga gaagaagga tgttgggttg 300
gagtgagatg agaccctaac g 321

<210> 74
<211> 321
<212> DNA
<213> Homo sapien

<400> 74
gcactgagag gaacttcaa gagagagaga gaggtaaac ctgtacttgg ggagagaaac 60
agaaygtgag aaagtcttgg gtctggaagc agcttctaag atctttctat ttgttcaatt 120
tcaaaagttc catgttgaca aagtgccatc ctgtggggtg ctgtttctcg agtccacgtg 180
ataactcatt tatcacaggg agataccccc aaaaaagtg agcaaatctt aaaaaggttg 240
cttgagttca gacttaataa ccatcttga atgacacaga gaagaagga tgttgggttg 300
gagtgagatg agaccctaac g 321

<210> 75
<211> 317
<212> DNA
<213> Homo sapien

<400> 75
gcactgagag gaacttcaa atgcactgag aaatgcattg tcaacaggac tgaagtcgtg 60
aaatcagttt ctcaagttca atctgattc aggtgtttac cagctacaca aacttaagca 120
agtcagataa ccttagcttc ctcatatgca aaatggaat gaagaatct catcgtgaa 180
ttgttttgag gattagaaaa acatcttgca tycagttaga attcaattag tatcatttt 240
cattcttcta aattaaaca ataggatttt tagtggtgga accttagaca ccaggaattg 300
gagtgagatg agaccct 317

<210> 76
<211> 244
<212> DNA
<213> Homo sapien

<400> 76

```

WO 01/90152

PCT/US01/16776

24

```

cgtaggggtc tctatccact cccactactg atcaaacctct atttatttaa tttattttat 60
cctactttta gttctgggat acsogtgcag catgcgcagg ttgttgcat aggtataaac 120
ttgccatggt ggtttgctgc accatcagt ccatcatcta cattaggtat ttctctaat 180
gctatccctc cctagcccc ttacaccccc aacaggctct agtgtgtgaa gtctctctca 240
gtgc 244

```

```

<210> 77
<211> 254
<212> DNA
<213> Homo sapien

```

```

<400> 77
cgtaggggtc tctatccact gaaatctgaa gcacaggagg aagcgaagca gtyctagtga 60
galggcaagt tctttacca cactctttta catttygttt agttttancc tttattttat 120
gataatcaag gttatattta ataatgattt attttaagge attccraat ttgcataatt 180
ctccttttgg agataccctt ttaetctcag tgcaggtctg gatcaagty ataszmagaa 240
gttctctctca gtgc 254

```

```

<210> 78
<211> 355
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)... (355)
<223> n = A,T,C or G

```

```

<400> 78
ttcgatacag gcaaacatga actgcaaggag ggtggtgacg atcatgatgt tgcagatggt 60
cgggatggnc acgaagaagc actgganccg gtgottacgt ccttttgctc tgttgatggc 120
cctgagggga cgcaggaccc ttatgacctt cagatcttc acaacgggag atggcactgg 180
attgantccc atgacacca gagaacccc aaccaccagn atatcantat attgatgtag 240
ttctgttga nggccccctt gggagagaaa gctccatnag ttggtcatct kaacaggat 300
ctcaacagtt tccagtggtt gtgatgggca tagtoatant taacnctgn togaa 355

```

```

<210> 79
<211> 406
<212> DNA
<213> Homo sapien

```

```

<400> 79
taagagggtg ccagcagaaa ggttagtato atcagatago atottatagc agteatatgc 60
ctgctatttg aagtgtantt gagagggaaa atttttaggt gctcacigec ctgacctgag 120
ccocagtgac agctaggatg tgcattctcc agccatczag agactgagtc aayttgttcc 180
ttaagtcaga acagcagact cagctctgac attctgattc gaatgacact gttcaggaa 240
cggaaicctg togtttagac tggacagctt gtggcaagtg aatttgctty taacagacaa 300
gattttttta aatttatatt gtaataatg tglgtgtgtg tgtgtgtata tatatatata 360
tgtacagtta totaagttaa tttaaaagtt gtttggtaac ctctta 406

```

```

<210> 80
<211> 327
<212> DNA
<213> Homo sapien

```

```

<400> 80
tttttttttt tttactogge toagtotaat cctttttgta gtcactcata ggcacagatt 60
aggactagga tgaatgattaa taagagggat gacataacta ttagtggcag gttagtgttt 120

```

WO 01/90152

PCT/US01/16776

25

```

tgttagggctc atgttagggg taaaaggagg gcaatttcta gatcaaatat taagaaggta 180
atagctacta agaaagaattt ttatggagaa gggacgggg ggggggatat aggttgaag 240
ccgcactcgt aaggggtgga tttttctatg tagccgttga gtigtggtag taaaaatga 300
ataattatta gtatgaagcc taggaga 327

<210> 81
<211> 318
<212> DNA
<213> Homo sapien

<400> 81
tagtctatgc ggttgattcg gcaatccatt atttgtgga ttttgtcat tgttttgca 60
attgcattca taaktattat tgaatttatg attgkatct ctaagtcatg gtatataatc 120
catgcttttt atgttttgc tgcattaac tttttctaga ggcatttcca cacaggatt 180
caataaatat taacacagtc tacatttatt tggatgaat tgcatactg ctgtactga 240
agcacattaa gtaacaaagg caagtgaaga gaatgaaga cactactcan aacagtata 300
atgattgcgc atagacta 318

<210> 82
<211> 338
<212> DNA
<213> Homo sapien

<400> 82
tcttcaacct ctactccac taatagcttt ttgatgact ctagcaagcc tggctaacct 60
cgcttacc ccactatta acctactgg agaactctc gtgctagta cccgttctc 120
ctgatcaaat atactctcc taattacag actcaacata ctagtccag cactatactc 180
ctctacata ttaccacaa caaatgggg ctactacc ccaccatta acaacataaa 240
acctcattc acacgagaa acacctctt gtctatcac ctatcccca ttctctctc 300
atccctcaac ccgacatca ttccggggtt ttctctt 338

<210> 83
<211> 111
<212> DNA
<213> Homo sapien

<400> 83
agccatttac ccccatcca caaaaaaa aaaaaaaag aaaaattatc aggaataaaa 60
atagactttg aacaaaaag sacatttgct ggcctgagga ggcataccc g 111

<210> 84
<211> 224
<212> DNA
<213> Homo sapien

<400> 84
tcgggtgatg cctctcagg caagaagat aaagcttcag accctaaca catttccaa 60
aaggaagaaa ggaagaaaa ggcatactc cccgttcca aggttcagg aggaagaaat 120
tgagggtgat tcaagattg cgaacaact atttgtgac aagcaggtg cagccggaga 180
ctggggagag cgaaccaat aggttttga gttctctca gtgc 224

<210> 85
<211> 348
<212> DNA
<213> Homo sapien

<400> 85
gcactgaggg gaacttcgtt ggaacgggt tttttctatg taaggctaga cagaagaatt 60

```

WO 01/90152

PCT/US01/16776

26

```

ctcagtaact tccctgtgtt gtgtgtatto aactcaase gttagaagat ccttacaca 120
gagcagactt gtaacactct twtltggaa ttgcaagtg pagatttcag ccgctttgaa 180
gtcaaggtta gaaaggaaa tatcttcta taaaactag acagattgat tctcagaaac 240
tcccttgtga tctgtgcgtt caactcacag agtttaacct ttctttcat agsagcagtt 300
aggaacact ctgtttgtaa agtctgcaag tggatagaga ccttaacg 348

<210> 86
<211> 293
<212> DNA
<213> Homo sapien

<400> 86
gcactgagag gaacttcyit gtgtgttttg yattcaactc acagagttga aawtswitt 60
acabagvksa ggcttkcnaa cactctttt gtmgastyg onagvgaak ttttrrcrc 120
tttgwggyow wysktngaan mgsuatac ttowystmra amctagacag aakcattctc 180
akaawstyyy ytgtgawgs tgcrttaac tcacagagkt kascmttyct kytsatrgag 240
cagttwkaa actctmttc ttggcttct gcaagtggt agagacccta aag 293

<210> 87
<211> 10
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer for amplification from breast tumor cDNA

<400> 87
ctcctaggct 10

<210> 88
<211> 10
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer for amplification from breast tumor cDNA

<400> 88
agtagttgca 10

<210> 89
<211> 11
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer for amplification from breast tumor cDNA

<400> 89
ttcgtttatg c 11

<210> 90
<211> 10
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer for amplification from breast tumor cDNA

```

WO 01/90152

PCT/US01/16776

27

```

<400> 90
tggttaaaggg                               10
<210> 91
<211> 10
<212> DNA
<213> Artificial Sequence
<220>
<223> Primer for amplification from breast tumor cDNA
<400> 91
toggatcatag                               10
<210> 92
<211> 10
<212> DNA
<213> Artificial Sequence
<220>
<223> Primer for amplification from breast tumor cDNA
<400> 92
tacaacggagg                               10
<210> 93
<211> 10
<212> DNA
<213> Artificial Sequence
<220>
<223> Primer for amplification from breast tumor cDNA
<400> 93
tggattggtc                               10
<210> 94
<211> 10
<212> DNA
<213> Artificial Sequence
<220>
<223> Primer for amplification from breast tumor cDNA
<400> 94
ctttctaccc                               10
<210> 95
<211> 10
<212> DNA
<213> Artificial Sequence
<220>
<223> Primer for amplification from breast tumor cDNA
<400> 95
ttttggtccc                               10

```


WO 01/90152

PCT/US01/16776

28

<210> 96
<211> 10
<212> DNA
<213> Artificial Sequence
<220>
<223> Primer for amplification from breast tumor cDNA
<400> 96
ygaaccacac 10
<210> 97
<211> 10
<212> DNA
<213> Artificial Sequence
<220>
<223> Primer for amplification from breast tumor cDNA
<400> 97
togatccagg 10
<210> 98
<211> 10
<212> DNA
<213> Artificial Sequence
<220>
<223> Primer for amplification from breast tumor cDNA
<400> 98
ggtactaagg 10
<210> 99
<211> 10
<212> DNA
<213> Artificial Sequence
<220>
<223> Primer for amplification from breast tumor cDNA
<400> 99
agtctatgag 10
<210> 100
<211> 10
<212> DNA
<213> Artificial Sequence
<220>
<223> Primer for amplification from breast tumor cDNA
<400> 100
ctatccatgg 10
<210> 101
<211> 10

WO 01/40152

PCT/US01/16776

29

```

<212> DNA
<213> Artificial Sequence

<220>
<223> Primer for amplification from breast tumor cDNA

<400> 101
tctgtccaca 10

<210> 102
<211> 10
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer for amplification from breast tumor cDNA

<400> 102
aaqagggtag 10

<210> 103
<211> 10
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer for amplification from breast tumor cDNA

<400> 103
cttcaacctc 10

<210> 104
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer for amplification from breast tumor cDNA

<400> 104
gctcctcttg ccttaccac 20

<210> 105
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer for amplification from breast tumor cDNA

<400> 105
gtaagtagag cagtgtag 20

<210> 106
<211> 20
<212> DNA
<213> Artificial Sequence

```

WO 01/90152

PCT/US01/16776

30

```
<220>
<223> Primer for amplification from breast tumor cDNA

<400> 106
gtaagtcgag cagtctgatg                20

<210> 107
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer for amplification from breast tumor cDNA

<400> 107
gacttagtg aaagaatgta                20

<210> 108
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer for amplification from breast tumor cDNA

<400> 108
gtaattcgc caacogtagt                20

<210> 109
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer for amplification from breast tumor cDNA

<400> 109
atggttgat gatagtgga                20

<210> 110
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer for amplification from breast tumor cDNA

<400> 110
acggggacc ctgcattgag                20

<210> 111
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer for amplification from breast tumor cDNA
```

WO 01/90152

PCT/US01/16776

31

<400> 111
tattcttagac cattcgtctac 20

<210> 112
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer for amplification from breast tumor cDNA

<400> 112
acctaccac tttagcgttc 20

<210> 113
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer for amplification from breast tumor cDNA

<400> 113
cgggtgatgc ctctcaggc 20

<210> 114
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer for amplification from breast tumor cDNA

<400> 114
agcatgttga gccagacac 20

<210> 115
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer for amplification from breast tumor cDNA

<400> 115
gacaccttgt ccagcatctg 20

<210> 116
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer for amplification from breast tumor cDNA

<400> 116
tacgtgcaa cactgtggag 20

WO 01/90152

PCT/US01/16776

32

```

<210> 117
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer for amplification from breast tumor cDNA

<400> 117
cgtttaggtc tctatccact                20

<210> 118
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer for amplification from breast tumor cDNA

<400> 118
agactgactc atgtccacta                20

<210> 119
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer for amplification from breast tumor cDNA

<400> 119
tcctogctcg gtgactcaag                20

<210> 120
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer for amplification from breast tumor cDNA

<400> 120
caagattcca taggctgacc                20

<210> 121
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer for amplification from breast tumor cDNA

<400> 121
acgtactggg ctgaagggtc                20

<210> 122
<211> 20
<212> DNA

```

WO 01/90152

PCT/US01/16776

33

```

<213> Artificial Sequence
<220>
<223> Primer for amplification from breast tumor cDNA
<400> 122
gaagcctggc cacttgacac                20
<210> 123
<211> 20
<212> DNA
<213> Artificial Sequence
<220>
<223> Primer for amplification from breast tumor cDNA
<400> 123
gtatcgagct agtggctctc                20
<210> 124
<211> 20
<212> DNA
<213> Artificial Sequence
<220>
<223> Primer for amplification from breast tumor cDNA
<400> 124
tagtgacatt acgacgtggtg                20
<210> 125
<211> 20
<212> DNA
<213> Artificial Sequence
<220>
<223> Primer for amplification from breast tumor cDNA
<400> 125
cgggtgatgc ctctcagga                20
<210> 126
<211> 23
<212> DNA
<213> Artificial Sequence
<220>
<223> Primer for amplification from breast tumor cDNA
<400> 126
atggetattt tggggggtg aca                23
<210> 127
<211> 22
<212> DNA
<213> Artificial Sequence
<220>

```

WO 01/90152

PCT/US01/16776

34

```

<223> Primer for amplification from breast tumor cDNA
<400> 127
ccggtatctc ctgtgggta tt                22
<210> 128
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer for amplification from breast tumor cDNA
<400> 128
ctgctgagc cacaatg                    18
<210> 129
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer for amplification from breast tumor cDNA
<400> 129
ccggaggagg aagctagagg ata            24
<210> 130
<211> 14
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer
<400> 130
ttttttttt ttag                      14
<210> 131
<211> 18
<212> PRT
<213> Artificial Sequence

<220>
<223> Predicted Th Motifs (B-cell epitopes)
<400> 131
Ser Ser Gly Gly Arg Thr Phe Asp Asp Phe His Arg Tyr Leu Leu Val
1      5      10      15
Gly Ile

<210> 132
<211> 22
<212> PRT
<213> Artificial Sequence

<220>

```

WO 01/90152

PCT/US01/16776

35

```

<223> Predicted Th Motifs (B-cell epitopes)

<221> VARIANT
<222> (1)...(22)
<223> Xaa = Any Amino Acid

<400> 132
Gln Gly Ala Ala Gln Lys Pro Ile Asn Leu Ser Lys Xaa Ile Glu Val
1      5      10      15
Val Gln Gly His Asp Glu
20

<210> 133
<211> 23
<212> PRT
<213> Artificial Sequence

<220>
<223> Predicted Th Motifs (B-cell epitopes)

<400> 133
Ser Pro Gly Val Phe Leu Glu His Leu Gln Glu Ala Tyr Arg Ile Tyr
1      5      10      15
Thr Pro Phe Asp Leu Ser Ala
20

<210> 134
<211> 9
<212> PRT
<213> Artificial Sequence

<220>
<223> Predicted HLA A2.1 Motifs (T-cell epitopes)

<400> 134
Tyr Leu Leu Val Gly Ile Gln Gly Ala
1      5

<210> 135
<211> 9
<212> PRT
<213> Artificial Sequence

<220>
<223> Predicted HLA A2.1 Motifs (T-cell epitopes)

<400> 135
Gly Ala Ala Gln Lys Pro Ile Asn Leu
1      5

<210> 136
<211> 9
<212> PRT
<213> Artificial Sequence

<220>
<223> Predicted HLA A2.1 Motifs (T-cell epitopes)

```


WO 01/90152

PCT/US01/16776

36

```
<221> VARIANT
<222> (1)...(9)
<223> Xaa - Any Amino Acid

<400> 136
Asn Leu Ser Lys Xaa Ile Glu Val Val
1 5

<210> 137
<211> 9
<212> FRT
<213> Artificial Sequence

<220>
<223> Predicted HLA A2.1 Motifs (T-cell epitopes)

<400> 137
Glu Val Val Gln Gly His Asp Glu Ser
1 5

<210> 138
<211> 9
<212> FRT
<213> Artificial Sequence

<220>
<223> Predicted HLA A2.1 Motifs (T-cell epitopes)

<400> 138
His Leu Gln Glu Ala Tyr Arg Ile Tyr
1 5

<210> 139
<211> 9
<212> FRT
<213> Artificial Sequence

<220>
<223> Predicted HLA A2.1 Motifs (T-cell epitopes)

<400> 139
Asn Leu Ala Phe Val Ala Gln Ala Ala
1 5

<210> 140
<211> 9
<212> FRT
<213> Artificial Sequence

<220>
<223> Predicted HLA A2.1 Motifs (T-cell epitopes)

<400> 140
Phe Val Ala Gln Ala Ala Pro Asp Ser
1 5

<210> 141
<211> 9388
```

WO 01/90152

PCT/US01/16776

37

```

<212> DNA
<213> Homo sapien

<400> 141
gctgcggccc ggcagctcaa ttaaccccca ctaaaaggag tgcactoyat cagactgtta 60
ctgtgtctat gttagaagaa gtagacataa gagattccat ttgtttctgt actaagaaaa 120
attctctctg cttgagatgc tgttaacttg taacctagc cccaaacctg tgctcacaga 180
gcactgtgct gtgttgcttc aaggttccat ggatttaggg ctatgctttg ttaaaaaagt 240
gtttgaagat aattgtgctg ttaaaagtca tcaacttct ctactctcaa gtacccaggg 300
acacatacna ctgcgggaag ccgcagggaac ctctgtctag gaaagccagg tattgtccaa 360
gattctccc catgtgatag cctgagatat ggctctatgg gaaggttaa cctgtactgt 420
cccaagagcc gaactccccc agccagacat ccccgagcc gacacccga baggtctgt 480
gctggggagg cttagtcaaa gaggaaggcc tctttgcagc taggttaaga ggaaggcatc 540
tgttccctgc tcttccctgg caatagaat gtcttggtc aaaaaccgat tgtatgttct 600
acttactgag ataggagaaa acatccttag ggttggaagt gacacagctt ggcggcaata 660
ctgtctctta atgacccgag atgtttgtat aagtgcacat caaggcacag cactttctct 720
taactctatt tatgacacag agaccttgt tcaagtttct ctgctgaccc totcccacat 780
attaccctat tggcctgcca catcccccct tccgagatgg tagagataat gatcaataaa 840
tactgaggga actcagagac cagtgtccct gtatgtctct cgtgtgtcga ggcgggtccc 900
cttgggctaa otttttttct tctatacttt gtctctgtgt ctctttcttt tctcagttct 960
tcttccccc tgaagagaaa taaccacagg tgtgaggggg caggccaccc cttaataaat 1020
tactagacct gttcctgac aacaaagact gtggtgcaga aggttgggtc ttggtgttca 1080
ccgggtggcc ggcattggcc aggtgagag gtctccaggg cctggtgcaa atctccaaag 1140
aaatgcagga aaacagacaa aggtgagtg taattttga tttaggcgg caggttagcca 1200
ttccagcgca aaatgcgca ggaagcttt tgcgtgtct gttagcaggt aggcaccaag 1260
cactctctat tggctaaltt ggaagggaac tgcactcca ttgctgaaa tctccgcta 1320
tttagagcag actgagcgg ttccctttct ctggttgcc tggaaacgga ctgtctgctt 1380
agtaacatct gatcaagttt cccattggcc gcccgttccc gaagcccgcc ctccatttc 1440
cggaagccct ggcgaagttt ggtctgcagg tgcctccag gtgcaagtg ggaagtgtga 1500
gtctcagctc ttggctatt cggccacgtg cctgcgcgac atgggaagct ggaggtcag 1560
cagcgtggag tctggcctt ttgctccac ggttgggaaa ttggccattg ccacggcggg 1620
aactggtgac caggtgccc cccggccgtt tctcactcgt ccaacggact cgtggggcgt 1680
cgcaactggc ctgagtgtt tctcgaact ctgacccgta ttgtctccag attaaagtta 1740
aaaaaggggc tttttcagcc caatcggtta aaacgcttt tgaattctag gcaggtgttt 1800
tgttgcagcc ctgggagggg gtgacccgca ggttgaggtt tattaaata cattctggt 1860
ttatgttatg tttaataaa agaccccaaa ccttaacaaa altcacttt ttgcaatgg 1920
tatttttag tggactgtct ctgataaaga cagcaagtta aaatggant ttgttgtg 1980
taataaacc aatttttagt ttgtgttt gtctaatag caacaacttc ttaggcttta 2040
taaaacata tttctgggg gaaatttctg tgaaggccac agcgagttag tttagantg 2100
tttaaaaga agtaagttcc bgtttttgat atcttagtag tgaatgccc aacctggtt 2160
tactaacccc tgtttttaga ctctcccttt ccttaactca cctagccttg ttccacttg 2220
aattgactct ccttagcta agagcgccag atggaactca tcttgctct ttcaactgga 2280
gccccttctc caaggactta actgtgcaa gctgaactcc agcaactcca agaattgaat 2340
taantgttaa gatactggg caagctatat cgcagttccc gaggaattca tccgattgat 2400
tatgcccaca agcccccgt ctatccctt gtaataatct taagccccc gcacctggaa 2460
ctattaaact tctgtaacct atttatcctt ttaactttt tgottacttt atttctgtaa 2520
aattgtttta actagaactc cctccctctt totaaaccaa agtataaag aagatctagc 2580
ccctcttca gacggagag aattttgagc attagccatc tottggcgcc cagctaaata 2640
aattgacttt taattgtct caaagtgtgg cgttttctct aactcgctca ggtacgaact 2700
ttggaggccc cagcgagaaa ctccacccgg agaaactgca cccggcgaga gccggcccc 2760
ctgtgtgctc cccgggaagg acagccagct tgtagggggg agtgccactt gaaaaaaaa 2820
tttccaggtc cccazaggtt gacctcttc ccgaggacag cggatngact aacctgggg 2880
tgccacccaa aattccacct ctgactctc aactgtgac cccgggttca ggtaggtag 2940
atttgaattt ggtttgaca gagggaagcg accctgatga ggtgtctct ottttgactc 3000
tgcctcttct tctagatgc tagagggtag agccctggtt ttctgttaga cgcctctgtg 3060
tctctgtctg ggaagggaag ggccttgaca gggcccatcc cttgagtagc tccacatccc 3120
aggatgctgg ggaactgaat cctgtttctt ggcagactgg tctctctct tctcttttct 3180
tatctctaat ctttctgtt tcaagtttct tggagactct ctgggaagga aaaaagaaa 3240

```

WO 01/90152

PCT/US01/16776

38

acgtgtatbaa abctctgtgtg naatgggtgaat gaatggggga ggacaaggac ttgcgcttgt 3300
 cctccagttt gtatctccac ggogaagct acggagttca agtgggcoct cactcgnggt 3360
 tccgtgggga cctcataagg chtaaggcag catccggcat agctcgatcc gagccggggg 3420
 ttatatacgg cctgtcaatg claaaggagg oncaagtcce ctlaaggggga gggccaggac 3480
 gggcatctga ctgacccctt caccggagcc cctcccttg tttgtctaaa aaaaaaata 3540
 gaagaactgt tcatatctgt ttacatgccc tagggtaaac ttgtgtttt atgtttattg 3600
 ttctgttctg tgtctattgt cttgttttagt ggttgtcaag gttttgcatg tcaaggacgtc 3660
 gatattgccc aagaactgtg ggttaagaact tctgcaaggt ccttagtgct gattttttgt 3720
 caagggaggt laaattttct atcaatcatt taggtggccc acccaagtc tgtcttttct 3780
 gccagagga agtcagggtg tgttaccggga atgagtgtaa aaaaaaatc gctgatttg 3840
 gatttctgc aocatgatg ttgtatttag attgtctaac ccccaatcca gtttgattgg 3900
 acctctctta aactaaacta gtgttaggtt caaaacagcc aocctgcaaga ttctgtgt 3960
 cactcttttg gtacattctgt aacttttct gtgccttaa atagucact gtgtagggaa 4020
 acctaccctc gtaactgttt acttcttcta gattcttact ctgttctct gtgtgtaact 4080
 tcccatctta aaaaagatcc aagtggtctt ttctctctc cctgcccctt accccaacaa 4140
 tetctgttct cagtgcgaca gcaagttcag cgtctccagg acttggtct gtctcactc 4200
 ctgaaacctt laaaagaaaa agctgggttt gacttattg ccttgagtc atggagagac 4260
 aaaaagttat taggttacag ctctagaaga agggagagaa cacttagatc caactgacct 4320
 agagagatct gggctggcct ctactctctc tccctcaatc ttaaaagtac agtgatgtg 4380
 caagtgtgat tttagctgttg tggtttttct gctctttctg gtctatgtga tctgttctt 4440
 tctgactctc agcccccag gggatgagtt tctctgtg tctgtgttt gatctctatg 4500
 ttcaaaactt ataaatttgc ctcaaaata aaaaaaata gggaaacact tctctccagc 4560
 ctgttaaggg ttggagccct ctccagata tcttctctc ttgttctcag gtttctcag 4620
 aggtttatg agtccgctt aaaaaagga agctctggac actctgctaa gtatattgg 4680
 caaagtttgg agttgagtg ccccttgag gctccctgaa cctcaaat gtctcaagtg 4740
 tgtggcgggt tgtactgaa actccggcc tccctgata gtttctctc attgatcat 4800
 ggctgagttt ggtcaggagc accccttcca tggctcaact catgacact tcaataattt 4860
 acctccaaag tctctctgag ccagacogtg ttctgcctc gacccctcag cgttctcagc 4920
 cgcctgttac tgcctctctc tgaagaagag gagagttctc ctacccagct cccacgcct 4980
 caaaacagc ctactcctt aggtctatcc catgtctctc cggctatgtc cctgtaggc 5040
 tcatcaacaa ttgctctgtt gttgcaaccg ttgtggagag aagtagctcc totactacca 5100
 ctgagagagc cccagctccc tctgggtgat gactgtctca cccctctctt ggtttatgtc 5160
 ccttcttctt acttctgact tgtataattg gaaaaacat aatctctct tcttgaaaa 5220
 gcccccgggt ttgcctctac tgaaggagtc tgaactctg acaaatggc caaactggga 5280
 tgaactctca cagctctctt tgaactctt caactctgaa gagaggggaa gtaactcaag 5340
 agagggcaaa aagtaaaccc tcaactcaaa caatagggag gaggaggaag ctaggaggaat 5400
 agtgattaga gacccaattg ggaactaatt gggacccaat ttctcaaggt ggaaggaga 5460
 cttttgacga ttctccacgg tatctctctg ttggatctca ggaagctgct cagaacacta 5520
 taactctgtc taaggcagct gaactgtctc aggggcatga tgaactcaaa ggaagtgttt 5580
 tagagcactt ccaggaggtt tctcgattt acacccctt tgaactggca gcccccgaaa 5640
 atagcaatgc teltaatttg gacttctgtg ctacggcagc cccagtgagt aaaaaggaaac 5700
 tccaaaact agggggttt tcttggaatg aataccagtc agctttaga gatagctaa 5760
 aaggtttttg acagtaacga ggttgaaaaa caaaaacag cagctcagcc agctgaaaaa 5820
 agccactga aaagcatctt ggaatataag agtttactgt tagatagcc tcaattgact 5880
 tccctctcca catgtgtttt aaatcagct acactacttc ctgactcaaa ctccactatt 5940
 cctgttctat actgtcagga actgttggaa actactgaaa ctggccgacc tgaactctaa 6000
 aatgttcccc taggaaggtt ggaatgcacc gttgtccagc acagtagcag ctctctcag 6060
 aagggactac gaaggccgg ttgactgtgt accatggaga cagatgtgtt gtgggtctag 6120
 gcttctcag caaacacctc agcaaaaag gctgaattga tctgctctac tcaagctctc 6180
 cgaatgggta agatattaa cgttaacct gaacgaaggt acgcttttgc tactgtcat 6240
 gtacgtggag ccatctacca ggaagctggg ctactcaact caacaggtg ctgttaacca 6300
 ctgttaaggga catcaaaagg aaaaacaggg ttgtccctgt ggttaacaga aagtgatctc 6360
 agcagctcaa gatgcagtg actttcagc caagctctca aacttctgc caacagctct 6420
 ctttccacag ccagatctgc ctgacaaac cgtatctca acagaagaag aaaaagggc 6480
 tcaagaacta gagccatcaa aaatcaggaa ggttgggtga ttctctctga ctctagaatc 6540
 ttcatatccc gaactcttgy gaacacttta atcagctccc tacagctctac caacctatta 6600
 gggaggacaa agctactctc gctctctcgg agccttitta agatcccaaa tcttcaaacg 6660
 ctacacagatc aagcagctct cgggtgcaca aactgcccac aggttaaatg caaaaagggt 6720

WO 01/90152

PCT/US01/16776

39

cctaaaccca gccacggcca ccgtctccaa gaaactccac caggagaaaa gtgggaaatt 6780
 gaotttacag aagtoaaacc acacggggct gggtacaaat acottctagt actggtagnc 6840
 acottctctg gatggactga agcatttgct accaaaaacg aaactgtcaa tatggtagtt 6900
 aagtttttac tcaatgaat catccctcga cgtgggtcgc ctgttgccat agggctctgat 6960
 aatggacccg ccttcgccct gtctatagtt tagtcagtca gtaaggcglt aaacattcaa 7020
 tggagactcc atgtgtccca tcgacccccc agctctgggc aagttagaac catgaactgc 7080
 accctaaaaa scactcttac aaatttaac ttagaaaccc gtgttaattg tctaagttct 7140
 otctctttag cctactctag agtaaggtgc acccttact gggctgggtt cttacttttt 7200
 gaaatcatgt atgggaggga gcggccatc tgccttaagc tangagatgc ccaattggca 7260
 aaattctcac aaatttaatt attacagtc ctacagtctc ccaaacaggt acaagatato 7320
 atctcgccac ttgttcgggg aaacctccac atctcaattc ctgaacagac agggccctgc 7380
 catctattcc cccacgggtga cctgttggtt gttaaaagtc tcaagagaga aggaactcct 7440
 cctgtttgga agagacctca caccgtatc acgagccca cggctctgaa gctgagtgcc 7500
 attctgtgct ggatttaca ctcccgcatc aaaaaggcca acggagccca actagaaaca 7560
 tgggtcccca gggtgggtgc agggccctta aaactgcacc taagtgggtg gaagccatta 7620
 gattaattct tttcttaatt ttgttaaac aatgcatagc ttctgtcaa cttatgtate 7680
 ttaagactca atataacccc ctgtttataa ctgaggaato aatgatttga tttcccaaaa 7740
 acacaagttg ggaatgtagt gtccacccgt gtttttacta accctgtttt tagactctcc 7800
 ctttcttcta atcaactcgc ctgttttccc cctgaattga ctctccctta gctaaagacg 7860
 ccagatggac tccactttgc ctcttcaact ggaagcggct tctcaagga ctttaactgt 7920
 gcaagctgac tccacagcac tccaaagatg cautttaact ataagatact gtgcgaactg 7980
 atctccacag tcccaaggaa ctgttcaact tgattacacc caaaagccccc gcgtctatca 8040
 ccttgtaact atcttaaaag cctgtcaact ggaactatca acgttctgtt aaacttttat 8100
 ccttttaact cttttgctca ctltatttct gtaaatgtt ttaactaga ccccccctct 8160
 cctttctcaa ccaaggtata aaagcaaatc tagccctctc ttcaagcgca gagaatttgc 8220
 agcgttagcc gtctcttgcc caccgcctaa ataacggat tctctatgtg tctcaaatgc 8280
 tggcgttttc tctactcgc tcaggtacga ccgtggtagt atttcccca acgtcttatt 8340
 tttagggcac gtatgtagag taacttttat gaaagaaccc agttaagcag gttttgggat 8400
 ttcttttctc aactgtaata ctggttttga ttttttatt attttttat tttttttgag 8460
 aaggagtttc actctgtgtg cccagcctgg agtgcactgc tgcgatcttg gctcaclgca 8520
 acttccgcat cccaggttca agngattctc utgctcagc ctgcagagta gctgggatta 8580
 taggcactgc ccccaacac cagctaattt tttattttta gtaagatgg ggtttcttca 8640
 tttgtgtcaa gctgtgtgtg aactccccc ctcgggtgat ctgcacgact cggcctccga 8700
 aagtgctggg attacaggtg tgatccacca caccagccgc atttatatgt atataatca 8760
 catctctca accaaaagt agtttcttcc tccatcttga atatagctg tagaccctgt 8820
 gggtaagga catgttcaa catgttaacca cagcagtttt taigtctatc gacagcatct 8880
 ccaatagccc ttcattgttg tcaatgtctc ccaagcaat tccaataaac tcttccactc 8940
 gatgaactcc tacttgcatt tgttacttaa tttgttaagg tggctgttac agaaactatt 9000
 agtatgtcac gaattacacc aaattttagt ggtcaaaccc atcattttat taigtatgtg 9060
 gattctcatg gtcaggtcag gatttcagac aggcacacag ggtagccca tttgtctgtg 9120
 ctatgatgtc tggcctcagc acaggagact caacagctgg ggtctgggac catttggagg 9180
 cttgttccct caactctgat acctggcttg ggaatgttga agaggggggt agotgagact 9240
 aggtgtctat atgtagtgtt tccatattgc cttgaacttc ttacagccctg gcagccctcag 9300
 ggttagtaca attcttagga ggcacaaggg tcaagggcag atgctgaagg ctcttttatg 9360
 aggtagcaca gcaaatccac ccaggatc 9388

<210> 142
 <211> 419
 <212> DNA
 <213> Homo sapien

<400> 142
 tgaagtcca gcaatgtgat ggaaggatg gtctttggag agagcatato catctctcc 60
 tcaactgtcc ctactgtcat gaaatcaact gacagacatt aaacagggtta gtcttaacca 120
 caactatttt agctaccttg tcaagtaaat ggttaagaa cacttttgggt ttacacttgt 180
 tgggtatag aagttgcttt ccgcactcac gcaataagtt tgtgtgtaat cagaaggagt 240
 taccattagg tttaagtgtc attctttagt taacttggga gctgtgtaat ttagggtttg 300
 cgtattattt cacttttgtt ctccacttat gaagtgtatt tgtgttcgag tgtgtgtgcg 360

WO 01/90152

PCT/US01/16776

40

tgcgcattgtg ctccggccag ttaacataag caaataccca acatcacact gutogactt 419

<210> 143
<211> 402
<212> DNA
<213> Homo sapien

<400> 143
tgtaagtca gcaagtgat gtccactca gtgtgttggat gggaaagtt aatgagcaaa 60
ttgtatacaa tggctagtac attgacggg atttgttga gctgggtgag: gttatgactt 120
agcctgttag actagtctat gcaatggct ctggtaact acggtctct cacttctcca 180
gataaatccc caatgcttta tattctcttc ccaacatact atctctatca ccacatagtt 240
cctttgttaa tgcctttgtc tagactttcc cttttctgtt ttottattca aacctatct 300
tctttgcata gattgtaat tcaaatgccc tcaagggtga ggcagttcat gtaaggagg 360
gaggtacgac agtgcagatc gcatcacact gctgactta ca 402

<210> 144
<211> 224
<212> DNA
<213> Homo sapien

<400> 144
togggtgag ctctctcagg ccaagaagat aaagcttcag accctcaaca catttccaaa 60
aaggagaaa gggaaaaaaa ggcacatcct ccggttcaga aggttcaggg agggagaaat 120
taagggtgat tcaagagttg cggacaactc ctttgatgcc aagcaggtg cggccggaga 180
ctggggagag cgaagcaatc aggtttttaa gtctctctca gtgc 224

<210> 145
<211> 111
<212> DNA
<213> Homo sapien

<400> 145
agccatttac caccatcca caaaaaaaa aaaaaaaag aaaaatctca aggantaaa 60
atgactttg aacaaaaagg acaatttgc tggcctgagga ggcctcacc g 111

<210> 146
<211> 585
<212> DNA
<213> Homo sapien

<400> 146
tagcatgttg agcccgaca cttgtagaga gaggaggaca gttagaagaa gaagaaaagt 60
ttttaaatgc tgaaggttac taaagaaag ctttggcttt ggtgagact tttaaagatg 120
cagaggatgc tttagcagaa cttcataaat atatgcaggt gattccttat ttctctctag 180
aaatttagtg atatttgaat taatgcccac aattaatatt ctctcgagga aactattct 240
acattacta agtaaggcat tatgaaggt ttctttttag gtatagtttt tcttaattgg 300
gtttgacatt gcttcaatgt gctctgttt ttgtccataa tgaagatga agtagctgt 360
ggaaaacta ttacataaat ttgtgtatgt gttttgagaa atgtcctat agggagctca 420
cctggttggt tttaaatat ttgtgtact ataatgggc taattatata aaccttttg 480
agaentattt taaattgtct ttctctgtaa tactgatgat gatgttttct catgcatttt 540
cttctgaatt gggacattg ctgctgtgtc tgggtccaca tgcta 585

<210> 147
<211> 579
<212> DNA
<213> Homo sapien

WO 01/90152

PCT/US01/16776

41

```

<220>
<221> misc feature
<222> (1)...(579)
<223> n = A, T, C or G

<400> 147
tagcatgttg agcccaagaca ctgggcagcg qgggtggcca cggcagctcc tgccagcccc 60
aagcgtgttt gtctgtgaag gacccctgacg tcaactgcca ggcctagggag gggccaatgt 120
ggagtgaaatg ttcaaccgact ttccnaggag ttgccaagaag ccaggtgcaa cttggtttgc 180
ttgtgtcaat caaccctcaa gatatgcaca ctgcttcca aataagcat caactgtcat 240
ctccagatgg ggaagacttt ttctcaaac agcagggcag ttccatcca ctccagacac 300
agcacgtcca cttctctggg cagcaccaag tccccaact ctgctggtt caagtgatg 360
atgcaagaca agccgtttctg cagacccaag tcccctgtgt gctgtgcat ctcactggcc 420
tccacccggt acacccgtct aggcgcgcga tanlgtgcac agaanaaatg atgatccagt 480
cccacagccc acgtccaaaga ngactttatc cgtcagggat tctttattct gcaggtatgc 540
ctgtgtgatt aattgttgt gtctgggctc aacatgcta 579

<210> 148
<211> 249
<212> DNA
<213> Homo sapien

<400> 148
tgacaccttg tccagcatct gcaagccagg aagagagctc tacccaagat ccccacccc 60
ttggcaccag gctcttggac ttccatctc cgaactgtg agaataagt attgtgcct 120
aataaatct ttgtgttttc agatatttag ctatagcaga ccaggtgac taagagaaac 180
ccataaagag ttacatactc attaatctc gtctctatcc ccaggtctca gatgtggac 240
aaggtgtca 249

<210> 149
<211> 255
<212> DNA
<213> Homo sapien

<400> 149
tgacaccttg tccagcatct gctattttgt gaatttttaa taatagccat totgactggt 60
gtgagatggt aactcaattgt ggttttggtc tgcatttctc taatgctcag tgatattcag 120
ctttttttaa atatgcttgt tgaccccatg tatatcatct tttagaagt gtctgttcat 180
atcctttgac caatttttaa tttttttatc ttgtasatit gtttaatttc cttacagatg 240
ctggacaagg tgcca 255

<210> 150
<211> 318
<212> DNA
<213> Homo sapien

<400> 150
ttaagctgca aactgttggg gcccaagctg ggtacacttc ttcattotaa ctggagagga 60
gggaagtica agtccagccag aggttggttg ggtagacagt ggcactcaga aatgtcagct 120
ggccccctgt ccccgcatag gccagacagc aaggtcttgg ctctccagg ccagctgaag 180
aacagggcac tttctccgct gccacaaagc gtcagagact cccatctttg aagcaaggcc 240
ttctgtctct tctgtcaact cctgtttctg tttagaacct ggttatagac aaggtctctc 300
cacagtgttg cagcgtaa 318

<210> 151
<211> 323
<212> DNA
<213> Homo sapien

```

WO 01/90152

PCT/US01/16776

42

```

<220>
<221> misc_feature
<222> (1)...(323)
<223> n = A,T,C or G

<400> 151
tncgngcgc acnntgtaga ganggnaagg cnttcccac attnccctt catnanagaa 60
ttattonacc aagmntgacc natgccttt atgacttaca tgennactnc ntaatctgtt 120
tcnngcctta aaagennntc cactacatgc ntcancactg tntgtgtnac ntcatnaect 180
gtongnaata ggggencata actacagaaa tgcanttcct actgtctcca ntgcctccg 240
cgtgtgggct tncctactct tcttntatc caagtatcat ctctgganbg cttcccact 300
ctccacttg ttgcagcnet aat 323

<210> 152
<211> 311
<212> DNA
<213> Homo sapien

<400> 152
tcaagattcc ataggctgac cagtcacaag agagttgaaa tcatgaagga gactctatct 60
ggagagagct gtatgtttga ggtgtgcaaa gacttaggat gtagttggtg ggtgtggtta 120
gtctctaaag ttgattttgt tcatataatt catgccctga atgccttctg tgcctcacc 180
tgytccaaag cttagtgaac acataaaagt ctctgtcttc ttgctctcca aactctctc 240
gagatttcc tcaagttgtc taccattcaga tcgaagccag ttggcacaac agatgcagtc 300
cagaggggtca g 311

<210> 153
<211> 332
<212> DNA
<213> Homo sapien

<400> 153
cagaattcca taggctgacc aggagcttat tcaagatctc tggcagttga ggaagtctct 60
ttaagaaaaat agtllaaaca atttgttaan atttttctgt cttaactctat tctgttagca 120
gttgatatct ggtgtctctt ttataatgc agagtgaggaa ctttccctac catgtttgat 180
aaatgttgc caggtctcat tgcctataat gtgtgttcca aaatgcctgt ttagttttta 240
aagacggaaac tccacctttt gtttgtctct aagtatgtat ggaatttat gataggacat 300
aqlagtagcgt gtgttcagcc tatggatctc tg 332

<210> 154
<211> 345
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(345)
<223> n = A,T,C or G

<400> 154
tcaagattcc ataggctgac ctggacagag atctctagg tctggccag gacacaggg 60
tcaagctcag tggagaaggt ttccatgacc ctacagattcc cccaaacctt ggaattgggtg 120
acattgcate tctctagaga gggaggagat gtanctctgg gcttccacag ggaactggta 180
ttttgggac aggttacccg tggcctcagg cttgcatcat tcaanagcctg ggggtggaa 240
ggctggcagc ctgtggccc attgaaatag gctctggggc actccctctg ttctanttg 300
aacttgggta aggaacagga atgtggctcan cctatggaaat ctbga 345

```

WO 01/00152

PCT/US01/16776

43

```

<210> 155
<211> 295
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(295)
<223> n = A,T,C or G

<400> 155
gaagcttggc cacttgacac attaaacagt ttgcataat cactanostg ttttttagt    60
ttgctgtctg ctgtgatgac ctgcctgat tctctggcgt taatgatggc aagcctaato    120
aaacgclgtt ctgttaattc caagttataa ctggcattga ttaagcattt gtctttcaca    180
actaaactgt tcttcataaa acagcccata ttattatcaa attaagagac aatgtattcc    240
aatatccctt anggccaata tatttnatgt cccttaatta agagctaotg tccgt      295

<210> 156
<211> 406
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(406)
<223> n = A,T,C or G

<400> 156
gaagcttggc cacttgacac tgcagtggga aaaccagcat gagcngctgc ccccaaggaa    60
cctcgaagcc caggcagagg accagccatc ccagcctgca ggtaaagtgt gtcacotgtc    120
aggtgggctt ggggtgagtg ggtgggggaa gtgtgtgtgc aaaggggtgt tnaatgtcta    180
tgcgtgtgag catgagtgat ggtagtgtgt acgtcatgtc agggagtgtg aaacagcgtg    240
cgggggtgtg tgtgcaagtg cgtatgcata tgagaatatg tctctgtgga tgagtgcatt    300
tgaagctotg tgtgtgtgag tgtgtcatgt anggtaanit antgactgag caggatgtgt    360
gggtgtgcat ggaacactca ntgtgtgtgt caagtggcon aocgtc      406

<210> 157
<211> 208
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(208)
<223> n = A,T,C or G

<400> 157
tgaagcttgg ccaactgaaa cactaaaggg tgttactcat cactttcttc tctctgggt    60
ggcatatgag tgcctctatt cacttgccac tcaattgttt ggcagtgaot gtaancana    120
tctgatgcat acacacgctt gtaaatgaa taatgtctc taatactatg tgcacacat    180
anggtanagg tgaaggagag gggagaga      208

<210> 158
<211> 547
<212> DNA
<213> Homo sapien

<220>

```


WO 01/90152

PCT/US01/16776

44

```

<221> misc_feature
<222> (1)...(547)
<223> n = A,T,C or G

<400> 158
cttcaacctc cttcaacctc attcaacctc ctggattcaa acaatcatcc cactcagac   60
tccttagtag ctggagactac agctcaccgc cactacatct ggctaaattt tigttagagal  120
agggtttcat cctgttgccc tggctggctc caaactcctg aactcaagca atgtgcccac  180
ctcagacctc caaagtgtgc ggattacagg oataagccac catgcccagt cctnttttaa  240
ttttctccac caactcttta ccacatttcc ttttatgttt agatacataa atgcttaccg  300
ttatgataca attgcccaca gtatttagac agtaaatgc tgcacagggt tctagcctag  360
gaacagttag caataccaca tagcttaggt gtgtgttaga ctataccac taggtttgtg  420
taagttaacc tttatgtgtt ttacacatg acsaaccat ctatgatgc atttctcaga  480
atgtatccct gtcagttagc tatgatgac agggaaact gcccaaggac aongatattg  540
tacctgt.                                     547

<210> 159
<211> 203
<212> DNA
<213> Homo sapien

<400> 159
gctctctctg ccttaccac tcaccagta tgtcagcaat tttatcgtct ttaactacga   60
aacagccctg atccaaacac ttaacacact aacttgaaa gthcaggcaa caatgccttt  120
ctcatgggtc tctctgctcc agttctgaac cttctctttt tcttagaaca tgcatttarg  180
tcgatagagg ttctcttcag tgc                                     203

<210> 160
<211> 402
<212> DNA
<213> Homo sapien

<400> 160
tgtcagtcga gcaagtctgat gggctggaac gggctgtaag cagtaattgc aaactgtatt   60
taacacataa taataatatt tagcatttat agagcacttt atatctbcaa agtactttga  120
aaoattayct aattaatatc cctctctgat tatactctgg atacaatgc aottaaacto  180
aggaacagggt catgagaaaa gtatgcattt gaagtttgt gctagotctg olttaaaaa  240
ctatcacatg atgggaaagt tagagttcag attctgttgg actgttttg tgcatttcag  300
ttcagcctga tggcgaattt agatcatatc tgcactcgat gactytgctt gataacttat  360
cactgaatcc tgcgtgttga tcatcacact gctcacttia ca                                     402

<210> 161
<211> 193
<212> DNA
<213> Homo sapien

<400> 161
agcaatgttg gccacagcac tgaccaggag aaaaaaacac caatagaaac acgcccagac   60
actgacccgg agaaaaacca accaataaaa acagggccgg acataagaca aataataaaa  120
ttagcggaca aggaatgaa aacagctatt gtaagagcgg atatagcgtt gtgtgtctgg  180
gctcaacatg cta                                     193

<210> 162
<211> 147
<212> DNA
<213> Homo sapien

<400> 162

```

WO 01/90152

PCT/US01/16776

45

```

tgttgagccc agacactgac caggagaaaa accaaccat aaaaacaggg caggacataa    60
gacaaataat aaaattagcg gacaaggaca tgaaccagcg tattgtaaga cgggatatag    120
tggtgtgtgt ctgggtcaa catgcta                                     147

```

```

<210> 163
<211> 294
<212> DNA
<213> Homo sapien

```

```

<400> 163
tagcatgttg agccacagca caaatcttct cttagccat aaatcatttc tgeatatgtt    60
tttaaaacaa cagctagccc atgattatct aaaaagacta ttgtatgggg tatittgatt    120
tgggttctta tctcctcacc attatcttca ttctatctat tgacctctta tcccagagac    180
tctcaaaactt ttatgtttata caaatccatct tctgtctcaa aaatatcttc aaccattctt    240
ctctgtcttc tgcgtgtgta tgtgtgtgtg tgtgtgtctg ggtccacat gcta       294

```

```

<210> 164
<211> 412
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1) .. (412)
<223> n = A,T,C or G

```

```

<400> 164
cgggattggc tttagctgac agatgctgac tgtgacgca cccggcgtgg aacagaaac    60
caccctggct caagtgcgcc agagccgcccc tgaactcgtg ctgctgtggg gctggggcgt    120
gatgaactcc accgccctga aggaagccca gccacacgga taccctcccg ccaagatgta    180
cggcgtgttg tggcccggtg cggagcccca tgtgctgac gtggggcgaag gcgccaaagg    240
ctacaacggg ctggctctga accgctacgg caagcagctc aaggtgatcc angacatcct    300
gaacacgctg caagacaagg gccagggcac ggggcccaaa gacgaagtgg gctcgtgtct    360
gtacacccgc ggcgtgatca tccagatgct ggacaagggt tccatcacta at         412

```

```

<210> 165
<211> 361
<212> DNA
<213> Homo sapien

```

```

<400> 165
ttgacacott gtccagcctc tgcctctgat gagagcctca gatgcttacc actaatggca    60
gaaggcuaag gagaacaggg attgtatggc agaaaggaa gaaagagaga ggggagaaag    120
gtcgtaggct cttttcaaca accagttctt gatggactg agagtaagag ctaagggcca    180
ggtgtgtgta ctccaaacag taatcccaac attttagagg gctgaggcag gcagatgtct    240
tgaccccatg agtttgtgac cagctgaac aacatcatga gactccatct ctacaataat    300
tacaaaaatt aatcaggcat tgtggtatgc cctgtatgcc cagatgctgg acaaggtgtc    360
a                                                                 361

```

```

<210> 166
<211> 427
<212> DNA
<213> Homo sapien

```

```

<400> 166
twagctgact catgtccctt acacccaact atcttctcca ggtggccagg catgatagaa    60
tctgatcctg acttagggga atatittctt ttacttccc atcttgatcc cctgcgggtg    120
agtttctcgg ttacgggtaa gaaaggagct caggccaaag taatgaacaa atccatccto    180

```

WO 01/90152

PCT/US01/16776

46

```

acagaagtag ggaataaag aacwtggacw tagccagcwg aacmcaaltg aaamcagaac 240
mcttaacatg gatracacmc merrazatar ktgcymcmc wtataataga aaccaaactt 300
gtatctaatt aaataittat ocaaygtcag ggcattagtg gtittgataa ataogctttg 360
gctaggattc ctgaggttag aatggaaraa caattgcamc gagggtaggg gacatgagtc 420
aktctaa 427

```

```

<210> 167
<211> 500
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1) ... (500)
<223> n = A,T,C or G

```

```

<400> 167
aacgtgcgat gctcccgccg gccatggccg cgggatagac tgactcatgt cccctaagat 60
agaggagaca cctgctaggt glaaggagaa gatggttagg tctacggagg ctccegggtg 120
ggagtagttc cctgctaagg gagggtagac tgttcaacct gttcctgcto cggcctccac 180
tatagcagat gcagccagga gtaggagaga gggaggttaag agtcagaago ttaigtgttt 240
tatgcgggga aacgcattat cgggggagcg ctagttatta ggggcaatx taghyartcw 300
agtlagcact caaagcggcg gagttttccc atatggttgg acctgcaggc ggcgcattta 360
gtgatlagca tglgagccoc agacacgcac agcagaagg acctaaactc agtccctgtg 420
ctgattactt mncatgaatt attgtattta tttacaaact ttgagttatg aggcatttta 480
ttaggcccat attacutgga 500

```

```

<210> 168
<211> 358
<212> DNA
<213> Homo sapien

```

```

<400> 168
ttcatgcctc ggtgactcwa gccgtgaac ccgaaccttt gggaggccga gggagcaga 60
tcacctgagg ttggagttt gagcccgcc tggccaaat ggtgacaacc cgtctctgt 120
aaaaatacaa aaattagcca agcatgttgg catgcacttg taatcccagc tactggggag 180
gctgagggcg gagaatcact tgagcccgag aggcagaggt tgcagttagg cagaggttga 240
gctcgtgcc cttgcactcca gcttgggcaa cagagttaag ctccactcca aaaaaaaa 300
aaaaaagaa tgatcagagc cacaataaca gaacaccttg agtcacagag cgtatgaa 358

```

```

<210> 169
<211> 1265
<212> DNA
<213> Homo sapien

```

```

<400> 169
ttctgtccac accaatctta gagctctgaa agaatltgic tttaaatata ttttaaatgt 60
aacatgtatt ttatggacca aattgaatt ttgaatatt ttttccaaa aaagtcagg 120
tgaatttccg caccatgagt tgggaatttc ttatccaga agnccgacag agcaatttca 180
tattttatit agattgtatc catactcgt ttccaaggag aatcctgca gtctccttaa 240
aggtagaaca aatcctttct atttttttt caccattgtg ggattggact ttaagaggtg 300
actctaaaaa accagagcac aaatatgtct cagtgtgatt aagcacggac ccatattatc 360
atatctactt aaaaaaatga ttctctgtgc accttttggc aacttctott ttcaatgtag 420
ggaataactt agtcaccccg aaaaacacaa aatataataa aacttctaga tgtgggcaya 480
argtttgggg gtggaatgt talgtgttta aattaaaccc tgtatcactg agaaagctgtt 540
gtatgggtca gagaatgtga atgttagaa gctgttccaa tcttcaagag cagaagcaca 600
ccacatgtct cagctatatt attatttatt ttttatgcac aaagtgaact atttctctg 660
tattaatttc caaagggttt taacctctat ttaaatgtct tgaataacag tgcattgaca 720

```

WO 01/90152

PCT/US01/16776

47

```

atgggttgat atttttcttt aaagaaan tataattatg aaagccaaga taacttgaag 780
cctgttttat tttaaaactt ttatgttct gtgttgatg ttgtttgttt gttgtttt- 840
attttgttg tttttaactt tgtttttgt ttgtttgt ttgtttgt cactactacat 900
gcagtttct taaccaatgt ctgtttgtgt aatgtattt aagtgtttaa ttatattgag 960
tgcatttcaa ctatgtcaat ggtttcttaa tatttattgt gttagagtag tggtaatttt 1020
tttatttcaa atatgtttaa agagataaca gtttgatag ttatcatgt ttatagcag 1080
aagttattta ttctatgtgc attcagcgg statttgggt gtttgcggg catgcagtaa 1140
atatttgtta cagttatgtg acagtattca gcaacgcctg atagcttctt tggccttatg 1200
tttaataaaa agacctgttt gggatgtaaa aaaaaaaa aaaaaaaa aaaaaaaa 1260
aaaaa 1265

```

<210> 170
 <211> 383
 <212> DNA
 <213> Homo sapien

```

<400> 170
tgttagtcga gcagtgatg gacgatattc ttcttattaa tgttggaatt gaacaaatga 60
tcctgtatcc tgactctgag ctaggagcgc ctgttcagtt aatgggactt ctctgtactc 120
taattgtacc agagaacatg ctggataaca ctatataaac cgaataaagt gaatttctaa 180
atttttttta caacaaatgt atgcagtgtt tcaacgcacc acttttgacc aatactctcg 240
agacaaatgt tgaataaggt aatagattg gatcaacaa aaacaaacaa atttgcctcg 300
ataattatca aacagcacag ctacttgctt taattttaga gttaactaca ttttgcctcg 360
aacatcaac tggctgactt aca 383

```

<210> 171
 <211> 383
 <212> DNA
 <213> Homo sapien

```

<400> 171
tgggcacott caatatgca agttaaaat aatgttgagt ttattatact ttgacctgt 60
ttagctcaac aggttgagg catgtaaga atgtggactt ctgaggactt ttcttttaa 120
aagaaacaaa tgaagtaaca tttaattac tcaaggacta ctcttggttg aagtttataa 180
ctatagatac ctacttttt gttillyctg ttgcacagtt caaaagagcc ttacgaatt 240
taacaggttaa aatgttgtaa gtatggagg tgaactgaa atttaaat atttcttaa 300
tactatagg aaagaggctg agcttagaatt ctlttggtt ttcatgtgtt ctgtgtcttt 360
atcatcaac tgcctgactt aca 383

```

<210> 172
 <211> 699
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)... (699)
 <223> n = A,T,C or G

```

<400> 172
tcgggtgatg cctctcagg cttgttgta gtgtacag agctgtctat gaagcgacag 60
cggctgccc bggacttca gaactcttc ctctacactt ttgttgctct tctgtatca 120
ggctgtcatg ctggcggcgg ctctggcaca ggcctcctgg aaagtttctc aggtaggca 180
gcactgtgtg tctgagcaca ggcactaat ggcctgtca tgtgtgtctg catggagcat 240
ggcagcagca tcaacgcct ctttgggtg tctgtctgc tgggtgtcaa cgcctgtctc 300
tcagcagcct tgcacggct gcagctcava gcgcctctct tctgtgtcaa attgtctatt 360
ggcctggcaa tgcctctgta ctatggcaga cgtctgtccc tgacaacttc caactgatt 420
cggacacctg tagattggg gcaacacaa gatccctctc ccaggcttc ctccctctcc 480

```

WO 01/90152

PCT/US01/16776

48

```

catcagcggc cctgtaacaa gtgccttggt aaaaaagctg gagaagtgg ggcagccagg 540
ttattctctg gaggtttggt gatgaagggt taacctatgg agatgtgaag tgtgggtttg 600
gttaaggaaa tgcctaccat ccccccccc oacccaagtt ntccagact aagaactae 660
ggtacatca atacctaygc ctgaggaggc atccccga 699

```

```

<210> 173
<211> 701
<212> DNA
<213> Homo sapien

```

```

<400> 173
tcgggtgatg cctcctcagg ccagatcaaa cttgggggtg aaaactgtgc aaagaaatca 60
atgtcggaga aagaattttg ccaagaaaa atgctaate agtaactaatt taataggtca 120
cattagcagt ggaagaagaa atgttgatat ttatgtlcag ctattttata atcaccagay 180
tgcttgatct catgtaagcc atctcgtatt cattagaact aagaacaatt tiattcgtcg 240
gaaagaactt ttcaatttat agcatcttaa ttgtcagga ttttaattt tgataaagaa 300
agctccactt ttggcaggag tagggggcag ggaagagga ggtcccatcc acaaggacag 360
agacaccagg gccagtaggg tagctgtggt ctggtacagt cacacaggac tgactttatgc 420
catgagaaga aacacacctc aactctcagt tgcctaatat cacacaggct catttcttgc 480
tcacgtttac tgcctatgt agatcaacag caggtgactc aggggaonag gctccatctc 540
catatgagct tccatagtca ccaggaaacg ggccttgaaa gtgtcctcca tgcagggaac 600
catgcctctt ccttctatgt ggaagagcaa gtcccttate gccagaagtc acactgcagg 660
gaagtgcant cctgtgtat gcttgaggag gcatcacccg a 701

```

```

<210> 174
<211> 700
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> {1}...{700}
<223> n = A,T,C or G

```

```

<400> 174
tcgggtgatg cctcctcagg ccactaaatc agagtcagg gtacagaca caggagacag 60
ggaagacat agatittaac cggccccctt caggagattc tgaggctcag ttacatttgi 120
tgacgtttga ccaggaggcag caaggctagt ggttaggggc acggtctcta aagctgcact 180
gcctggtatc gctccacagg tctgcacaga accagctgcy tggccttgag ctgctgcacc 240
gcagaaagcc cctgtggac ccagctctct cgtctgtaag atgaggacag gactctagga 300
aacctttccc ttggtttggc ctcaclitca caggctccca tcttgaaact tatctactct 360
ttcttgaaa cctgttaaaa gaaasagtg ctagectggg caacatggca aaacctgtc 420
tctacaaaa atcaaaaaat tagttgggtg tggtagcatg tgctgtagt cccagccact 480
tgaggaggtc tgagggtggg ggtacacttg agccggggag gtggaggttg cagtgcacaa 540
agatcatgcc actgcactcc agcctgagta atagagtaag actclgtctc aaaaacacaa 600
acacaaacag tgagtggtcc tctgtttcgg ggttggtagg ggcacacat ttatgcatct 660
ctcagatttg gacgtgcag cctgaggagg catcacccga 700

```

```

<210> 175
<211> 484
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> {1}...{484}
<223> n = A,T,C or G

```

WO 01/90152

PCT/US01/16776

49

<400> 173
 tataggggga attaggcccg agttgcacgn tccgggcccg catggcccg ggattcgggt 60
 gatgcctcct cagggttgto tgccacaagc tactttctct agctcagaaa gtgccccttg 120
 atgagggaaa atgtcctact gcaatggaa tttctcagtt caattttacc tccagtcct 180
 ccttctacac cagttatata attcattcca caagtattta ctgattacct gcttggtcca 240
 gggactattc tcaggctgaa gaaggtggga gggaggggcg gaacctgagg agccacctga 300
 gcaagcttta tattcaacc atggctggcc catctgagag catctcccca ctctcgccaa 360
 cctatcgagg catagccag ggaatccccc agggggccca ggttagatgc gtcccttttg 420
 ctgtcagtg atgcataca ccttagctgc ttactgtgtg ctggcctgag gaggcctcac 480
 caga 494

<210> 176
 <211> 432
 <212> DNA
 <213> Homo sapien

<400> 176
 tggggtgatg cctctcagg gctcaaggga tgagaaatga cttctttctg gagggacagt 60
 tcattgccc caggatgaaa atggatagg acccactgg aggcattgct gatattgttg 120
 gaacaaatgc aggtagcgga atggatctg gtccaggagt tatccaggat agattttcac 180
 ccccaatggg aagtcactgt tcaatccac tcttcaatgg ccatggggga cacatcatgc 240
 ctcccaacaa atgcctattt ggaagatgg gaggcagtt tatgaanaag caggggctaa 300
 gcaagcttta ccaaacccag agtcaggga tcttatccca gctgcaagga cagtggaagg 360
 atatggccac tgggtttctt aagaaggac agcttaatga agatgagatt agcctcagga 420
 ggcataccc ga 432

<210> 177
 <211> 788
 <212> DNA
 <213> Homo sapien

<400> 177
 tagcatgttg agccacagaa cagtagoatt tctgcaaat tctggtlaga atggtgacaa 60
 catgctggag ccaagtgcata acatgccttg gtccaaggga tggaaagtc cccgttagga 120
 tgcacatgac agtggaaaca cgtgcttga ggcctctggc tgcctcctac ccccaactcg 180
 ccccaactgac aagccttgc gctgcctct ccaggatgic tacaacattg gtggatttgg 240
 taatgttctc gtggccggag tggagacttg tcttccaaa cccgctatg tggtaacatt 300
 tgcctcagtc aacgttaca cgaagataa atctgtcgaa atgcacatg aagctttgag 360
 tgaagcttct cctggggaca atgtggctt caatgtccag aatgtgtctg tcaaggatgt 420
 tctgtgtgac aacgttgctg gtgacagcaa aaatgaccca ccaatggagc cagctggctt 480
 caatgctcag gtgatttatcc tgaacacccc aggcacata agtgcggct atgcctcgtt 540
 attggtatgc ccaacggctc acattgcctg caagtttgct gactggaagg aaaaagattg 600
 tgcctgttct ggtaaaaagc tggagatgg cctcaaatte ttgaagtctg gtgatgctgc 660
 cattgtgat atgttctctg gcaagcccat gtgtgttgag agcttctcag actatccacc 720
 ttgggtgcgc ttgtgtgttc gtgatattgag acagacagtt ggggtgggtg tctgggtcca 780
 acatgcata 788

<210> 178
 <211> 786
 <212> DNA
 <213> Homo sapien

<400> 178
 tagcatgttg agccacagaa cctgtgttct tgggagctct ggcagtggcg gattcatagg 60
 caattgggtt ccaatttgaa tgacacactt ggcctttatg gattcactag tttttaaaa 120
 atgtgtgttc gttctcttct attaaaggtt taatcagaca gatcagacag ctaattttg 180
 tatttaataa cagaaaagtt ggtacatttc ttcatgaat agcttgcatc ctgaagcagag 240
 agcctacaaa aggcacttgt tataaatgaa agttctggct ctgaggccca gtactctgga 300

WO 01/90152

PCT/US01/16776

50

```

gtttcagagc agccagtgat tgttccagtc agtgatgcoo agttatatay agcaggagta 360
cactgtgcac toctotaggt gtaagggfat gaaacttttg atcttaaaat togttacaca 420
tacacaccttt atatatatgt atgtatgtat gaaacacatg aattagtttg tcaaatatgt 480
gtgtgtttag kattttagct tagtgcacat aittccacat takttatua attgatckaa 540
gacactttot tgttgacacc tgaatatata atgttccagg gtgcaatgt taltocttta 600
gattgttaaa gtttaatlac tatgttttgt agtaaatata cttttataat gtatttgaga 660
cattctgtag tgtcttaggg ctcttaccag gtgggaaaga ttttaatttt ccagttgcta 720
atlgacacag atggcctcat tatatatitt gatttatagg agtttgtgtc tgggctcaac 780
atgcta 786

```

<210> 179
 <211> 796
 <212> DNA
 <213> Homo sapien

```

<400> 179
tagoutgttg agccagagca ctggttacaa gaccagacot gcttcotcca tutgtaaca 60
gcttttaaaa agccagagaa ccttttaaat actttggcaa ccttctttca caggcaaga 120
acacccccat cccccccttg ttggagtg agagtttggc ttgtgttctt tgcottgcoo 180
ggagtatact totaattcct gttgtctctg acaagotgaa taacagagta cccaccccca 240
cccaggccag gtttccactc atttattact ttatglltot gttccattgc tgggtccacag 300
aaataagttt tccrttgagag gaabtgatb atacccottt aatttctctc ttgtgttttt 360
ttttaatato attggtatgt gtttggccca gaggaaactg aatttccaca tcatottgao 420
tgccaatccc attaccatgc tttttttaaa aaagttaatt ttctctgcoo taacttgcca 480
gagtagccct tcclygctac tggcttaatt tagtcaacta gttctatagg ggcattaggc 540
atgagacctg aagccacagac tgtcttaccu caaaaggtga caagatctca aacttagcc 600
aaagggtcat gtcaggtttc aatgtatctc gottctgttc ctgtctacig ttctgggttt 660
tgtctctctc catccctagc aacagaattt cccagtctcc ctccctacct tccctgtttt 720
taattotaat ctatccgcaa tataactttt caaatgtttt aacccgtatc tccatgtgtc 780
tgggtccaac atgcta 786

```

<210> 180
 <211> 488
 <212> DNA
 <213> Homo sapien

```

<400> 180
ggatgtgtgt caaggcgatt aagttgggta acgcaagggt ttcccaagtc aagcagttgt 60
aaccagcagg cagtgaaatt gtaatacagc toactakagg gccaatggg cccagcgtcg 120
catgtctccc gacgcacatg ccggcgggata gcatgttgag cccagacacc tgcaggctat 180
ttggagagat tttaacgltt aacggcttga tggttttttt caggggagaa gacactgago 240
actcccaagg tgaggttgaa gatttctctc agatagccgg ataagaagac taggagggat 300
gcatagaaaa tgattagcat gcaaatctct acotgccatt toagaactgt gtgtcagccc 360
acattcagct gcttctgtg aactgaaaag agagaggtat tgagaacttt ctgatggccg 420
ctctaacatt gtaacacagt aactctgtgt tgtgtgggtg tgtgtgtgtg tctgggtcca 480
acatgcta 488

```

<210> 181
 <211> 317
 <212> DNA
 <213> Homo sapien

```

<400> 181
tagcatgttg agccagagca cggcgaagggt aontgatgag tgggggtgat gcaactgtga 60
aaaggaggaa cgtcatcccc catgatattg gggacccaga tgatgaacca tggctcccg 120
taactgcatc tttaatccat gatactgctg attggaagga cctgaaactg aagtttgtgc 180
tgaaggttta tgggactat tacctcagg gtgataaaaa cttoctgaag gacatgtggc 240
ctgtgtgtct agtaagggat gcacatgcag tggcagatgt gccaggggta tggttgtgt 300

```

WO 01/00152

PCT/US01/16776

51

```

ctgggctcaa catgcta                                317

<210> 182
<211> 507
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(507)
<223> n = A,T,C or G

<400> 182
tegcattgtt agcccgaca ctggctgtta gccaaatcct ctctcagctg ctccctgtgg      60
tttggtagct caggattaca gaggcatcct gtttcaggga acaaaaagat tttagctgac      120
agcagagagc accacatata ttagaatggc aaggactgcc acctccttca agaacaggag      180
tgaggggtgtt ggtgaatggg aatggaagcc tgcattccct gatgcatttg tgcctcttca      240
aatcctgtct tagtcttagg aaaggaagta aagtttcaag gacggttccg aactgcttiti      300
tgtgtctggg ctcaaatgac tatcccgctg ccatggcggc cgggagcatg gacgctgggg      360
cccaattcgc cctatagtga gtctatttac aattcaotgg ccgtctgttt acaacgtcgt      420
gactgggaaa accctggcgt taccacaatt aatgcocttg cagcacatcc ccttttccca      480
gtggcgtaa tancgsaaag gcccgca                                507

<210> 183
<211> 227
<212> DNA
<213> Homo sapien

<400> 183
gatttacgct gccncaactgt ggaggtagcc ctggagcaag gcaggcatgg atgcttctgc      60
aatccccaag tggagcctgg tatttcagcc aggaactga gcagagcccc ctctaattgt      120
agcaatgata agttattctc ttgtttcttc aaccttccaa tagccttgag ctccaggggg      180
agtgtcgtta atcattacg cctggtctcc acagtgttgc agcgtaa                                227

<210> 184
<211> 225
<212> DNA
<213> Homo sapien

<400> 184
ttaagctgca acactgtgga gcagattaac atcagacttt totatcaaca tgactgggg      60
tactaaaag acaacaaalc aatgcttca aagttctnag gaataatttc gatacttcaa      120
ctttataaaa cctgacaaaa ctatcaatca agcataaaga cagatgaaga acatttccag      180
atktttggca atcagatatt ttacotccao agtgttgag cgtaa                                225

<210> 185
<211> 597
<212> DNA
<213> Homo sapien

<400> 185
ggcgcagcgt cgcattgctcc cggccgcacat ggccgcggga ttctgttagg tottatcca      60
ctgggaccca taggttagtc agagtattta gatttagatt cctttctgct tcccagaatt      120
tgaagaaaaa ggagttaggt gatagactga agagataaga ttgctctctg aagcctgttc      180
aagatgtatg tgcctagacc ccaccaatgg ggcctgtggg tgaggtcctg ggcacttatt      240
tgaatgaatt gctgaagggg agcaatagc caaggaaggg gaacccatcc tggcacttgg      300
acaggggtca ccttatccag tgcctagtc ttcttttctg ctactgtgtt ttctctcata      360
tgttaggggc agttaagaag aagtcgcccg tcttctgcca gttttagaac atctaccagt      420

```


WO 01/90152

PCT/US01/16776

52

```

aagtggggaa gtttcacaaa ggcgcagctt tgttttgtt attttcaact tcagttagaa 480
gaggagagct gtgagatgaa tgttagltga gtgaaaga cgggttaagt tagtgatag 540
agaccctaao gaatcaatag tgcggcgccc ttgcaggtcg acaatatggg agagctc 597

```

```

<210> 186
<211> 597
<212> DNA
<213> Homo sapien

```

```

<400> 186
ggcccgaaat tgcagtgtcc cggcgcgaat ggcgcggga ttogttaggg tcttatcca 60
ctacctaaaa aatcccaaac atataactga actctcaaca ccaattgga ccaatccatc 120
accccgaggg cctacagatc ctctttgat acataagaaa atttcccaa actacctaac 180
tatatacttt tgcagatttt gttttacca attttgatgg cctttctgag cttgtcagtg 240
tgaacaccta ttacgaacga ttgatatta actgcacctc accgtccagg tctagctgge 300
aacatcaagt gcagtaaaata ttcatttaagt ttccacctac taaggtgctt aaacaccta 360
gggtgcacgt tcggtagcag atcttttgtt ttgttttat ttccataag ggtcctgttc 420
aaggtcaatc atacatgtag tctgagcagc tagtcaatat cgcataactt ggaggtgat 480
aatagagccc tctttgtctg ttaagaacct ctgtcccg cctgtcaag tggatagaga 540
ccctaacgaa tcaatagtgc ggcgcgttgc aggtcagaaa tatggggagag ctcccaa 597

```

```

<210> 187
<211> 324
<212> DNA
<213> Homo sapien

```

```

<400> 187
tcgttagggt ctctatccac ttgcaggtta aatccaatcc tgtgtatata ttatagtctt 60
ccatatgtag tggttcaaga gaactcagtt ccagaaagac tagccgagcc catccatgtc 120
ttccacttaa cctgtctttg ggttacacat cttaactttt ctgttcaagt ttctcttgtt 180
agttttatag atgagtattg ggaataatgc ctgaacctg acatgagatc tgggaacac 240
aaacttactc aataagaatt tctccatat ttttatgtg gaaaaatttc acatgcacag 300
aggagtggat agagacctta acga 324

```

```

<210> 188
<211> 178
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(178)
<223> n = A,T,C or G

```

```

<400> 188
ggcgggggat tgggggtgal acctctcat gccaaaatcc aacgtntaat ttccaaactt 60
gccttccaat ttacgcattt tcaatttgc ctcctctttt gttgagtcac acaaaacac 120
attgcccaga aacagtattt acctaacatg cacatactct taacctact catccctt 178

```

```

<210> 189
<211> 367
<212> DNA
<213> Homo sapien

```

```

<400> 189
tgacaccttg tccagcatct gacacagctc tggctcttgg aaaaatttgg ataatgaaa 60
atgaatttct ttgcaagtg gtataagctg agaataacg tatcaatat cctcatttta 120
agacacatcc agtgtccctg aaattagaat aggaacttaca ataagtgtgt tcaatttttc 180

```

WO 01/90152

PCT/US01/16776

53

```

aatagctgtt attcaattga tggtaggcoo taaaagtcaa agaatgaga gggcatgga 240
aaaaagctc acatcacctg atcattagaa aacttccatt caaacccura algagatacc 300
atctcatacc agtcagaatg gctattaita aaaaagtcata aataacaga tgcaggacaa 360
ggtgtca 367

```

```

<210> 190
<211> 369
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(369)
<223> n = A,T,C or G

```

```

<400> 190
gacacccctgt ccagcatctg acaacgctaa cggcctgagg agatctttat ttatttattt 60
agttttttct ctggctaggg agalggtgyc taaacattc atttcccat ttattcattt 120
aattgtttctt gcaagggcta tggatagagt attgtccagc actgctctgg aagctaggag 180
catggggagt aacaagatag gctacatctt gttccacag aacttccact ttagtctggg 240
aaccagatga tatctccaa tatataatg aattcaggta gttttaagta cgaagaagat 300
aagaagcag agtcatgatt tanaatgctg gaacagggg ctattgcttg agatattgaa 360
ggtgccaa 369

```

```

<210> 191
<211> 369
<212> DNA
<213> Homo sapien

```

```

<400> 191
tgacacccctg tccagcatct gacacgggaa aagaacctat tatcagagtg aacaggcaac 60
ctacagaattt ggaagaaatt ttgcacatct atccatctga caaagggcta atctccagaa 120
tctacaaaga acttatcaa atttacaaga acaaaacaaa caaacacac ctcaaaaagt 180
gggtgaagga tgtgaacaga caattctcaa aagaagccat ttatggggcc acaaacata 240
tgaacaaag ctcatcatca ctgctcacta gataastgca aatcaaacac acaatgagat 300
accatctcat tccagttaga atggcaatca ttaaaagtc aggaacacac agatgctgga 360
caaggtgbc 369

```

```

<210> 192
<211> 449
<212> DNA
<213> Homo sapien

```

```

<400> 192
tgacacccctg ccacttgaca ctccatcttt gcacagaaaa acttctttac agatttaatt 60
caagactggt ctagtgaag tctccagac attttttcat ttgttccata taactgggat 120
tttaaaatca tgtttcatca gtttgaatg atttggggctt ctactcaaca caattggatc 180
gactgttcta ctacacaca ggaacatgt tatctggcag ctgttggaaga aacactaacc 240
attgattttt ctgtgcttt taaggacttt gttccagata ctgttaaac caartctctt 300
ttaagagag agagtgtga tcttcatttg ttctacacag actgcaacc tagtaaatat 360
tctttattt tgcgtgtaaa aaattgccat caaaataaga tgattcatga tactgttatt 420
ccgtctgagt gtcaagtgcc caaggtca 449

```

```

<210> 193
<211> 372
<212> DNA
<213> Homo sapien

```

WO 01/90152

PCT/US01/16776

54

```

<400> 193
tgacgcttgg ccacttgaca ccagggatgt akcagtgna tataatcctg caattgtaca 60
tattggcaat ttcccctcaa acattctaga aagagacac caggattgct agccatcaa 120
agctgcaata aataactggt aattgcagta atcatttcag gccattcaa tccagtttgg 180
ctcagaggtg cctttggcty agagaagagg tgagatataa tgtgttttct tgcacattct 240
tggaaagata actccacaat agtctgagga ctagatacaa acotatttgc cattaaagca 300
ccagagcttg ttaattcccg taactgataag ttttggagat tagactccag tgtgtcaagt 360
ggccaaagct ca 372

```

```

<210> 194
<211> 309
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(309)
<223> n = A,T,C or G

```

```

<400> 194
tgacgcttgg ccacttgaca cttatgtaga atccatcgtg ggcgatgca agccctttat 60
ttaggcttag tgttttgggc accttcataa tcacantaga gacaaagccg acaagatctg 120
cagaacacat cagttctgan cactgcaatg gcagataac tttttgtgtt gtaacatttc 180
acattataaa aacaaacoto tgaantctca cgttacaada aaactgactg ctgtaaaata 240
tttagaaggg gtaaaggata caactataaa caaagtaact tacaactagt gtcaagtggc 300
caagcgtca 309

```

```

<210> 195
<211> 312
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(312)
<223> n = A,T,C or G

```

```

<400> 195
tgacgcttgg ccacttgaca cccatctctg cacttcatec tcccagcacc tgatgaagta 60
ggaetgcac tatcccccct tcccagatga ggggacccan gtacacatta ggaccoggat 120
gggagcacag atttgtccga tcccagatcc caagcactca gcgtcactcc agyacagcgg 180
ctttcagata agttcacaaa catgaatggc tccgacaacc ggaatcagtc cgtgctgagt 240
taaggcaatg gtgcacagga tgcagtgtn acctgtaatg gtccatccta agtgcacagt 300
ggccaaagct ca 312

```

```

<210> 196
<211> 288
<212> DNA
<213> Homo sapien

```

```

<400> 196
tgtatcgacg tagtggcttc ctaagccatg cagaactgtg actcaattaa acctttttcc 60
tttatgaatt acccaatctc gggtagtgtc ttatagtag tgtgagaatg gactaataca 120
agtacatttt acctagtaat aataataaac aaatatatta catthtttbtg tatttaactc 180
accataattt ttattgttat tgtagtgtac acctctact tattaaaga aataggcccg 240
aggcgggcag atcacgaggt caggagatgg agaacactac gtcgatac 288

```

```

<210> 197

```

WO 01/90152

PCT/US01/16776

55

```

<211> 283
<212> DNA
<213> Homo sapien

<400> 197
ttgggacccct tcaatctcat gacagggtat gtgataacca agaaggctac taagtatta 60
atgggtgggt aatgtataca gattaggtac actggacaga ggggtaatc atagccaagg 120
caggagagagc agaatggcae aacatttcat caccctactc aggatagcat gcagttaaaa 180
aootataagt agtttatitl tggaaatttc cacttaatat ttccagacty caggtaacta 240
aactgtggaa cacaagaaca tagataaggg gagaccaata cgtcgatac 289

<210> 198
<211> 288
<212> DNA
<213> Homo sapien

<400> 198
gtatcgacgt agtggctctcc caagcagtggt gaagaaaacy tgaaccaatt aaaaagtac 60
agatacccca aagaagaagg cttgagtaaa gattcnaagt gggtaaccaat ctcaatcttt 120
aaaattcagg ctgtcaaaag gatttgctat gaggttgcct tcaatgactt caggcaccgt 180
cggcaggaga ttgaagccct ggcattgtc aagatgaagg agctttgtgc catgtatggc 240
aagaagaacc caattggaggt ggaactctgg agaccctac gtcgatac 288

<210> 199
<211> 1027
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(1027)
<223> n = A,T,C or G

<400> 199
gctttttggg aaaaacncaa ntgggggaaa yggggntttn lngcaagggg ataaaagggg 60
aancccaagg ttcccccatt cagggaggtg taaaaagnc gccaggggat tgaanagga 120
ttcaataata gggggagtggt gccngaggt tgcaggttc cmgcccgcga tncocggg 180
ntttagtgac attacgacgt tggtaataa gtgggaccaa waaatatttg tgatgtgat 240
ttttagccag tgaacccatt gwaacagacc tcaattctty tgagatgrta gcaatcatca 300
gataaaagrt tgaagtytt tctgaacgtt aacagcatca ttaaatggag tggcatcacc 360
aatttcaccc ttgtttagcc gataccctcc ccttgaagcc attcaattac gtagccacac 420
gtcatcagag aggggatggc atggggattg atgatgatat cagggggtgat accctcacg 480
gtgaagacca tatcctcttg tctatcctga alacaaacag tacccttttg accatgtcga 540
ctagcaaat tgtctccaat ctgtgtwato octaacagag ogtaccctta ttttcaaaa 600
tttatatcct tctgtattga gatttccat aacctgatcc scaatgccg tctcgtwgt 660
tctgagaaaa gtgtcacagt ctctcttggt atagcgtcta ttggtgctct ccaattcato 720
tcaatttctc aggcagggg aactgttttg cctataataa cmtcatctcc tgatacmoga 780
aaccctkagg cctatcaaac catcacatc cagcgttckk walgtymcta aatccctail 840
gggycgcgct gcaggtcaac atatngaaa acccccaccc ccttnggagc ntaacttga 900
ttttccatat gtccntaaa tlanctngnc ttanctgpc cntaacctnt tccggtttta 960
attgttttcc ccccccctcc cncctttna accggaaacc ttaattttta accnggggtt 1020
ctctatcc 1027

<210> 200
<211> 207
<212> DNA
<213> Homo sapien

```

WO 01/90152

PCT/US01/16776

56

```

<400> 206
agtgaattta cgaagctggc catcttgaat ctagggcat gaagttgac caaagttcag 60
cacttgatga agcctgafcc ctctggitta toanaagaa taggatggga taagagaagt 120
ggacccttaa ataagctata aattatatgg tocttgtcta gcaggagacc actgcacagg 180
tatactacca gctgtgtaat gtaacta 207

<210> 201
<211> 209
<212> DNA
<213> Homo sapien

<400> 201
tgggcacatt caatctctat taaaagcaca aatactgaag aacacaccaa gactatcaat 60
gaggttacat ctggagctct cgtatatca ggaaaaaatg aaytgaacat tcacagaggt 120
ttactttttt gggaactcaa atgctagaaa agaaaagggc gccctcttct tctggcttcc 180
tggctcatatc cagcgtcgtg atgtcacta 209

<210> 202
<211> 349
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(349)
<223> n = A,T,C or G

<400> 202
ntacgtgga acaactgtga gcaactggtt ttatttcccg gcaggttato cagcaaacag 60
toactgaaca caaccgaagc cgtggtatgg taacogttca gactaatgt tccagtcgtc 120
tgccgggaccc cgaagagcgt cactgggtac agccagatt cagccggaag agaaagcgcc 180
gcagggaagag actgaactc cactccgtg gtagcagac ccattgtttc aactcgaagt 240
tcaaacggca ttgggttata taccatcagc tgaacttca acactctcc ttgaacacac 300
tggaaatcta lttcttgtt cctctcttct ccacagttgt gcacagtaa 349

<210> 203
<211> 241
<212> DNA
<213> Homo sapien

<400> 203
tgctctcttt gctttaccaa ccaaaagccc actgigaact atgaagtga tgacaaaatt 60
cagttttcaa cgaatatag tatagtttat ctgattcttt tgatctccag gacactttca 120
acaaactgta ccaacacac caacctaggg atttaggatt ctcaacagac agaaatttat 180
ttctctcttg agtttcaggc tctcttgga ctctgttca tcaatgggtg gtaaatggtt 240
a 241

<210> 204
<211> 249
<212> DNA
<213> Homo sapien

<400> 204
tagcaattta ccaaccatct gcaaaccewg acmwwnargx cywgewekys ggcgatttga 60
agtaactgta atgctctgat catgttagti acataagttt ggtcagttta caaaatttca 120
cagaactaaa tactcaatgc tatgtttca tgtctgtgtt tatgtgtgtg taatgtttca 180
attaagtgtt tttaaaaaaa agagatgatt tccaastaag aaagccgtgt tggttaaggca 240
agaggagc 248

```

WO 01/90152

PCT/US01/16776

57

```

<210> 205
<211> 505
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(505)
<223> n = A,T,C or G

<400> 205
taagctgaag aactgtggag caattoatac agtccctaa ttaaggaaca agtgattatg 60
ctacatttgc aaggttaagg taacggggcc gtttaacatg tgtactggg caggcggtgc 120
ctctaatact ggtgatgcta gagggtgatgt ttttggtaaa caggcggggt aagatttgc 180
gggttccttt tacttttttt aacctttcct tatgagcatg coctgttttg gtgacagt 240
gggttaataa tgacttgggt gttgattgta gatattgggo tgttaattgt cagttcagtg 300
ttttaatctg acgcaggctt atgcggagga gaattgtttc atgttaotta tactaacatt 360
agttcttcta taggggtgta gattgttoca attgggtgtg aggaattcag ttatatgttt 420
gggatttttt aggtagtggg tgttgancit gaacgctttc ttaattggtg gctgctttta 480
agcctactat ggggtgtaaa tggct 505

<210> 206
<211> 179
<212> DNA
<213> Homo sapien

<400> 206
tagactgaat catgtccctt accaaagccc atgtaaggag ctgagttcct aaagactgaa 60
gacagactat tctctggaga aaaaataaat ggaattgta cttaaaaaa aaaaaaatac 120
ggcggggcat ggtagcacac acctgtaata ccagctanta ggggacatga gtcagtcta 179

<210> 207
<211> 176
<212> DNA
<213> Homo sapien

<400> 207
agactgaact atgtccctta ccccaacttc tgctgtgtgt ongtgttctt aacaggtcac 60
agactggtac tggtcagtgg cotgggggtt ggggaactot attatatggg atacaaattt 120
aggagtggga attgacacga tttagtgact gatgggtat ggttggtaaa tggcta 176

<210> 208
<211> 196
<212> DNA
<213> Homo sapien

<400> 208
agactgaact atgtccctta ttttaacagg totctagtcg tctgaaaaaa aaaaatgctg 60
aacattgcat ataatata tgttaagaa taactgtaca tgactttatt gcctctgggt 120
agctgtttag catgaaggat gccagaagc ttaaggata tgggtgttaa atggctaggg 180
gacatgagtc agtcta 196

<210> 209
<211> 345
<212> DNA
<213> Homo sapien

```

WO 01/90152

PCT/US01/16776

58

```

<220>
<221> misc_feature
<222> (1)...(345)
<223> D = A,T,C or G

<400> 209
gacgctggc cactggacc cttttattt ttacggatc ttaagtcatt tangtnactt    60
tgtaaattt tctgtgccc cstaagaaat gatagcttta aaattatgc tggggtagaa    120
aagaagatac ttctagcttt agaatcttta ggtatagcca gaattcttgg ggggaggggt    180
gatttagagc aaatttctta ttctccttgc ctcatctgla acatggggat aataatagaa    240
ctggcttgac aaggttggaa ttatatttac atggttaata catgtataat gtttgaatg    300
gtgcacaata totaggaggt acttgggcat ggtgtgtaa tggct                345

<210> 210
<211> 178
<212> DNA
<213> Homo sapien

<400> 210
gacgctggc cactggacc tagagtggg ttggccaac ttittotata aaggaccaga    60
gagtaaatat ttccagcttt gtgggttgg cagctctct tgcacactac caagcttggc    120
atttagatag agaatcagc catagacagg acagaaatga atgggtgga aatggcta    178

<210> 211
<211> 454
<212> DNA
<213> Homo sapien

<400> 211
tgggcacott caataictat ccagcgcac taaattcgct tttttcttga ttaaaaattt    60
caccacttgc tgtttttgct catgtatacc aagtagcagt ggtgtgagcc catgcttgtt    120
tttgatctcg atatacgac cgtataagag cagtgccttg gcaatttatt tatcttcatt    180
gtagaacaga taggttagag tggatctcnc atactcaat ggaattcttg gctcaatgac    240
atgtctcagc aacattaaag caacttcato ttctctggat tgtacggcct ttgtcagaga    300
tgtctctctt ttgtgttcac ggcatttaag ttgacatcgt ctgtccagca cgaattttac    360
tacttttgaa ttcccatlgt cagagggcag atgtagagca gtctctcttt gcttgnocct    420
cttgittaca tcagtgtccc tgagcstaac gga                454

<210> 212
<211> 337
<212> DNA
<213> Homo sapien

<400> 212
tcogtatgc caccagaaa acctacttga gttacttatt aacatcaagy ctggaacctc    60
tttccctcag tctatcttga ttcatagca catggttatt actgatcgaa ttgaaaacat    120
tgatcacctg ggtttcttta ttatcagact gtgtcatgac aaggaaactt acaaatgca    180
aagcagagaa actatttaag gtattcagaa aogtgaagcc agcaattgtt togcnaatcg    240
gcattttgaa aacaaatttg ccgtgggaa ttaatttgt tcttgacag tcagaaaaa    300
cattattgag gaaatttaat atcacagcat aacggaa                337

<210> 213
<211> 715
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature

```

WO 01/90152

PCT/US01/16776

59

```

<222> (1)...(715)
<223> n = A,T,C or G

<400> 213
tcgggtgatg cctcctcagg catcttccat ccatctcttc aagattagct gtcccaaatg    60
tttttccctc tcttctttac tgataaatgt ggactccttc ttgacactga tgacagcttt    120
agttctcctc ttgtcacctt gcagacttta aacataaaaa tactcattgg ttttaaaagg    180
aaaaaagtat acattagcac tattagctt ggcttgaaa catcttctat cttttattaa    240
atgtcgggta gctgaacaga attcalttta caatgcacag tgagaaaagg agggagctat    300
atgcatttga gaatgcacag attgtcaaat aaacatttta atgtctttct taaagtggagc    360
aacatcaaaa atacatttag atattgaaa gtgttttgc ttgtgtacta cttaattaggc    420
aagcaccttg tatagttcct cttotaaaat tgaagttagat tttaaaaaac catgttaatt    480
aatgagctgc tcagttcaga ttttagggag attttaacag gattttggtt ttgtctaaat    540
tttgtcaatt ttttagttta atctgtataa ttttataaat gtcaaaclyt atttagtccg    600
ttttcagctc gctatgaag aaatacccan gacagggtta tttataaagg gaaagangtt    660
aatttgactc ccagttcaca ggctlgaggc ngnatonccc gaaatcccta ttgug      715

<210> 214
<211> 345
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(345)
<223> n = A,T,C or G

<400> 214
ggtaangmc atactcgggt gctccggccg ccgggctcgg gggattcggg tgaigctcc    60
tcaggccccc ttgggctctgc ttttccaaa tggcagctcc tctggacatg ccatctcctc    120
tcacccctgc ctgaltcttc atagtltggg tgtcctgttt tttctggtgc talttctctga    180
ctgctgttca gctgcactgc tctgcacaa cctgcctttt taaatgcttc accattcctt    240
catttgttcc ttaaatatgg gaaagtgaag tgcacactga ggcggggcac agtggctcac    300
gctgttaac ccagcacctt gggagcctga ggaggcatca cccga      345

<210> 215
<211> 429
<212> DNA
<213> Homo sapien

<400> 215
ggtgatgctt cctcaggcga agctcaggga gacacagaac ctcccgctga gcagaagggc    60
aaaagctcgc ttgatcttga ttttcagtac gaatacagac cgtgaagcg gggcctcagc    120
atcctctcga ccttttgggt ttttaagcagg aggtgtcaga aaagltacaa nagggtaac    180
tgctcttgga oggccaagcg ttoatagcga cgtgcctttt tgatccttg atgtcggctc    240
ttcctatcat tgtgaagcug aattcaccaa gcttgggatt gttcacccac taatagggaa    300
cyl'gagctgg gtltagacgg tcttgagaca ggttggtttt accctactga tgatgtgttg    360
ttgcaatggg aatcctgctc agtacagag gaaccgcagg ttcaaacatt tgggtgatgt    420
gcttgcctt      429

<210> 216
<211> 593
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(593)

```


WO 01/90152

PCT/US01/16776

60

```

<223> n = A,T,C or G

<400> 216
tgacacatat gtcnccatc tgttcacagt ttccacaaat agccagccct tggccacctc 60
tctgtccctga ggtatacaag taatacagga ggtgtatacc ttctcttctc ttcccccaca 120
aagagaaacat gcaggctctg gaagctgtct taggagccct tgggtctaga atttccagagt 180
cttgggtacc ttggatgttg tctggaagga gaaacatttg ctctggataa ggagtcagag 240
cggaggaggg tcacagagcc ctccagctca gccctctgtc cttagtctaa asgcagcctt 300
ggatggagaa gaaggttaag taacatccgt agcgtacac aggtagaag tgctggggagt 360
cagatctga caggtctcag gagttagtacc tcaatcaatg agggcaaatc aactgaagaa 420
agaaagacaa tttaagaaat gctttagggg aaggatcag gctatcaatg agatcttct 480
aggaagatta ttgtttanaa ttatgaagg antagggcag ggaacgggac agaagtanaa 540
gaaacatttg octatancce ttgtcttga cccagctgtt ggacaggtg tca 593

<210> 217
<211> 335
<212> DNA
<213> Homo sapien

<400> 217
tgacaccttg tccagcatct gaagtgaaga tgagcagctc agaggaggtg tectggattt 60
ctgggtcttg tgggtctcgt ggcattgaat tottctgtga agtggatgaa gactacatcc 120
aggaacaaat taatcttact ggcctcaatg agcaggtccc tcaactatga caagctctag 180
aaatgatctt ggaacctggag cctgtatgag aactgaaga caactccacac naggatgac 240
tgatttgaga ggcagccgag atgcttattg gattgatcca cgcncgtac atocttaca 300
aocgtggcat cgcncagatg ctggacaagg tgtca 335

<210> 218
<211> 248
<212> DNA
<213> Homo sapien

<400> 218
taagtactgg tcttgaaggt cttaggtaga gaaaaaatgt gaatatttaa toaagagcta 60
tgtatgaat gggactgtaa gtacagaggy aagpytgycc ctatctgcca gaagttggta 120
gatggctccc cytatgtgaa tgttgttga ctgcccagaa ttgctgcat tactgaatt 180
cgttgaatt agtgcacatt ctaagttgt tcatctaga ttatggttcc atgtttctag 240
tactttta 248

<210> 219
<211> 530
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)... (530)
<223> n = A,T,C or G

<400> 219
tgacgcttgg ccacttgaca caagtayggg ataaggacaa agacccatna ggtggcctgt 60
cagccttttg ttaactgtgc ttccctgtca ccaaggcccc ctctgtagg gttgtctgtg 120
ctctgtggac atkggtgcct tttaacacac accattctct ttctgttca cagaagtcct 180
gaggcggag cacaaggaac taacttgtca gatgagata atgatgtctg gccaaotcac 240
cccccacact totcactagt tatangaaga gccangccta naacpttcta tctgmcccc 300
tggccctatg acotcatccc tgttccatgc octattctga tttctgtga acttgggagc 360
agcctggttt ntctctctca ctccagccto tctccatacc atggtanggg ggtgtgttgc 420
caoncaaaag gtacaggtgtg tctgggaat octnananct gcncggagtt tconangcat 480

```

WO 01/90152

PCT/US01/16776

61

```

tcttaaaaaac cttcttgcoct aatcanatng tgtccagtg gcaacnctcn      530

<210> 220
<211> 531
<212> DNA
<213> Homo sapien

<400> 220
tgacgcttgg ccacttgaca ctaaatagca tcttctaaag gcttgatlea gagtltgga      60
aaattctccc agtgcacagg abtgtaagga acagggtgac tctgtgtctc actttacotg      120
ctgtgtttct gctggaagag gagggaagag gaatggctga tttttacctc atgtctccca      180
gtttttcata tcttttttag atctcttctc ctgacacctg ttccttttg gctttcttct      240
tctgtctcag agagcaggtc tctttaaac tgaagagaga gaatgagcaa atgattaaag      300
aaaacacact tctgaggccc agagatcaaa tattagctaa atactaaac gottgcctgc      360
tgtgttcaat ttctcctctc ttacatgct ctatcctctc atccccacac taitcatatg      420
gctttttatc gccaaagtat ccggcctctc atcaaccttc tccctagcc tactggggga      480
tatcaatctg ggtctgtctc tgggtgtatt gtgtcaagt gcaagcgtc a      531

<210> 221
<211> 530
<212> DNA
<213> Homo sapien

<400> 221
attgcgctt ggccacttga caacggcctg cctgcaatc tgggpcagg gcttccactg      60
ctttcctccc accagctgac actgcacaca gagatcaga atgtaccaa ccaagactgt      120
tggtcctcag cctctctgag gagaaagaga agaagcctgg agtccagaag agaagctaga      180
tgggtcagg ccttgccagc cagctcccc accgttgcca ataaagtct gcatgctta      240
acaaaggagg caactcctga gaaacacatt gtaggcaat tggcgctgtg ttcacagag      300
catatttaca caaacctcga tagtgcagcc tactatccac tattgtctct acgtgcaaa      360
cctgaacaga atgggaactgt actgaatact ggaagcagct ggtgatgga cttatttctg      420
tatctaaaca caggagaaggt acagtacaaa taigtatca taaacttaca gggacgcaca      480
tcttatatgc agtctgttgt gacaaaaatg tgtcaagtg ccaagctca      530

<210> 222
<211> 578
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(578)
<223> n = A,T,C or G

<400> 222
tgtatcgagc tagtggctc cgggtctacta ggcgcttctg tcttggtagt aactggttca      60
ctgaanaggg catctccctc cccgctctgc cctgaagcag ggggaggact tggccacgcc      120
aaggcagttg tatgagtttt agctgggga cttttagacc tttagagcca cctctctcag      180
gagctctccc cgaataagga agcaggggta aggtatcttt cttccccag aacccagaa      240
caccacccca cccccctat tctggcgcc catatacatc agaacgaaac aaaaataaca      300
ataaaccaa acccaaaaaa aaagagagag gggcaatgta tatgtctgtc catctgtttg      360
ctttagcctc tcaactccta agggcaggg aactgtctt ccaaatgctc tgtgcagcgc      420
cgactgcggg aagtatcaga ggaagagca gagtcagcag aagttgaac gtcggcccg      480
cggctctctg gggctgtgtg tgaactcga gacgcttctc gctttttgct ttagatttcc      540
gtttgtcttl tggagtggga naccactacn tcnataca      578

<210> 223
<211> 578

```

WO 01/90152

PCT/US01/16776

62

```

<212> DNA
<213> Homo sapien

<400> 223
tgtatcgacg tagtgggtct ctcttgcaa ggaatggctg gtgaatggit tccctgaatt    60
atggacttac cctaaacata tcttatcact attaacagtt gcaaatatatt agaattggtt    120
gtccactgtt catttgattc ctagaaggtt agtcttagat atgttacttt aacctgtatg    180
ctgtagtgtt ttgaatgcat tttttgtttg catctttggt tgcccacact gtcaattata    240
gctgcttagg tctggactgt cctggataaa gctgttataa tattacacag tccagccata    300
ttacaagcta attaagtcac ctaaatgctt cctgtttttg ccagacttgt tatgtcaatc    360
ctcaattctt aggttcattt tgggtgacct aaactcttgg gtgtgacttt cttagcatcc    420
tgtaacatcc atttccaaagc aggcacact tcaattataa ctttcagaa gtctatttgt    480
gaagccttct cttcacccag cggagcaact tgattttctt caacttccct catcagagcc    540
acaagagtat gggatctgga gaccaactac tcatataa
578

<210> 224
<211> 345
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(345)
<223> n = A,T,C or G

<400> 224
tgtatcgacg tantgggtct ccaaggtgct gggattgcag gcatgagcca ccaclccacg    60
gtggactctt tcttttatcc ttacttcatt aggtttctgt tattcaagaa gtgtagtgtt    120
aaaagtcttt tcaattctaca tggttaataa atgataagctt gggaaataaa tagaaatttt    180
ttctttcttc tttaggttga ataaagaaac agaaaaataa gaacatactg aaataatctt    240
aagttccaac cttagagagc ctgcagaaga aatgaagaaa gtgatgatga tttagatttt    300
gatattgatt tggagagcac aggagagagc caotangtct ataca
345

<210> 225
<211> 347
<212> DNA
<213> Homo sapien

<400> 225
tgtatcgacg tagtgggtct ccaactgagg tatgtgtgac actagacacac aaagccttcc    60
aacaggggcg caggccacag cagtttaagc ggaactcigt tctaaataaa ttccacactt    120
ctctaatgtat tctttcttaa aactgataca gttgtgaagc ctgtgtcttt tcccaactcc    180
cctttgacaa cagccttcaa ctaacacaa gaaagggcatg tctgacactc ttccctgagt    240
tgactctgat aagtgttct gatgtctaaa gagctccaga acaccaaagg gacattccag    300
aatgtctgtg tataacagac tccaatggag accactacgt cgataca
347

<210> 226
<211> 281
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(281)
<223> n = A,T,C or G

<400> 226
agggngggga ntgtatcgac gtagtgggtc ccccaacgtc tgtcattcag tctgcaggtg    60

```

WO 01/90152

PCT/US01/16776

63

```

tcagtgtttt ggacaatgag gcacaaattgt caattattga ctccctagct ctataatgotg 120
aaattaaatc ttgtcatgac aagtctggaa ttctctatga ggtttttcaa agtatttttg 180
atcaatactc caacaaatca gaagccaga aagaggatuc ttcaatatt gcgaabccac 240
gagtggattt acacacotca ggagacacat acgtcgatac a 281

```

```

<210> 227
<211> 3546
<212> DNA
<213> Homo sapien

```

```

<400> 227
gggaaacact tctctccagc cttgtaaagg ttggagccct ctcaagtata tgcgcagaa 60
ttttttctct ggtttctcag aggtataggg agtcgccttt aaaaaaggca agctctggaa 120
actctgcaca gtagaatggc caaagcttgg agttgagtg ccctctgaag ggtcaatgaa 180
cctacaactt gttcaagctg tgtgggggtt tgttactgaa actcccgccc tccctgatca 240
gtttccctac attgataaat ggcctgagtt ggtcaggagc accccttcog tggctccact 300
ctgcacccat tcataatttt acotccaagg tctctctgag ccagaccgtg ttttgcctc 360
gacccctcag cggttcgctt cgcctgtatc tgcctctctc tgaagaagag gagagtcctc 420
ctcaacccag cccacccgct taaaacccag ctactccctt agggctatcc catgtctctt 480
cggctatgtc cccgttaggc tcatcaccca ttgcctcttg gttgcaaccc tgggtggagg 540
aagttagccc tctactacca ctgagagagg caaaagtccc tctgggtgat ggtgctcca 600
ccccctctct ggtttatgac cctctctctt acttctgaet tgtataattg gaaaccccat 660
aatccctccc ttctctgaaa gcccaagctt ttgaactcac tgaaggagtc tgaactctgg 720
accactctgc cccctctgga tgaactctca cagctctctt tgaactcttt cccctctgga 780
gagagggaaa gtatccaaag agagggccaa agtctcaacc tcaactcaac caataggccc 840
gagagggagc ctgagagaaat agtctatgaa gucccaattg ggaactaatt gggacccaaa 900
ttttctcaag ggagggagaa ctcttgacga ttctcaaccc tatctctctg tgggtattca 960
gggagctgct cagaacccca taaacttgc taaggccact gaagtctctc aggggcatga 1020
tgaactacca ggaagtgttt tagagccact ccaggaggct tatcagattt acaccccttt 1080
cccagatgga gcccccgaaa atagccatgc tcttaatttg gaattttgtg ctcaaggcagc 1140
ccagatagat aaaaaggaaac tcaaaaaact agagggaatt tgcctggatg aataccagtc 1200
agcttttaga gtagccctaa aggttttttg aagctcaaga ggttgaaaaa caaaaacag 1260
cagctcaggg agctgaaaaa agccactgat aaagctctct gggatctcag agtttactgt 1320
tagatcaagg tcaattgact tccctctcca catggtgttt aaatccagct acactacttc 1380
ctgactcaaa ctccactatt cctgtctcag actgctcagg acgttctgaa actactgaaa 1440
ctggcggacc tgaacttcaa aattgtcccc taggaaggtt ggaagccacc atgtctcagg 1500
acagtcagcc ctctctcagc aagggaactac gaagggccgg agccacaaag gctgaattga 1560
cagctgtgtt gtggctcag cgtttactac caaacctctc agccacaaag gctgaattga 1620
tcgctctcac tccagctctc cgtttactac aggtatkaa cgttaacact gacagcaggt 1680
acgctcttgc tactgtgcat gtactgtgag ccatcaccac ggaagctggg ctactcactt 1740
cagcaggtgg ctgtaatcaa ctgtaaggga catcaaaagg aaaaacaggg tgttgcctg 1800
ggtaacaaag aagctgatto agcagctcaa gatgcaggtt gactttcagt cagcctctca 1860
aacttgcctc cccagctctc ctctccacag ccagatctgc ctgcaaatcc cgcatactca 1920
acagaagaag aaaaatggcc tcaagactca gagccaataa aactcaggaa ggttgggtga 1980
ttctctcaga ctctagaatc tctatacccc gaactcttgg gaaaacttta atcagtcacc 2040
taagctcac caacattta ggaggagcaa agctaacctca gctctccgg agccttttta 2100
agatcccca tctcaagga ctcaagagtc agcagctctc cgggtgcaca aactgagccc 2160
aggttaaatg caaaaaaggt cctcaaccca gccccaggca cgtctcaca gaaactcac 2220
caggagaaa gtgggaattt gactttacag aaqtaaaacc aacccggctt ggttaaaatt 2280
acottctagt actgttagc acottctctc gatggactga agcatttgc tccaaaaagc 2340
aaaotgtcaa tatggtagtt aagtttttac tcaatgaant catccctcga catgggtgc 2400
ctgtttgcaa taggtctgta taaggagcc gcttctccct tctctatagt ttagtcaagt 2460
agtaaggctt taacattca atggaagctc cattgtgctt ctgcaccca gagctctggg 2520
caagtgaac gcatgaactg caccctaaaa aacactctta caagattaat cttagaacc 2580
gggttaaat gtgaagtct ccttctttta gccctactta gactaaggtg cacccttac 2640
tgggtgggtt tctactctt tgaactcatg tatggaggg tgcctcctat ctgtcctaag 2700
ctagagagtg cccattggc aaaaatata caacttaatt tattacagta cctacagct 2760
cccaacagg tacaagatat cctctgcaa cttgttgag gaacccatcc caatcaatt 2820

```

WO 01/90152

PCT/US01/16776

64

```

cctgaacaga cagggccctg coattcattc cggccaggtg acctgtgttt tgltaaaaag 2880
ttccagagag aaggactccc tctgtcttgg aagagacotc acaccgtcat caggatgcca 2940
acggctctga aggtggatgg cattctctgg tggattcatt antccogcat caaaaaggcc 3000
aacagagccc aactagaaac atgggtcccc agggctgggt caggcccttt aaaaactgac 3060
ctaagttggg tgaagccatt agttaattc tttttotlaa ttttgtaaaa caatgcatag 3120
cttctgtcaa acttatgtat cttaagactc aatatauucc ctttgttata actgaggaaat 3180
caatgatttg attcccccac aaacacacag ggggaatgta gtgtccaaac tggtttttac 3240
taacccctgt tttagactct cctttccttt taatcactoa gcttgtilcc acctgaattg 3300
actctccctt agctaagagc gcagatgga ctccatcttg gctcttloac tggcagccgc 3360
tctctcaagg acttaacttg tgcagctgta ctccagacc atccaaagat gcaattacat 3420
gataagatac tglggcagc tatatccgca gtccctggga atttgtccaa ttgatccag 3480
ccctctacac cttcagcacc caccaccccg atcagtcagc agccatcagc accgagccaa 3540
ggcctccacc aagcacaagc attctgactc actgaagact tggatgatac ttagtatttt 3600
tagcagttaa gttttttttt cttttctttt cttttttttt cgtgac 3646

```

<210> 228
 <211> 419
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1) ... (419)
 <223> n = A,T,C or G

```

<400> 228
taagagggtc caagatctca gcacagccgt caatgcagaa cacagaacgt agcctggtaa 60
gtgtgttaag agtgggaatt ttgggattac aggttaagcc acctaacctt agctgggggt 120
tggtagccgt cccagatggc ttccagacga aagtgtcttg agatgagttt ttaagaatga 180
ataagatagc acacaagtga ggaactgactt ggcagtgggt aatggtgggt ggcacacac 240
ttcgaatata tggaaactgc acgtacagga atgaagaatg agactgtgtg gtgttttaatg 300
agctgcaaat actaatttta tctgaaagt ttggaagatg taactacaaa gtatttttta 360
glaaggaaat aacctacat ttcagggtta ttgtltgttt anattattgaa ggtgcccac 419

```

<210> 229
 <211> 148
 <212> DNA
 <213> Homo sapien

```

<400> 229
aagagggtac ctgtatgtag ccaatggtgc aatgagagac tgattactac ctgctggaga 60
ttgtttaagt gagttaatat attaaggata aaggagacca ggttttttga ctgttggaga 120
aggaattacc agatattgaa ggtcccaa 148

```

<210> 230
 <211> 257
 <212> DNA
 <213> Homo sapien

```

<400> 230
taagagggtc cmaaaaaaaa aaattgagac gaatgagtaa gacctactat ttgatagtac 60
aacagggtga ctatagtcac tgataactta attatcatt taactatagc tgaatttga 120
ttgtttgtaa ctgaaggat aaatgcttga gaggatggat accccattct ccatgtgta 180
cttattccac attacatgcc tqtatcaag catctcatat accctataaa tatgtacac 240
tactatgac cctotta 257

```

<210> 231
 <211> 260

WO 01/90152

PCT/US01/16776

65

```

<212> DNA
<213> Homo sapien

<400> 231
taagagggtta cgggtatttg ctgatgggat ttttttttct ttctttttct ttggaaacaa    60
aaatgaagac cagaacaaaa ttattgaaca aaagacaggg actaatcttg gagaatgaa    120
gtccctccac ctgactgcca ttctattcta ttgaccttc cagttaaggt taggagaata    180
gggggtggag gggattcaatc tgatacaggt atatttaag caactctaca tgtgtgccag    240
aagtcacatg taccctctta

<210> 232
<211> 596
<212> DNA
<213> Homo sapien

<220>
<221> misc feature
<222> (1)...(596)
<223> n = A,T,C or G

<400> 232
tgctctcttt gctttaccaa ccaaaaatta gaacataat gagatgtcac ctctaccttg    60
gtgggattaa cattatttaa aaatcagaa gtattgacaa ggaatggaag aaattagaac    120
atctgtgcac tgttgggtgg aatgtaaaaa aggtgtggcc actatgggtc acagcatgaa    180
ggttctctaa aaaaaatttt tttaaatcta ctctatgac gatcttgagg ttgtttatgc    240
aaagaaactg aatcaggat ttgtgggaaa tattcaatt ccaactcca ttctgtcttt    300
attcataata ctcaagagat ggaacaaacc taatgtcca tccgggatg aatggataaa    360
cacagtgtgg tatatgcata caatggaata ttatttagtc tttaaaaga aaattctat    420
catatactac aaottanatr aacottgagg aacaaatgct nagtreaata agccacggaa    480
ggaagaaatc tgcattatct cttatatga agtatctaaa gtgtcaaac tcttanagca    540
naaagtaaaa atgggtgtgt gccanacagt tggttagcgn agaaganaaa cctant      596

<210> 233
<211> 96
<212> DNA
<213> Homo sapien

<400> 233
tctctgaag acctttcgcg actcttaage tctgtgttg taaggcaaga ggaacgttg    60
taaggcaaga ggaacgttg taaggcaaga ggaaga      96

<210> 234
<211> 313
<212> DNA
<213> Homo sapien

<400> 234
tgtaagtcca gcagtgtgat gataaaactt gaatggatca atagttgctt cttatggatg    60
agcaaaagaa gtatgttctt gtatggaaat ctgctcctgg caaaaatgct gtgaacgttg    120
ttgaaaagac aacaaagat ttagegtagt acataaattt agaataglac ataaacttag    180
aatagtaaat aaacttagta cataaataat gcaagaaaga ggggaagga ttgagagaat    240
tgacttaaat ttgaaagag tatatacgtt aggttagatg ctctcaaca gcatcaact    300
gtctgaccta caa

<210> 235
<211> 550
<212> DNA
<213> Homo sapien

```

WO 01/90152

PCT/US01/16776

66

```

<400> 235
aacgaggaca gatcotta aaagatgttg agtgaaaaaa gtgaaaaa agataatctc 60
caagtcocag tagcattatt taacattttt taataaatat actgataaaa attttgtaca 120
tttcccaaaa atacattatg aagcacagca gcatgaatgc ctatgggrtt gaggataggg 180
gttgggagta gggatgggga taagggggga aataaaaaaa agagagggagt ctacacatt 240
tcataaccca agagatctaa ttatttcaac tatttgtacc wgaagtccag aaagagtga 300
ggcagaaggg ggaagaaggg gcgaagaaac gtttttggga gagggtccc aaagagaga 360
tttttgggat gtggcgctac atcgtttttt ccaggatgcc tttagctctg cccctatttt 420
ttctctctac taatattaga ttaaaccttt gaagacagc gtctgtggtt tctctctctc 480
agcttccctt ccgtgtcttg caacacagtag ctgttttaca agggttgaac tgactgaagt 540
gagattattt

<210> 236
<211> 325
<212> DNA
<213> Homo sapien

<400> 236
tagaotgatc catgtccctt accagagtag ctagaattaa tagcaccaagc ctctacaccc 60
aggaactcac tatgaatac ataactggaa ttatttcagc cttaaaaaat ttggaggga 120
attttggcat atgttaaac atggatgaac cttagagact ttatgatag taagaagaga 180
cagtcataaa aggaataata ttgatgatt ccacttatat gaggtaacct gagtagtcaa 240
ttctatagaa acataaaaaa gattgggttt tgcacgggct tttagagaaa agggatgac 300
aagttagagg acatagatca gtcta

<210> 237
<211> 373
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1) ... (373)
<223> n = A,T,C or G

<400> 237
tagactgact catgtccctt atctactcaa catttccact tgaagtctga taggcatctc 60
agaactatct tgttcccaag caactctctt atttcttctc atctagtctt ttatttcttg 120
tgctgtctta ccatctctaa agagtggca aaatccacca agttgctgaa acagaaactct 180
aagaaatctc ctgtattctt ctcttctcca tctacttccc ttctaatcca tttagaata 240
atctgtttca gaaaaccaaa cactctatgt tctactctat aagggggagt tgacaaatga 300
gaacacacag acaacgggag ggaacatca caccaccagg cccgtcaggg agtangggac 360
atgagtcaagt cta

<210> 238
<211> 492
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1) ... (492)
<223> n = A,T,C or G

<400> 238
tagactgact catgtccctt ataactctcc caggcatcag aaagcatctc aaactgggac 60
tgacacacat gcaagagttt caggttaagt acaaaagggg tcttaagaa ttgtccctca 120

```

WO 01/90152

PCT/US01/16776

67

```

atatacagat gattagaaga agtggacaga gatacccaag ttaaacatet gcaagataaa 180
aaaaatatgg caattgtgaa cacacactac aggaggaaaa taaggaaact aatagcatat 240
tgtgtctatfa tgatgatgaa gaacutctct anagaaaaa ataaccaag aaacaaagaa 300
aattctctgn aatgttitaal gctatagaag aaattacaa aaacatetat tcaatgaatt 360
cagaaaegt agcagggtcan aggaanacaa atcaaaagac agataatcc cattttagat 420
tgtogagtta actanaacag aaagaatacc actggaaatt gaattctac gtangggaca 480
tgantcunte ta 492

<210> 239
<211> 482
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(482)
<223> n = A,T,C or G

<400> 239
tggaaagtat ttaatgatgg gcaacttgct gtttacttcc taatatccc atcatcttct 60
gtattttttt aaataacttt ttlttgatlt tttaesgtaa ccttattctg agaggtaaca 120
tggattcaact acctctnago oattagggaga ctctatgtta aaacaaaggy aatgtttact 180
agatctctat ttgataaala ggtatggata atcatcactt ttctgcbota atggaaaagt 240
actanaaaca tggaaacata atcttatagt aacaaactta gaatttgaac taattctaaq 300
gaatttcaag aattcggcaa atgtcgggca gtgacaaac abttcaagac ggggacgat 360
ctacaaactt atggagataa gggcaacccc tccctctgta cttacagtc oattctctac 420
acagtctttg attaaatatt cacatttttt ctctacttaa agaccttcaa gaacagtacy 480
ta 492

<210> 240
<211> 519
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(519)
<223> n = A,T,C or G

<400> 240
tgtatagaag tagtggcttc cccatgtgat agtctgaant atagccctat gggatggag 60
gctgtgcccc agcccacac cagttaacgg tctgtgctga ggtggattag taaaagagga 120
aagcccttga gttgagatag aggaagggca ctgtctcttg cctgccctg ggaactgaat 180
gtctcgggat aaaaaccgat tgtacatttg ttaattctg agataggaga aaacccaccc 240
tatggcggga ggcagagcat gttggosgca atgctgcctt gttatgcttt actccacaga 300
tttttgggag gaggaaaca laaatclgpc ctacgtgcaac atccagcat agtaactccc 360
tttgaactta ataatgacac agattccttt gctcaaatgt ttttttgctg aactctctct 420
tatatacacc ctgctctctt accgcatcc ttgtctgag ataatgaaa taatatcaat 480
aaaaacttga nyygaactcg agaccactac gtogataca 519

<210> 241
<211> 771
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(771)

```


WO 01/90152

PCT/US01/16776

68

```

<223> n = A,T,C or G

<400> 241
tgatcgacg tagtgggtct cactcccgcc ttgacggggc tgctatctgc ctccagggcc 60
actgtccagg ctccggggtg gaagtcacit atgagacaca ccagtggtgc ctgtgtgget 120
tgaagctcct cagaggaggg tgggaacaga gtgacggagg gggcagcctt ggctgacct 180
aggacgggtc gcttgggtccc tccgccaacc ccgacagttg tgctgcttgt atatgagctg 240
cagtaataat cagctcgttc ctccgctcgg agcccaagaga tggcagggga ggcctgtgtg 300
ccanacttgg agccagagaa ggcattagaa acccctgagg gccgattacc gacctcalea 360
atcctgattt tgggggctct gctcgggtgc tgttgggtac angagacatt attataacca 420
ccaaggtcac tgtgtgtccc atgcaaggga aatgtgttga tccactgttc ccagaaacc 480
actacgtcca taccactcca claatggccm gcgcctgcca gttccacaa tattggggaa 540
naactccccc ccgccttttg ggttgnocat naacotttga aatttttccc tattantgt 600
ccctcaaaaa taacccttgg ggccttaate catggggccc atahcttttt tncocgttt 660
ttaaaanttg ttatcccgcc cncocnattt ccccccacac ttcccacaa ccgaaacant 720
tnaaatfint tnaaacctg ggggggtccc mnaattnnan ttnaahctnc c 771

<210> 242
<211> 167
<212> DNA
<213> Homo sapien

<400> 242
tgggcaactt caatctggg ctoatcgata acatcacgct gctgatgcty olgttcttg 60
tcctctctag gacctctcgg attttcaaat totttgagga atctatccaa attatctgoc 120
tcctctctct tcctctcttt totaaggtct totgttaca ggggtca 167

<210> 243
<211> 338
<212> DNA
<213> Homo sapien

<400> 243
ttgggcaact tcaatctcta ctgatctaaa tagtgtggtt tgaggccctct ttttctctgc 60
taaaatctct tggcagaggt caactctcac ttcccatag aggtaaaaat cttaaatggt 120
atattcttga caagctagc atagagagag caatttaca caaggtatth tccactgttt 180
taatacaagt ggttttctca caacctatgg gtgcaccaa gggagagtg cbaagttgca 240
gaacacaaat aagatactga agacacacat acttaccatt tccgtatag ctacaccaca 300
gttccactgt acatgtatgt tottatgggc aatcaaga 338

<210> 244
<211> 346
<212> DNA
<213> Homo sapien

<400> 244
ttttgggtct caatcacgca cactctctg ggaatgtct gttctaaggt caacccctaa 60
tgcaaaaaac abcatatcac ttgagatgc cctgtctcagg tacaatgtat ttaatttat 120
cactgtatca atgtatccaa taccagtttt agtctggcct tgaalcaaat cactgttctt 180
gttgtataaa aagagaataa tttagcttat atttaagtac catattgtaa gaanaaagat 240
gottatcttt acatgctaaa atcatgtatct gtacattggt gcagtgata ttactgtaaa 300
agggagagag gaatgaagcc gagctaagga tattgaaggt gcccaa 346

<210> 245
<211> 521
<212> DNA
<213> Homo sapien

```

WO 01/00152

PCT/US01/16776

69

```

<220>
<221> misc_feature
<222> (1)...(521)
<223> n = A,T,C or G

<400> 245
accatccca caccgatct gagggaacg tatatcatcc ctttcatcc ctacagcagc 60
aacttcacta ggcaggagtt attagtccca ttttaagaa gaggaactg agacttaggy 120
agatcaagta atttgcacag gtgcacacat tagtgcctga gccagggctt gaagcgagct 180
ctgtcttaag ccaatgcacc ctgcagatta ttgagagcac tgtttctcac aacagtgtaa 240
gccttttgcg anaggtccag gtccaaagg gcagagattt tggctgttt tgcctattgc 300
tccttcacca ttgcttagag cagggtctgc caccgaacag gtintcaatg oahgtttat 360
aatgtatat aagagcaaac atatgttaca gagaacttc tgtatgctg tcacttacat 420
gaatcacctg tganatgggt atgcttggtt cccatgttg cagatnaaga tattgaangt 480
gcccacaaca ctanttgagg gcgctgcan gtccancata t 521

<210> 246
<211> 482
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(482)
<223> n = A,T,C or G

<400> 246
tggaaaccaat ccaaatatccc atcaatgata gactggatna agaaaatttg gcacatgite 60
aoccatgaat actatgcagc cataaaaag gatgagttca talcctttgc agggacatgg 120
atgagctggy agacccatcat tctcagcaaa ctacaagggg aacagaaac caaacatgc 180
atgttctcac tottaagtgg gagctgaaca atgagaacac atggacacgy ggaagggaac 240
atcacacagt ggggcctgct ggtgggtagg ggtctagggg aggcatacga ttaggagaaa 300
tacctaatgt agctgacggg ttgatgggtg cagcaaacca coactgacag tgtataccta 360
tgtacaacac ctgcattgtc tgcacatgta ccccaaacct taaagtgtta ataaaaaat 420
taagaaesaa gtttaagtatg tcatagatad ataaatatt gtonatattg aagytgocca 480
aa 492

<210> 247
<211> 474
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(474)
<223> n = A,T,C or G

<400> 247
ttcgatacag gcacagagta agcagaaaaa tggctgtggt ttaaccaagt gaglacagti 60
aagtgcagga ggaagcagga agacaggggc atatgcaggg ggtgattata acaggtgggt 120
gtgctgggga atgaggttac tgggggtgna ggaacagtga ahaagtygca ahaagtggt 180
agatcagtag attgtacttc tccgaattt gatttctggn ggaagtcacat aactatcag 240
tttgggttat catanggcaa cagttgaggt ataggaggtg gaegtoung tgggataatt 300
gaggttatga anggtttggt actgactggt actgacaaag tctgggttat gaccatgga 360
atgaatgact gtanaagcgt anaggtgaa actattccac ganaaagggg tccnaaaact 420
aaaaannnaa gnnnnngggg atatatttt atgtggatat tgaangtgcc caaa 474

<210> 248

```

WO 01/90152

PCT/US01/16776

70

```

<211> 355
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(355)
<223> n = A,T,C or G

<400> 248
ttcgatacag gcaaacatga actgaaggag ggtggtgagc atcatgatgt tccgatggt 60
oaggatggnc aggaagaagc actggaacac gtgttlacgt ccttttgcct tgttgatggc 120
cctgagggga cgcaggaccc ttatgacctc cagaatcttc acaacgggag atggcaactg 180
attgantccc atgacacca gagacacccc aaccacacag atatcattat attgatgtag 240
ttcctgtaga nggcocccct gtggaggaaa gctccatnag ttggtcatct tcaacaggat 300
ctcaacagtt tccgatggct gtgatgggca tagtcattat taacattgta tggaa 355

<210> 249
<211> 434
<212> DNA
<213> Homo sapien

<400> 249
ttggatttgt cctccaggag aacaaaggga aaaaggtgac cgagggtccc ctggaaactca 60
aggtatctca ggaacaaaag gggatggggg aattcctggt cctgctggtc ccttaggtcc 120
acctggtccc cagggttac caggtctcca aggcacaaag ggtacacaaag gctctactgg 180
aocgctggcc cagaaagggt acagtggctc tccaggccct cctgggccc cagctccacc 240
tggtagagtc attcaacntt taacctcttt gtctccaaa aaaaacgaga gacatactga 300
agggatgcaa gcaatgcag atgataatat tcttgattac tggatggaa tggaaagaa 360
atttggtccc ctcaattccc tgaacaaga catcgagcat atgaaatttc caatgggtcc 420
tcagaccat ccaa 434

<210> 250
<211> 430
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(430)
<223> n = A,T,C or G

<400> 250
tggatttgtc aactggcaga gacaggatto caaggcagtg agaggagget acaatgcttc 60
caactagtta ttattattta tttattttt gaggatgaag ctgctttgt ctccaaaggt 120
ggagagcaggt ggtgcgatct tggctctctg caaccocaga ctcaagaat tctctgtct 180
tagctctccc ggtagatgga attacaggcc cccacagcca tgcacaccta attttttgt 240
gtctctagta ggaacagggt ttgacatgt tgggcagggt ggtcttgaa cctgacctc 300
nagtcatctg cctctctgg cctcacaaag tgcctgaatt acaggcatg gctgctgac 360
ccagtcaact tctcaatagt tatggctta tcaatttcc ccaattctat tggcccaaaa 420
aaaaaaaaan 430

<210> 251
<211> 329
<212> DNA
<213> Homo sapien

<400> 251

```

WO 01/90152

PCT/US01/16776

71

```

tggtactcca coatyatggg gtcaacggcc atcctcgccc tctcctggc tgttctccaa 60
ggagctcttg ccagggtgca gctgtrtcag tctggagcag aggtgaaaaa gtccggggag 120
tctctgaaga tctcctgtaa ggttcttgga taacacctta agatctactg gatcgcttgg 180
gtgcgcoagt tgcctgggaa aggcctggag tggatggggc taatctttcc tgatgactct 240
gataccagat acagcccgtc cttcaaggc caggtcacca tctcagtcga taagtcacac 300
agccaccgct atctgcagtg gagtaccaa 329

```

```

<210> 252
<211> 536
<212> DNA
<213> Homo sapien

```

```

<400> 252
tggtactcca ctcaguccaa ccttaattaa gaattaagag ggaacctatt actttctcc 60
caggctctctg tgccttaacc aggcctcttg gacagtatta gaaaggatg tctcaacag 120
tatgtagatc ctgtactggc ctaagaagtt aaactgagaa tagcataaat cagaccaaac 180
ttaatggctg ttgaagcttg tgtcctggag cagctgggat aggaacctt ttgggcagca 240
agaggaagaa ctgctctgaa gggggctca tgttaaaat tacaaggga acccacacca 300
ggcccccctc caagctctca gcttagagta ttgacatttc tcagctagag actcacact 360
tccttgctta gaatgtgca cggggggag tccctgtggg tgatgaggt ctaaggagt 420
agatggcat cctatcttct gtgtgcccac agggactgg ccagagactt agcaggtgaa 480
gtttcgggtc caggctttgc ccttgaacta ctatgtgacc tctgtggag taccaa 536

```

```

<210> 253
<211> 507
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(507)
<223> n = A,T,C or G

```

```

<400> 253
ntgttctcat cccagtaact cgggaagctg aggggggagg atcacclgag ctcaggaggt 60
tgagccgcca gtgagccggg acacagccac tccactccag cctggggcat agatgagac 120
cctccagagc agaaagaaa agaaaggag ggaagggaa aggggaagg aaaaagaaa 180
ggaaaaggaa aaggaagaa cagacaaaa cagacttga atttgatct cctgaactca 240
atttattgtt cttctacac caaacttct ctgactacta agatgataat ttgaacaccc 300
ctcgtttcct tctttacagc agcttggaag ttgtgtcaag taattacaat aatagtaaca 360
aatltgaata ttatatgcca ggtgttttct attcctgctc taacttaatt ctcaccactc 420
tgatataaat acaattgctg cgggtgtgag tggctcatgc ctgtaatccc ggcactttg 480
gagaccgagg tggcgggats gcaacaa 507

```

```

<210> 254
<211> 222
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(222)
<223> n = A,T,C or G

```

```

<400> 254
ttggattggt cactgtgagg aagccaaato ggatccgaga gtctttttct aaaggccagt 60
actggccaca ctttctctg ccgccttctt caaagctgaa gacacacaga gaaaggcgt 120
tctgttttac tcccaaatgg taactccaaa caatagatgg ttactnccc tgcctatct 180

```

WO 01/90152

PCT/US01/16776

72

```

tccacatccc tgcattccag tatagtcctg ggaccaatcc aa . 222

<210> 255
<211> 463
<212> DNA
<213> Homo sapien

<400> 255
tgttgagatc cctccatgct gaaatggaaa taacaacacat gatgagggag gattaagtta 60
gggagggagc acatcaaggt ggccatgaa ttgttggaa gaagtgaclt ttgaacaagg 120
ccttggtgtt aagactgat gagggtgca agacaaggg ggccactggt acaatagagc 180
agatgggaga ggaattgaaa ggttgagaa ataggaggg gtittgtctg gtattggtct 240
agtgaacacc gaggcgagag gccctggttg gtgcagctcg agagttatgc agataacat 300
taggcctcgt gggggaactgt agactgtcag caataatcca cagtttggat ttattcttaa 360
gagtgatggg agccctggg aggggggtta agcaaggagt gaaattatca gatttacagt 420
gataaaaata aattggtctg gctactgggg aaaaaaaa aaa 463

<210> 256
<211> 262
<212> DNA
<213> Homo sapien

<400> 256
ttgatttggt caactgtctc aactctacvt ttctctcttc ttcttaaaa attaatgaat 60
ccatcacatt atgccaataa ccttgggtt ttatcaakat ttotgttaa aagtattatc 120
cagaactgga cataatacta cataatata catataaac ccttctcttg gatgcaaca 180
tctattaata tagcttaaga tcaatttca ttacagaag caacatctg ttgatgttat 240
tttgatgttt ggaccaatcc aa 262

<210> 257
<211> 461
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1) ... (461)
<223> n = A,T,C or G

<400> 257
gngggnnnnn nncasattcg actcngttcc ontggtannc ggtcgacatg gcgcgggat 60
taccgcttgt nncgggggt gtatggggga ctatgacccg ttgtagctgg ggggtgatgg 120
gggactatga ccgcttgtag mtggkgtgt atggggact atgacogott gtccgggtgt 180
cggataaacc gacgcaaggg acgtgatcga agctcgcttc ccgtcttttc gcacgtgtag 240
ggatcatgga agcaalato cgcaticgyc tgaaggcgtt cgaccatcgc gtgctcgatc 300
aggcgacccg cgaalagcc gacacggcac gccgtaccgg cgcgctcacc cgcggtcaga 360
tcocgcttcc cccgcgcacc gagaagtcca cggtaaacgg tggcccgcac gtgacaaga 420
agtgcggga gcagttcgag gtgcgtacct acaagcggtc a 461

<210> 258
<211> 332
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1) ... (332)
<223> n = A,T,C or G

```

WO 01/90152

PCT/US01/16776

73

```

<400> 258
tgacgccttg tagctggggg tgtatggggg actacgacgc cttgtagctg ggggtgtatg    60
ggggactatg accgccttga gctgggggtg tatggggggac tatgacgcgt tgtagctggg    120
ggtgtatggg ggaactaggac cgtttgtagc tgggggtgta tgggggacta tgacgccttg    180
tagctggggg tgtatggggg actacgacgc cttgtagctg ggggtgtatg ggggactatg    240
accgccttga nctgggggtg tatgggggac tatgacgcgt tgtgtgctct gggggatggg    300
aggagagttg tggttgggga aaaaaaaaaa aa                                332

<210> 259
<211> 291
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)... (291)
<223> n = A,T,C or G

<400> 259
taccgccttg gacgccttgt gaccgccttg gaccgccttg gaccgccttg gaccgccttg    60
gaccgccttg gaccgccttg gaccgccttg gaccgccttg gaccgccttg gaccgccttg    120
gaccgccttg gaccgccttg gaccgccttg gaccgccttg gaccgccttg gaccgccttg    180
gggtgtctgg ggggactatg nnganngina cnggggggtat ctgggggact atgannagct    240
gtgcnnectg ggggatcnga ggaganlngn ggtagnngat ggttngggan a          291

<210> 260
<211> 238
<212> DNA
<213> Homo sapien

<400> 260
taagggggta ctgggttaaaa tacaggaaat ctggggtaat gaggcagaga accaggatac    60
tttgggtaaa gggatgaaaa ctagaatttt ttctttttt tttgctgag aaacttgctg    120
ctctgaagag gccaatgtat taatgtgttt gatcttctct ttcttacagc cctttcaagg    180
gcagagccct ccttactctg aaggaaatct atccttagct atagtatgta ccttcttta    238

<210> 261
<211> 746
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)... (746)
<223> n = A,T,C or G

<400> 261
ttgggcacct tcaatatnaa tagctaaact ttattgagtg tttatcgatg gataaaacac    60
tgtttctaaq atttaaacgt actaatctat ttaattgctca taataacttt agnagtgpg    120
tactagtatt agtctcaattt acagatgcaa catgcaggca cagagaggtt aatkaacttg    180
cccaaggtta ccaagctaag aaatagaaaa aatattgcat ctggaaagtt gggtctctgy    240
gtaaaccaca gactcttana tgagcctggg gccctactca gtttgccttt acaaagcgaa    300
tgagtaacat caattaatte agtgggtagg ccaaatggag gtcagctacg agtttctgct    360
gtttcttgag tggactgaca gatgtttaca acgtctggcc atcagtwaat ggactgatta    420
tcatttggaw gtgggtgggc tgaatgttgg ccagtgaagt ttattcawgc catattttta    480
tgbttaggat gacttttggc tggctcctag gcaagctctg tctgscagg aacacagaat    540
wacacaggga cccctcaaat ttctggtgtg gctagaaaca tgaaccactg gttgggggaa    600

```

WO 01/90152

PCT/US01/16776

74

```

caagcggtga aaacotaagt ggcgcggcgt ggcaggggtcc acccatutgg ggaaaactcc 660
caagcggttt ggaatgcctn agctngaat attctaanag ttgtconent aaaattagcc 720
tgggcggtta toangggctn naagcc 746

```

```

<210> 262
<211> 588
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> [1]...[588]
<223> n = A,T,C or G

```

```

<400> 262
tgacggcttg tcatctcaca tggggctcct caegcttttg oottttagg aacotgaca 60
tttgtctgtt tcttcttctt othttcttc ccatatcttc ctaatttagg tttgacttgt 120
tttgtgagga ggcaggagct agagactgct gtgagctcat aggggtggga agtttatcct 180
tcaagtcccg cccactcctc actgctcttc acctccctct gaccaggctt acaagtgggt 240
tcttgccgtg ttccctcttg paccacacaa gccctgttaa tgagtgtgca tgactctgac 300
agctggggac tcaaggctct tggctacag tgcctgttaa aatctctcat ccagttctcg 360
caatttgyfa aaataaccaa atttcttaga ttccagtacc caaatcatgt ctttaccgac 420
tgctcctcaa acccagaagt ggcacacaaa ttcttgggga attattactt tttttttct 480
ctctnttncv gnnngnnncg gnnngnccag gaattaccac nttagaagac atgcoogaa 540
tttattatac aggggagccg attatttttc ctaacacaaa gcgggtca 588

```

```

<210> 263
<211> 730
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> [1]...[730]
<223> n = A,T,C or G

```

```

<400> 263
tttttttttt ttgggcctga gcaactgaaa ttatgaattt tccatatact caaaagagta 60
agactgcata agatttaaat gtaaaagtgg tcttgatata agtaatgttt aagatacctc 120
ttanatttat aattgaaaaa ttagggaatt tggatataca agttgaaat toaggagtgta 180
ggttgggctg gctgggtata taactgaaac tgcagtaca cagatgcat ctanaeccac 240
aaatctgggt ttatttttag agtgatatgt gtcaactcca caaaagcctt ccaatttggc 300
ctougcatac acaacaaagtc acctccccc agccctctac acataancaa attccttagt 360
ttagtccagg aggaatagcg cctttttctt tccgtcttag gtgacoycaa ggcacagttc 420
tgcacacaaa galyttaegy gaagtctgac aaagaggcat ctgaaaggaa ataaagggaa 480
tggaagtbac caaaaggaa aaccaagaa aaactttgga gacggtttct aganccttgy 540
oatttcaaaa caaaactcag gaacaaacct tgtctcaca atcatttaag cctttgttll 600
ggannagact ttctgaactg ggcactgaac ataancotea ttgaatgtcl tcacagcttc 660
ccagctgaag gcacaccttg ggcagaaggy ggaatctton aggtctcaa nacagggtcc 720
gccttttenc 730

```

```

<210> 264
<211> 715
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature

```

WO 01/90152

PCT/US01/16776

75

```

<222> (1)...(715)
<223> n = A,T,C or G

<400> 264
tttttttttt ttggccagt atgatatgt ctacactat attgaagctc ttaggtcatt    60
tcaacttaak gtggltatag atgctgttga gcttacttct accacottgc tttttctccc    120
gtctcttttt tgttcccttt ctcttttttt cctcccttat ttataaahg aatttttttag    180
gattctattt tatatagatt taccagctat aacactttgt atbtctttgt ttgttgcttc    240
ttctgttatt tcaatgtgga tcttaactc atcacactct attttcaant aatatactat    300
aaccttaact ataatgtang aacttaacc cctatacttc catttctccc ttccatccta    360
tgintgtcat attttttctt ttalatatgt tttaagaca taatagtata tgggaggttt    420
tgccttaaaa tgtgataaat attccttcaa ngsaacctaa aaatttcaaa taatattctg    480
ttattctcca aatnnaccta atatttctca ccatnctna taontttcaa gaatctgaag    540
gcatttggtt ttccggctt aagaacctcc totaaagcac tctaaagcga attaagtott    600
ctgggagagg aattctccca agcttggcco ttamttgia ctccntang gttaaanttt    660
ggccgggaaa tagaaattcc aagttaacag gntanttttt ntltttntn tcncc       715

<210> 265
<211> 152
<212> DNA
<213> Homo sapien

<400> 265
tttttttttt ttcccaaca caaagcaaca ttatctttcc tcaaatatt caadatagtt    60
tgattcccat gaagaggtta tgatttctaa agaaacatg gctactata tatcaatcag    120
ggttaaatct tttttttttg agacggagtt ta                               152

<210> 266
<211> 193
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(193)
<223> n = A,T,C or G

<400> 266
taaatctcgt ccccttttta stcaatatgg aggtaccca ctccacatta ccttcttttc    60
aagggaactgt ttccgttaact gttgtgggta ttccagacca ggcttctaaa cctcttaaaa    120
ctccccaatt ctggtgccaa cttggaaaac atgctttttt tttttttttt tttttttttt    180
gagacggagt tta                               193

<210> 267
<211> 460
<212> DNA
<213> Homo sapien

<400> 267
tgttgcgact ccttaagcat ggggtctatt aaaaaatgg tggagaagaa aatacctgga    60
atttaagttc tatctttaga gattgggaag accctgatgg aggcctgga gaacagcttc    120
ttcttgatgt tcaattccca agtaacacca gtgtgtcagg caattgctaa ggaacctaaa    180
ttgcagcaag gctacaatgc tatggggttc tccagggag gccaatctt gaggcgagtc    240
gtcagagatg gccttctccc tccatgatc aatctgatct cggttggggg acacatcaa    300
gggttttttg gactccctcg atgcacaga gagagctctc acatctgtga ctctatcga    360
aaaaacatga atgctggggc gtactcnaa gtgttctcag aacgctcgt gcaagccgaa    420
taatggcatg acccataaaa ggaggatgtg gatcgcaaca                    460

```


WO 01/90152

PCT/US01/16776

76

```

<210> 268
<211> 533
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> [1]...(533)
<223> n = A,T,C or G

<400> 268
tggtgagatc ggttgataga atagagacgt ggtaatgagt gcatggacag nctcgaatt  60
accttgcgcc gtggggaccc cagtagcgtc taaggcgtcg tcaattagag taacctctgg  120
acgcacgggc ggttgcgatt taccggaagc ggcagctgca gtggcttgc gccccgggcc  180
aaattcttltg gggggtttaa ggcgcggggg aatttgaagt atctctatca gtatgtagcc  240
aagtbtgaac agtcgccatt ccgcaaatcg ctctctttga atucgaccc cctccagcat  300
tgctcattc atcaacctga aggcacgcat aagtgcgggt tgtgtcttca gtagctccac  360
tcataaacta gcgcgctcga cctcgtcttc gtacgcgcca ggtcgtgctg tgggaattcc  420
caactcgtgt gagtgcgca tttaagttt cgaactgtt cgcctccacn atttggcatt  480
ttcacgcatg acacggaata aactcgtcca gtacgggaa tgggatcgca aca  533

<210> 269
<211> 50
<212> DNA
<213> Homo sapien

<400> 269
tttttttttt ttgcgctgaa ttactatag atctctctca caagcggtea  50

<210> 270
<211> 519
<212> DNA
<213> Homo sapien

<400> 270
tggtgagatc caataaaccc accagcttct tgcaaatcgc gcagaagcca ccgtctcttg  60
gctgagtcac gtgaacggtc agtgcaagca gcgcgctgcc agagcagagg tgcagcatgc  120
tgcaacacag ctcaaggctg acctcctcca gcaggatgga cagatggag ctgcgtatcg  180
tgtccacacc ctccgggcac tcttcgcgca gggacttcgg cagcttcgag caattttgt  240
caaaagcgctc gactatttct ttctcagctc tgtgtgtgto aatcagcttg gtcacctcct  300
tcaccaggaa ttacacaccc tcaacagaaa catcagaactt tgctggggcc cgtgtcttct  360
taattggctc caccagttcc aaggcaaggc tgacattctt ggaggccact ttggcgggga  420
ccagagctcg catgggcctc tctttacact catcacagaa ccaaccagc gcaacagatct  480
ccttgggttg catgtgcctc atcatctggg atcgcaaca  519

<210> 271
<211> 457
<212> DNA
<213> Homo sapien

<400> 271
tttttttttt ttgcggcggc gacggacgtt gcactcctcc agtagcggct gcaagtctgt  60
ccaatggccc gctatgagga gctgagcgtg tccgcttctg aggaattcca ccggcccggt  120
gaacagcaca atgcaagac catcttcgcc tactttangy gttctaaaga cccggggggg  180
aaaagctggc gcccgagctg cgtgcaggct gaaccagtcy taagagaggg gctgaagcac  240
atttggagag gatgtgtgtt catctactgc caagttaggy aagagcctta ttggaagat  300
ccaaataatg acttcagaaa aaacttgaaa gtaacagcag tgcctacact acttaagtat  360
ggaacacctc aaaaactggt agaactcgag tgtcttcagg ccaacctggt ggaatgttg  420

```

WO 01/90152

PCT/US01/16776

77

```
ttctctgaag attaagcttt taggatggaa atcaaga 457
<210> 272
<211> 102
<212> DNA
<213> Homo sapien

<400> 272
tttttttttt ttgggaacca acatgaatac ctlttcaagg ctctgcttg ggctcaagcc 60
cgccggggaa atgcaactgg ccaggctaca gggaatcaa ga 102

<210> 273
<211> 455
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...[455]
<223> n = A,T,C or G

<400> 273
tttttttttt ttgggaacca acaggtttta gtcttcggcc gaagttaac tctgttttt 60
ggcaatcaac aggtttaagt ctctggccga agttaatctc gtgtttttgg caatcaacag 120
gtttaagtct tggccgaag ttaatctcgt gttrtrgcca atcaacaggt ttaagtcttc 180
ggccgaagtt aatctcgtgt ttttggcaat caacaggttt aagtcttngg ccgaagttaa 240
tctcgtggtt ttggcaatca acaggtttta gtcttcggcc gaagttaac tctgttttt 300
ggcaatcaag aggtttaagt ctctggccga agttaatctc gtgtttttgg caatcaacag 360
gtttaagtct tggccgaan ttaatctcgt gttrtrgcca atcaacaggt ttaagtcttc 420
ggccgaagtt aatctcgtgt ttttggcaat caana 455

<210> 274
<211> 461
<212> DNA
<213> Homo sapien

<400> 274
tttttttttt ttgggaacca cocttgatga acatcaatgt gaaatcctc ggtcaatac 60
tggcaaacca aatccagcag cacatcaaaa agcttatcca ccagatcaa gtgggcttca 120
tccctggggt gccaggctgg ttcaacataa gaaatcaat aatgttaac catcacataa 180
acagaaacca agcaaaaacc cacatgatta tctcaataga tgcagaaag gccctggaca 240
aatcaacag cocttcctgc taacacatct taataacta gatattgat gaatgtatct 300
caaaataata agagctattt atgaacacc ccagccaat atcatatga atgggcaag 360
actggaagca ttcccttga aaactggcac aagacagga tgcctctct caccgctcct 420
attcaacata gtattggaag ttctggccag ggcacacag a 461

<210> 275
<211> 729
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...[729]
<223> n = A,T,C or G

<400> 275
tttttttttt ttgggaacca caagtcttc cagtgaggag gttttattat gttttcaac 60
```

WO 01/90152

PCT/US01/16776

78

```

catgaaacaa taggaagggtg gctgttaacg caaacatttc agatagaaga ataggccaaag 120
ctcccaaac ccacacttca cagcctcttc cacacgtctc ccanagcttg ttgtccttca 180
cttgcaaat canggatggt ggaagtngac atttonagtn gcnggaaccc catcagtcaa 240
ncantaagca gaantacgat gactttgana nacanctgat gaagaaacacn ctacnganaa 300
cccttctnt cgtgttanga tctnngtcc ntaactaatg cggccccctg cnggtccncc 360
atttgggaga actcccccnn cgttgatccc ccccttgagt ntccattctt ngtcocccan 420
acnngncttg nngnncanln cncctcnca cctgtttcc ctgungtnaa aatnngtttt 480
nccgccnccc naattccac ccaaatcaca gcgaanccng aaggccctcn naagtgttta 540
agcccnngg gtttccctnt ntanttgcaag ctaaccctcc cacttnnmt tncngtttg 600
tcggcccttg gncnccnln gtlcccttt nngnncacn cctugntcnn ngvncntcn 660
nmcttttca tnnacttago tngctntcc nncnngngg ncnngcaca ttnncnnac 720
tngttnnc 729

```

<210> 276
 <211> 339
 <212> DNA
 <213> Homo sapien

```

<400> 276
tgactcgaca tglagtagat acttaataaa tatttctgga atgaatggat gaagtggagt 60
taccagagaaa atagaaaag tacaattgt tctcagtgtt ttgaaggaaa attatgatct 120
ttcccnaagt tctgacttca ttctaagaca ggttagtat ctccatnact aattttactt 180
gcttttgaaa atcaattgag ataatctatt tagattgata attattttag actggglata 240
aactatttag tgcctgcaaa tatcaattt aatctcattt tccacctctt gtgatatagc 300
tatgtagggt ttgactttaa tggatgtcag gtcaatccc 339

```

<210> 277
 <211> 664
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(664)
 <223> n = A,T,C or G

```

<400> 277
tgactcgaca tccataacaa aatctttctc catttatctt ttctagggga atttcttga 60
aagactccca aggaatacaaa tgaatgtaag acgtgccaa gtggggggca gaccaccaag 120
taagaccaca gattttacat tcaacagcta gctcacagta ctttgccgca cactgtgggc 180
agaatagccc tctaatgtc agccctggct cgttattgcc atccaatgc gccatgctga 240
aagaagggtt tgcctcctgg tcaagatnaag aagcaatggt gtgctgagge aatcccatat 300
gaataagtga gacttcagaa cttgagctag caggaggagg actaangatga tgtgttagca 360
actctttgta atggctttcc tctaaatnaa cttggtacgt gccacagtt tccagagcaa 420
gtacagtgca aacgggaact tctgcagaca atccaataac agatactcta attttagctg 480
cttllagggt cttgattaaa tcaataatat tagatggato gcaagtgtga aggtgtclaa 540
aagatgatta gtaactctcg acttgtatgt ccaggccatgt tgttttaaa totgccttag 600
nccctgctta ggggaatttt taagaagat ggcctccctt gttaanggtc aatcacnaat 664
tgcc 664

```

<210> 278
 <211> 452
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(452)

60
120
180
240
300
360
420
431

WO 01/90152

PCT/US01/16776

80

```

<212> DNA
<213> Homo sapien

<400> 282
attgatttcg atgcttgagc ccaggagttc aagactgcag tgagccactg caettcaggc 60
tggacaacag agcagagtcgc tctgcacaaa aaaaaaaaa 90

<210> 283
<211> 764
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(764)
<223> n = A,T,C or G

<400> 283
ttttttttt ttgcgaagca cgtgcacttt attgaatgac actgtagaca ggtgtgtggg 60
tataaacgc tgtatctagg ggcaggacaa agggggcagg ggcacacagc ccagcgtgca 120
gggccacacat tgcacagtagg aatgcacaaag ttgcaggcta tggccggcta ctavtaaccc 180
cgttttttcc gttttatctg taacataata tggtagactg tcacagagcc gaatttcact 240
haacagatgc atccaaaggt cagagagatg ccacaaatca gggccacat attcaggcac 300
ttggcggtgg gggatagagc ctgggcccgc gtaacgtccc caacowctly cctgtcccta 360
cmcttgavtc cmcctctnn nninecnaa knigcccgcc cmctctctng ngtcacccng 420
natctgcact anctccctcn cccctnttgg antctctccc ttcacataaa nttatccctn 480
acncccccc cncctttccc ctnccncccn tnatcccngr nccnctatca ntcnncct 540
cncntnctcn cmctagttc cmcctnnaa ctacnctln nccnncct cactnncct 600
ngnnantttc ttccttccct cccacgcnn tgcgtgcgcc cgtctngcct nnnctccgna 660
cccnactttt attaccttt nccactagc actctactin acccancnc tscctacccc 720
ngnccacccc nncctnatic nctnctctn tcnctctntt cccc 764

<210> 284
<211> 157
<212> DNA
<213> Homo sapien

<400> 284
caagtgtagg cacagtgatg aaagccttga gcaaacacaa tctgtgggta attaacgttt 60
attctccccc ttccaggcac gtcttgcatg gatgatcaaa gatcagcttc tggccacat 120
aactaagcta gtttaagata cgttcccata cacttga 157

<210> 285
<211> 150
<212> DNA
<213> Homo sapien

<400> 285
attcgattgt actcagacaa caatatgcta agtggagaa gtcagtcaca aaagacaca 60
tactgtatga cttaatttac attaatgtc cagcaatagc aactccgtag agacagaag 120
tagatgagca gctgctagg tctgagtaca 150

<210> 286
<211> 219
<212> DNA
<213> Homo sapien

<400> 286

```

WO 01/00152

PCT/US01/16776

81

```

attcgatttt tttttttttt gccatgatga aatttctact cccctcagatt ttttgtotgg    60
ataaatgcaa gtctacccac cagatgtgaa attacagtta aatttgaagg aatctcctga    120
gcaaccttgg ttaggatcga tccaatattc accatctggg aagtcaggat ggotgagttg    180
caggtcttta caagttcggg ctggattggt ctgagtaca                    219

<210> 287
<211> 196
<212> DNA
<213> Homo sapien

<400> 287
attcgattct tgaggctacc agggcctagg aagaaggcca tggaaacaaat ttctcctaat    60
atccatactc agaagggaac aacctgcctg acaccttaat ttacgcttct gccctctaga    120
actgtgaagc agtaactttc tcttggttta agcccaagaga atctgtcttt tggtaattta    180
tatcataucc tcaaga                    196

<210> 288
<211> 199
<212> DNA
<213> Homo sapien

<400> 288
attcgatttc agtcacgtcc cagaaccac accgtcaatt actactctgt aagaattcca    60
tttgttgaaa ttcatgtagt aaaaatttta tgcctcctta atctatgcaa attacattgc    120
taggtactga agattcaagt gaccgagatg ctagcccttg ggttaagtg atccctctcc    180
cagagtgcac tggactgaa                    199

<210> 289
<211> 182
<212> DNA
<213> Homo sapien

<400> 289
attcgattct tgaggctaca aacctgtaca gtatgttaact ctactgaata ctgtaggcaa    60
tagtaataca gaagcaagta tctgtatgt taacataata aaaggtacag tgaacttca    120
gtattataat cttagggacc acaattatat atgtggtcca tcaattggca aaaaaaaaaa    180
aa                    182

<210> 290
<211> 1646
<212> DNA
<213> Homo sapien

<400> 290
ggcacgagga gsaatgtaat tccatatttt attgaaact tattccatat tttaatgga    60
tatlgagtga ttgggttctc aaacacccac aaactttaat ttgtttaaat ttatatggct    120
ttgaataga agtataagtt gctaaccttt ttgtataaca ttgaagata gtattttacc    180
atctttacac atcttggaaa atacaagttcc tgtgaaacac caacttttca cctagagca    240
tgaggccaaa agtaaggctc ttaattata acatattgga ttctttagtag tatgtttttt    300
tcttgaaact cagtggtctc atctaacctt actatctctc caacttttct ctaagactaa    360
actctaggtc ctttaaaato tgcctcaccc aactttagaa gctctgaaaa gaatttgtct    420
ttaaatatct ttttaataga acatgtattt tatggaccaa attgacattt tgcactatth    480
tttcccaaaa agtcaggtag atttcagcac actgagtttg gaatttttta tcccagaaga    540
ccacccaatt tcatatttat ttaagattga ttccatactc cgttttcaag gagaatccct    600
gcagtctctc taaaggtaga acaaatctct totatttttt ttcccccatt gtgggatttg    660
acttttaagc gtgactctaa aaaaacagag aacaactatg tctcagttgt attaagcagc    720
gaccatatt atcatattca cttaaaaaan tgaattctcg tgcacctttt ggcaactttc    780
cttttaaty tagggaaaaa cttagtcacc ctgaanaccc acaaaatana taaaactttg    840

```

WO 01/00152

PCT/US01/16776

82

```

agatgtgggc agaggggttg ggggtggaca ttgtatgtgt ttaaattaaa cctgtatata 900
ctgaagaagct gtgtatgtgg tcagagaaaa tgaatgttta gaagctgttc acatcttcaa 960
gagcagaagc aaaccacatg tctcagctat attatatttt attttttatg cataaagtga 1020
atcattttctt ctgtattaat ttccaaaggg tttacccttc tatataaatg ctttgaaaaa 1080
cagtgcaatg acaatgggtt gataittttc tttaaaagaa aaatataatt atgaagacca 1140
agataatctg aagcctgttt tatittaaaa ctttttatgt tctgtggttg atgttggttg 1200
tttgtttgtt totattttgt tgggttttta cttgtttttt tgtttgtttt tgtttgtttt 1260
kgcatactac atgcagttct ttaaccaatg tctgttttgg taagttaatt aaagtttgta 1320
ctttatatga gtgcatttca actatgtcaa tgggtttctta atatttattg tgtagaagt 1380
ctggttaatt tttatttca aatagtitta agagataaac agtttgataa gttttcagtg 1440
gtttatagca gaagtatttt atttccatgg catccagcgg gatattttgg tgtttgagag 1500
gcattgcagtc catattttgt acagttagtg gacagttatc agcaagcctt gatagcttct 1560
ttggccttat gttaaataaa aagacctgtt tgggtatgat tttttatttt taaaaaaaaa 1620
aaaaaaaaaa aaaaaaaa aaaaaa 1686

```

<210> 291
 <211> 1851
 <212> DNA
 <213> Homo sapien

```

tcatcaacat tgcagcagc ggcacggta gtcagggttt ctgggaatcc cacatgagta 60
cttcogtgtt ctccattctt cttcaataga cataaatctt ctagtcttgg ctggctgttt 120
tcacttcttt taagcctttg tgcctcttcc totgatgtca gctttaagtc tgtttctgga 180
ttgtgttttt cagaagagat ttttaaatc tgtttttctt tglaglcaga aagtaactgg 240
caaatctcat gatgatgact agaaacagca tacttcttgg cgtcttttcc agatcttgag 300
aagatacatc aacattttgc tcaagttagg ggcctgactat acttgcctga ccacacata 360
cagaagatga gagagcagtt ctccatata latccagcgc atttaaatc gctttttctt 420
tgattnaaaa tllcaccact tgcgtttttt gctcctgtat accaagttag agtgggtgta 480
ggcactgctt gttttttgat tgcattatcg caccgtataa gagcagtgct ttggcactta 540
atttatcttc attgttagaca gcatagtgta gactggtatt tccatactca totggaatat 600
ttggatcagt gccatgttcc agcaacatta aagcacatto atcttcttgg catglaacgg 660
ctttgtcag agctgtcttc tttttgtgt caaggacult aagttgacat cgtctgtcca 720
gcagagatt tactactctt gaattccat lggcagaggg cagctgtaga gcagctctct 780
tttgtgtgtc ctctgtgttc aatcctgtgt ccttgagcat gacgtgaga tccctctcgg 840
ggacttttcc ccacacggca gctctgtgga gcttgtccag atcttctcca tggacgttgt 900
acctgggata catgaaggcg ctgtctctgt agtctcccca agcgaacacg ttgctcttgc 960
cgctccctgt aagcagggga agcagtgcca gcaccacttg cactcttctg tcccaagcgt 1020
cttccacagg aagctcctgtt gctctccaga agtgcacag ttgctcttgc cgtctccctt 1080
gtcctccagg gagggaagaa atgcaggaaa tgaagatgac atgcacgatg gtaactctct 1140
cagccatcaa acttctggac agcaggtcac ttcacagaag gtggagaag cgttccaccc 1200
acagaggttg agatccagaa accacatat ccattacaaa acaaacactt ttcagccaga 1260
cacaggtatc gaattcatgt catctgcggc aacatgttgg aacctaccca atccacatc 1320
aagagatgaa gacactgag tatatctgca caacgtasta ctcttctacc ataacaaat 1380
aatataattt tctctcggag ccatatggat gaactatgaa ggaagaactc cccgaagag 1440
ccagtcgag agaaagccca ctgaagctct gtctcagcc atcagcgcca cggacaggag 1500
tgtgtttttt cccagtgat gcagctcaaa gttatccga agctgctgca gcacacggtg 1560
gctctcgaga aacacccag ctctctcggc ctacacaggg caagtcaata aatgtgttaa 1620
tccatcaaac agaattaaaa gcaagtcac ahaagcatct caacagacac agaaagacca 1680
tttgacaaaa tccagcatcc ttgtatttat tgltcagtt ctacagggaa atgcttctaa 1740
cttttcccaa tttagtatta tgttgctgtg ggguttgta taggtgtgtt ttattacttt 1800
aaggtatgta ccttctatga ctgttttgc gagggtttta attctctgac c 1851

```

<210> 292
 <211> 1851
 <212> DNA
 <213> Homo sapien

WO 01/00152

PCT/US01/16776

83

<400> 292
 tcatcaaat tgcacagac ggcacagtta gtacgggttt ctgggaatcc cacatgagta 60
 ctccgtgttt ctccatcttt ctccatcttt catcaatctt ctacgtctgg ctgggtgttt 120
 tcatctcttt taagcctttg tgcctctctt tctgatgta gctttaagtc ttgtttctga 180
 ttgtgttttt cagaagagat tttaacatc tgtttttctt tglagtcaga aagtaactgg 240
 caattacat gatgatgac agaaacagca tactctctgg cogtctttcc agatcttgag 300
 aagatcacat aacattttgc tcaagtagag ggtgactat acttgctgat ccacaccta 360
 cagcaagtat gagacagtt ctccatata tatccagcgc atttaaatc gctttttct 420
 tgattaaaaa tttaacacat tctgttttt gctcatgtat accaagtagc agtgggtgta 480
 ggcatgttt gttttttgat tcatatcag caacgtataa gagcagtggt ttggcaatla 540
 attttcttt attgtagca gcatagtta gagtgatct tccatctctc tctggaaatct 600
 ttggatcagt gcatgttcc agcaacata agcaacatc atctctctgg cttgtacgg 660
 ctcttgctag agctgtcttc tttttgtgt caaggaactt sagttgact cgtctgtcca 720
 gcaagagttt tactactct gaattccat tggcagagga cagatgtaga gcagtctctt 780
 tttgttgtt cctctgttt acatcgtgt cctgagcat gagatgaga tctttctctg 840
 ggactttacc caacacagca gctctgttga gctgtctcag atctctctca tggagctggt 900
 acctgggac catgaaggg cgtctctctt agtctctctc agcagccag ttgctcttgc 960
 cgtctctctt cagcagggga agcagtgga gcaacacttg cactctcttc tcccaagcgt 1020
 ctctacagag gagtcttgt ggtctcaga agtgcacag ttgctcttgc cgtctctctt 1080
 ttccatcag gagggagaa atgcaggaat tgaagatgc atgcagatg gctctctctt 1140
 cagcaatcaa actcttgac agcaggtcac ttcagacag gtgggaaag cgtctctctt 1200
 acagcagtt agatccagaa acccaatct ccatcccaa acaaacactt ttacgcca 1260
 ccaagttact gaatcatgt catctcggc aactgtgttg aactctctca atccacatc 1320
 aagagtgaa gcaatgtcag tatatttga caagtaata ctctctctc atcaacaaat 1380
 aatataatt tctctgtgag ccatctgtat gaactatga ggaagaaatc ccgagagag 1440
 ccagtcagag agaagccaca ctgaagctct gctctcagc atcagcccca cggacagga 1500
 ttgtttctt cccagtgat gcaagctcaa gttatccga agctcggca gcaacaggtg 1560
 gctctgtgag acaacccca ctctctcgt ctaacacag caagtcata aatgtgata 1620
 tcatcaaac agaatanaa gcaagtaac ataagcatc caacagacac agaaagagca 1680
 ttgcaaaaa tcaagcatc ttgtatttat tgttgagtt ctacagagaa atgtctctaa 1740
 ctctctctca tttagtatta tgttggtgt ggtgtgtgta taggtgtgtt ttattacttt 1800
 aaggtatgt cctctctgc ctgttttgt gaggttttta attctctgtc c 1860

<210> 293

<211> 668

<212> DNA

<213> Homo sapien

<400> 293
 cttgagcttc caaakeyga agactggccc ttacacaggt caatgttasa atgaatgcat 60
 ttoagtattt tgaagatnaa attgtgat ctatcccttg ttttttgatt cgtatcagc 120
 accctataag agcagtgctt tggccattca ttatctcttc atttttagaca gortagtga 180
 gagtgttatt tcatctctca tctggaatat ttgatcagt gccatgttcc agcaacatta 240
 angcaaatc atctctctgg catgtctcgg cgtgtcagta ttgaccccaa aacaaatla 300
 catatcttag gaattcaaaa taacattcca cagctttcac caactagita fatttaaag 360
 agaaaaatca tttttatgcc atgtattgaa atcaaaccca cctcatgtgt atattgttg 420
 ctactgcata cttttatcag agctgtcttc tttttgtgt caaggaactt aagtgaact 480
 cgtctgtcca gcaagagtt tactactct gaattccat tggcagagga cagatgtaga 540
 gcaagtcata gcaagtgaga agactttta ggaacttga gtgactagc taccagcata 600
 gcaatgato atgtaactgc aaacactgaa tagctgtcta ttactctgac ttcaaaaaa 660
 aaaaaaaa

<210> 294

<211> 1512

<212> DNA

<213> Homo sapien

<400> 294

WO 01/90152

PCT/US01/16776

84

```

gggtgcacca gggggagcgt ggggtttcct cgggtgggtg tgggttttcc ctgggtgggg 60
tgggtgcggc tggaaatccc tgcctggggt ggcaggtttt ggcctgggatt gacttttlytc 120
ttcaaacaga ttggaacccc ggaattaccc gctagtttgt gaaactgggt gtagagcgcg 160
atctgttgcc tactactggc ttctctggc tgtlaaagc agatgggtgt tgaagttgat 240
tccatgcccg ctgctttctc tgtgaagaag ccatttggtc tcaaggagcaa gatgggcaag 300
tgggtgctgc gtgtcttccc ctgtcgcagg gagagcgcca agagcaacgt gggcaactct 360
ggagaccacg acgactctgc tatgaagaca ctcaaggaca agatgggcaa gtgtgcccgc 420
caactgcttc cctgtctcag ggggagtgcc aagagcaacg tgggcgcttc tggagaccac 480
gaagaytttg ctatgaagac actcaggaac aagatgggca agtgggtgtg ccactgcttc 540
ccctgtctgc gggggagcgg caagagcaag gtgggcgctt ggggagacta ccatgacagt 600
gccttcattg agccaggtta ccaagtcctg ggaagagato tggaaagctt ccacagagct 660
gcctgggtgg gtaaaatccc caaaaaggat ctatctgtca tgcctcagg caactgacgtg 720
aacaagaggg caagcaaaa gaggactgct ctacactctg cctctgcaaa tgggaattca 780
gaagtagtaa aactctgctt ggacagacga tgtcaactta atgtctctga caaaaaaag 840
aggacagctc tgaayaaagg cgtacaatgc caggaaagat aatg-cggtt aatgttgctg 900
ggacatggca ctgatccaaa tattccagat gagtatggaa ataccactct rcaataygt 960
rtctayaatg aagataaatt aatggcaaaa gcaactgctc tataygtgc tgatatcgaa 1020
tcasaaacaa aggtatagat ctactaattt tatcttcaaa atactgaaat gcattcattt 1080
tcacattgac gtgtgtaagg gccagtcttc cgtatttggg agctcaagca taacttgaat 1140
gaaaactatt tgaatgccc taattatctm agactttatt taaatattg ttattttcaa 1200
agaagcatta gagggtacag tttttttttt ttaaatgcac ttctgttaaa tacttttgtt 1260
gaaaacactg aatttgttaa agttaactac tactattttt caatttttcc cctctaggat 1320
ttttttcccc taatgaatgt aagatggcaa aattggccct gaaataggtt tbaatggaa 1380
actccaaaga aagttaaaaa tgtttcaagt aatagagato ctgctcttlt ggaagltcc 1440
taaaaaacag taatagatcc gaggtagtgc gccctcaggt ggcaggtttt aagatatttc 1500
tgatctctgt cc

```

<210> 295
 <211> 1853
 <212> DNA
 <213> Homo sapien

```

<400> 295
gggtgcacca gggggagcgt ggggtttcct cgggtgggtg tgggttttcc ctgggtgggg 60
tgggtgcggc tggaaatccc tgcctggggt ggcaggtttt ggcctgggatt gacttttlytc 120
ttcaaacaga ttggaacccc ggaattaccc gctagtttgt gaaactgggt gtagagcgcg 160
atctgttgcc tactactggc ttctctggc tgtlaaagc agatgggtgt tgaagttgat 240
tccatgcccg ctgctttctc tgtgaagaag ccatttggtc tcaaggagcaa gatgggcaag 300
tgggtgctgc gtgtcttccc ctgtcgcagg gagagcgcca agagcaacgt gggcaactct 360
ggagaccacg acgactctgc tatgaagaca ctcaaggaca agatgggcaa gtgtgcccgc 420
caactgcttc cctgtctcag ggggagtgcc aagagcaacg tgggcgcttc tggagaccac 480
gacgaytttg ctatgaagac actcaggaac aagatgggca agtgggtgtg ccactgcttc 540
ccctgtctgc gggggagcgg caagagcaag gtgggcgctt ggggagacta ccatgacagt 600
gccttcattg akcccaaggtt ccaactcctt ggaagaatc tggacaagct ccacagagct 660
gcctgggtgg gtaaaatccc caaaaaggat ctacactctg tgcctcagg caackaggtg 720
aacaagargg caaagcaaaa gaggactgct ctacactctg cctctgcaaa tgggaattca 780
gaagtagtaa aactctgctt ggacagacga tgtcaactta atgtctctga caaaaaaag 840
aggacagctc tgaayaaagg cgtacaatgc caggaaagat aatgttgctt aatgttgctg 900
ggacatggca ctgatccaaa tattccagat gagtatggaa ataccactct rcaataygt 960
rtctayaatg aagataaatt aatggcaaaa gcaactgctc tataygtgc tgatatcgaa 1020
tcasaaacaa agcatggcct ccaaccactg ytaattggtr tactagagca aaaaacagca 1080
gtctgaaatt ttttaayaa gaaaagagcg aattlaaat gcrctggata gatatggag 1140
racgtctctc atactgtgtg tatgttggg atcagcaagt atagtcagcc yctactctga 1200
gcaaaatttt galgtatctt ctcaagatct ggaagagcgg ccagagagta tgcctgtttct 1260
agtcatctc atgtaatgtt ccagttactt tctgactaca aagaaaaaca gatgttaaaa 1320
atctctctgt aaaaacgcaa tccagaacaa gactlaaagc tgacatcaga ggaagagta 1380
caaaagctta aaggaagtga aaaaacagag ccagaggaat ggaactttt aatttaaac 1440
ttttgggtta atgttttttt tttttgctt aataatatta gatagtcaca aatgaattwa 1500

```

WO 01/90152

PCT/US01/16776

85

```

ccatagagac taggotttga gaataaatg attottttt taagaatott ttggctagga 1560
gcgtgtctct acgctgttaa ttccagcaac ttgaagagct gaagtggaac gatccagaga 1620
tcaggagato gagaccatcc iggttaacgc ggtgaacccc catctctact aaaaatacas 1680
aaaotttagct aggtgtgggt gccgggtgct gtagtcccaag ctactcagga rgttgaggca 1740
ggagaatggc atgaacccgg gaggtggagg ttgaagtga ccagatcccg ccactacact 1800
ccagcctggg tgacagagca agactctgtc toaaaaaaa aaaaaaaaaa aaa 1853

```

```

<210> 296
<211> 2184
<212> DNA
<213> Homo sapien

```

```

<400> 296
ggcccgagaa ttaaaacccct cagcaaaaca ggcatagaaq ggaataacct taagaataa 60
aasaccacct atgacaagcc cagcagcaac ataatactta atggggaaaa gttagaagca 120
tttctctctga gaantgaaac aataaalaca aggatgctgg attttgtcaa atgcttttct 180
tgtgtctgtt gagatgotta tgttgacttg ctcttaattc tgtttatgtg attatcaact 240
ttattgactt gctgtgttta gaccggaaga gctgggggtg ttctcaggag ccacccgtgtg 300
ctggggcagc ttccggataa cttgaggctg catcaotggg gaagaasacc aytccgtgtc 360
gtggcgctga tggctgagga cagagcttca gttgtgcttc tctcgactg gcttctctgg 420
ggagtctctc ctccatagtt catccatagc gctccagagg aaatttatac tatttcttga 480
tggatgaaga gtattacgtt glgcagatat actgcagtggt ctccatctct tgatgtgtga 540
ttgggttagt tccaccatgt tgcgcagat gacatgattt cagtaoctgt gctgtggtga 600
aaagtgtttt ttgtgaaatg galatttgtg ttcttggttc tcatctctgt tgggtggaca 660
gctttctcca cctgtcgggc agtgacctgc ttccagaag tttyatggct gaggagata 720
ccatctgaca tgcactcttc atttctctca ttcttctct cctggatgga cagggggagc 780
ggcaagagca acgtgggccc ttctggagac caaaccgact cctctgtgaa gacgttgtg 840
agcaagaggt gcaagtgtgt ctgcaactgc ttccctgct gcaagggagc ggcagagca 900
acgtgtgctc ttggggagac taagatgaaa ggccttctat gnatccagc taccacgtcc 960
atgggagaga tctggagcag ctccacagag ctgctgtgtg ggttaagtc cccagaagag 1020
ctctcatctc catgctcagg gacacggatg tgaacaagag ggacaagcaa aagaggagtg 1080
ctctcatctc ggcctctgcc aatgggaatt cagaagtagt aaactctgt ctggacagac 1140
gatgtcaact taatgtcttt gacaaacaaa agaggcagc tctgcaaaag gccgtacat 1200
gocaggagaa tgaatgtgcy ttaattgtgc tggacatgg cactgatcca aatattccag 1260
atgagtatgg aaataccact ctacactatg ctgtclacaa tgaagataaa tcaatggcca 1320
aagcaatgct attatccgtt gctgatatcg aatcaaaaa caagcctggc ctccacccac 1380
tgcactctgg tatcatctcg caaaacagc aagtgtgaa atttttaac agaaaaaag 1440
cgaaattcaa tgcctggat agataaggaa gaactgctct catctctgct gtaggtgtgtg 1500
gatcagcaag tatagtcagc cctctacttg agcaaatgt tgaatgtatc tctcaagatc 1560
tggaaagagc gcaagagagt atgctgtttc tagtcatcat catgtaahit gcnagttacl 1620
ttccactacc aaagaaaaac agatgttaaa aatctctctt gaaaacagca atccagaaca 1680
agacttaag ctgacatcag aggaagatc acaaaggctt aaaggaagtg aaacagaca 1740
gcaagagcca tggaaacttt taattttaa cttttgtttt aatgtttttt tttttgct 1800
taataatatt agatagctcc aaatgaatw acctatgaga ctaggctttg agaataata 1860
gattcttttt ttaagaatct ttgtgttagg agcgggtgct cagcctctga attccagac 1920
cttgagagcc ttgagtgagg agatcacgag atcaggagat cagagaccatc ctggctaaca 1980
cgtgaanacc ccatctctac taaaaataca aaacttagc tgggtgtggt ggcgggtgct 2040
tgtagtccca gclactcagg arctcgagc agagaatgg catgaacccg ggaggtggag 2100
gttgagtgaa gcaagatccc gcaactcac tccagcctgg gtgacagagc aagactctgt 2160
ctcaaaaaaa aaaaaaaa aaad 2184

```

```

<210> 297
<211> 1855
<212> DNA
<213> Homo sapien

<220>
<221> mi_sc_Feature

```

WO 01/90152

PCT/US01/16776

86

<222> {1...[1855]}
 <223> n = A,T,C or G

<400> 297

```

tgcacgcctc ggcacgtgtc tgtgacacgt acactgaacg cccctgagat gtgcacgctc 60
cgcgcgcacg tgcacgcgc ggcacgggct tggctggcct gtaacggctt gacgcgcac 120
gcgcgcgcgc cacaacgtc agactggcct gtaacggctt gcaggcgac gcgcacgcgc 180
cgtaacggct tggctgcctc gtaacggcct gcaactgcac gctgcacgcg cgttaacggc 240
ttgctggcgc tgtagcgcct tggcttgcc ttcgatttct tgcctggctk ggcgttgkty 300
tcttgagattg acgcltctc cttggatkgc ctttctctcc ttggatkgac gtttcytyty 360
tcgggttctc ttgctgaac tgccttttct tctgctgggt ttggcattcc ttgggggtgg 420
gclgggtgtt ttctcggcg ggggttgccc ttcttggggg ggcgtggggc cgcgcgcacg 480
gggcgtggcg ttcccccggg tgggtgtggg ttcttccggg gtgggtggg cgtgtgtggg 540
atccccctgc tgggtgtggc agggattgac tttttctctc aaacagattg gaaacccggc 600
gttaactgct agttggtgaa acggttggc agacgcgcat tgcctgact actgtttctc 660
ctgctgtgta aaagcagatg gtggttgagg ttgattcaat gccgctgct tcttctgtga 720
agaagccatt tggctcagcg agcaagatgg gcaagtgggt cgcacactgct tccccctgct 780
cagggggggc ggcacagagc acgtgggcac ttctgggac cacaacgact cctctgtgaa 840
gacgctggcg agcagaggtt gcaactggcg ctgcacactg ctccccctgc tgcagggggc 900
ggcacaagcg aacgtgkcg cttgggggga ctacgatgac agcgccttca tggakcccg 960
gtccccctgc cttggagagc atctgggaca gctccacaga gctgcctggt ggggttaagt 1020
ccccagaaag gatctcctg tcatgtctag ggcactgag gtgacaaga rggacaagca 1080
aaagagagct gctctacatc tggcctctgc caatgggcat tccgagctgt taaaactcgt 1140
gctggagaga gctgtgaac ttaactctc tgcacacaaa agagggagag cctctgacaa 1200
ggcgtgacaa tgcagggag atgaatgctc gttcaatgtg ctggaaactg gcaatgctc 1260
aataattcca gatgaatgt gaaataccac tctaacatct gctgtotac atgagatga 1320
attaatggcc aaagcactgc ttttatcagc tctgtatata gaaataaaaa acaggtata 1380
gatctactaa ttttatcttc aaaaactgc aatgcattca ttttaacatt gacgtgtgta 1440
agggccactc ttcogtattt ggaagctcaa gataacttg aatgaaaata ttttgaaagt 1500
ccctaattat ctaagacttt attttaanta ttgtatttt caaagagca ttagagggta 1560
cagttttttt ttttkaatg cacttctggt aaatactttt gttgaaaca ctgaatttgt 1620
aaaaggtaat acttaactat ttcaacttlt tccctcctag gatttttttc cctaatgaa 1680
tgtaaagtgg caaatllgc octgaatag gttttacatg aaaaactcaa gaaaagttaa 1740
acatgtttca gtgaatagc atcctgctcc ttggcnaat tcttaaaaaa cagtaataga 1800
tgcaggtgga tgcgcctgct agtggcaagg ttaagatat tctgtatctc gtgcc 1855

```

<210> 298

<211> 1059

<212> DNA

<213> Homo sapien

<400> 298

```

gnaacgtggg cactttctga gacacacag actcctctgt gaagacgctt gggagcaaga 60
ggtgcaagtg gtgctgcca ctgcttcccc tgcctgacgg gagcggaag agcaactgg 120
gcgcttgagg agactmcat gacagggcct tcatggagcc caggtaccac gtcctgtgag 180
aagatcttga caagctccac agagctgccc tgggtgggta aagtccccag aaagatctc 240
atcgtcatgc tcaagggaac tgaagtgaac aagarggaca agcaaaagag gactgctcta 300
cactctggct ctgcccattg gaattcagaa ttagtaaaac tctgtctgga cagaagatgt 360
caacttaatg tcttgacaa caaaaagagg acagcttga yaaagctcgt aaaaagcag 420
gaagatgaa: gtgcgttaat gttgcgggaa catggcaatg atcaaatat tccagtgag 480
tatggaaata ccaactctca ctaygchrtc tayaatgagc ataaattaat ggcacaaga 540
ctgctcttat syggtgctga tatgaatca aaaaacaagg tataatctca ctattttat 600
cttcaaaata ctgaatgca ttcatittaa cattgacgtg tgaagggoc agtcttctc 660
atttggagcg tcaagctaan cttgaatgaa aatattttga aatgcctaa ttatctaga 720
ctttatttta aatattgtha ttttcaaga agcattagag ggtacagttt tttttttta 780
aatgaacttc tggtaatac tttgttgaa aacactgcat ttgtaaaagg taatacttac 840
tatitttcaa tttttccctc ctgagatttt ttccctctaa tgaatgtaag atggcaaat 900
ttgccttgaa ataggtttta catgaaact ccaaganaag ttaaacatgt ttcagtgaat 960

```

WO 01/90152

PCT/US01/16776

87

agagatcttg cctctttggc aagtttctaa aaaaacagtaa tagataacag gtgatgcgcc 1020
 cgtcagtgcc aaggttttaag ttatttttga tctcgtgcc 1059

<210> 299
 <211> 329
 <212> PRT
 <213> Homo sapien

<400> 299
 Met Asp Ile Val Val Ser Gly Ser His Pro Leu Trp Val Asp Ser Phe
 1 5 10 15
 Leu His Leu Ala Gly Ser Asp Leu Leu Ser Arg Ser Leu Met Ala Glu
 20 25 30
 Glu Tyr Thr Ile Val His Ala Ser Phe Ile Ser Cys Ile Ser Ser Ser
 35 40 45
 Leu Asp Gly Gln Gly Glu Arg Gln Glu Gln Arg Gly His Phe Trp Arg
 50 55 60
 Pro Gln Arg Leu Leu Cys Glu Asp Ala Trp Glu Gln Glu Val Gln Val
 65 70 75 80
 Val Leu Pro Leu Leu Pro Leu Leu Gln Gly Ser Gly Lys Ser Asn Val
 85 90 95
 Val Ala Trp Gly Asp Tyr Asp Asp Ser Ala Phe Met Asp Pro Arg Tyr
 100 105 110
 His Val His Gly Glu Asp Leu Asp Lys Leu His Arg Ala Ala Trp Trp
 115 120 125
 Gly Lys Val Pro Arg Lys Asp Leu Ile Val Met Leu Arg Asp Thr Asp
 130 135 140
 Val Asn Lys Arg Asp Lys Gln Lys Arg Thr Ala Leu His Leu Ala Ser
 145 150 155 160
 Ala Asn Gly Asn Ser Glu Val Val Lys Leu Val Leu Asp Arg Arg Cys
 165 170 175
 Gln Leu Asn Val Leu Asp Asn Lys Lys Arg Thr Ala Leu Thr Lys Ala
 180 185 190
 Val Gln Cys Gln Glu Asp Glu Cys Ala Leu Met Leu Leu Gln His Gly
 195 200 205
 Thr Asp Pro Asn Ile Pro Asp Glu Tyr Gly Asn Thr Thr Leu His Tyr
 210 215 220
 Ala Val Tyr Asn Glu Asp Lys Leu Met Ala Lys Ala Leu Leu Leu Tyr
 225 230 235 240
 Gly Ala Asp Ile Glu Ser Lys Asn Lys His Gly Leu Thr Pro Leu Leu
 245 250 255
 Leu Gly Ile His Glu Gln Lys Gln Gln Val Val Lys Phe Leu Ile Lys
 260 265 270
 Lys Lys Ala Asn Leu Asn Ala Leu Asp Arg Tyr Gly Arg Thr Ala Leu
 275 280 285
 Ile Leu Ala Val Cys Cys Gly Ser Ala Ser Ile Val Ser Pro Leu Leu
 290 295 300
 Glu Gln Asn Val Asp Val Ser Ser Gln Asp Leu Glu Arg Arg Pro Glu
 305 310 315 320
 Ser Met Leu Phe Leu Val Ile Ile Met
 325

<210> 300
 <211> 148
 <212> PRT
 <213> Homo sapien

<220>

WO 01/40152

PCT/US01/16776

88

<221> VARIANT
 <222> (1)...(146)
 <223> Xaa = Any Amino Acid
 <400> 300
 Met Thr Xaa Pro Ser Trp Ser Pro Gly Thr Thr Ser Val Glu Lys Ile
 1 5 10 15
 Trp Thr Ser Ser Thr Glu Leu Pro Trp Trp Gly Lys Val Pro Arg Lys
 20 25 30
 Asp Leu Ile Val Met Leu Arg Asp Thr Asp Val Asn Lys Xaa Asp Lys
 35 40 45
 Glu Lys Arg Thr Ala Leu His Leu Ala Ser Ala Asn Gly Asn Ser Glu
 50 55 60
 Val Val Lys Leu Xaa Leu Asp Arg Arg Cys Glu Leu Asn Val Leu Asp
 65 70 75 80
 Asn Lys Lys Arg Thr Ala Leu Xaa Lys Ala Val Glu Cys Glu Glu Asp
 85 90 95
 Glu Cys Ala Leu Met Leu Leu Glu His Gly Thr Asp Pro Asn Ile Pro
 100 105 110
 Asp Glu Tyr Gly Asn Thr Thr Leu His Tyr Ala Xaa Tyr Asn Glu Asp
 115 120 125
 Lys Leu Met Ala Lys Ala Leu Leu Leu Tyr Gly Ala Asp Ile Glu Ser
 130 135 140
 Lys Asn Lys Val
 145
 <210> 301
 <211> 1155
 <212> DNA
 <213> Homo sapien
 <400> 301
 atgtgtgttg aggttgattc catgcgggct gctcttcttg tgaagaagcc atttggcttc 60
 aggaagcaaga tgggcaagtg gtgtgtcgtt tcttccctct gctgcaggga ggcgggcaag 120
 agcaacgtgg gcaattctgg agaccacgac gactctgcta tgaagacact caggagcaag 180
 atgggcaagt ggtgcggcca ctgtcttccc tctgtcaggg ggaglggcaa gagcaacgtg 240
 ggactctctt gaaacacaga cgaactctgt atgaagacac tcaaggacaa gatgggcaag 300
 tgggtgtgcc actgtctccc ctgtctcagg gggagcgcca agagaaggtt gggcggttgg 360
 ggaactatct atgaagatgc ctctcatgag acaaggtacc aagtccttgg agaaatcttg 420
 gacaagctcc acagagctgc ctggtggggt aagtcacca gaaaggtatc catgtctatg 480
 ctacaggaca ctgacgtgaa caagaaggac aagcaaaaga ggaactgctc acatctggcc 540
 tctgccaatg ggaattcaga agtagtaaaa ctctgtctgg acagacgatg tcaacttcat 600
 gtctttgaca acaaaaagag gacagctctg ataaagggcg tacaatgcca ggaagatgaa 660
 tgtggtttaa tgttcttgga acatggcact gatccaaata ttccagatga gtatggaaat 720
 acaactctgc actaagctat ctataatgaa gataaantaa tggccaaago actgtcttta 780
 tatgtgtctg atatogaato aaaaacacag catggcctca caaacctgtt ncttgggtga 840
 catgagcaaa aacagcaagt cgtgaatttt ttaatcaaga aaaaagcgaa tttaaattga 900
 ctggatagat atggaaggac tgcctccta attgtcttat gttgtggata agcaagtata 960
 gtaagctctc taattgagaa aaatatgtat gtatcttctc aagatctata tggacagacg 1020
 gcaagagagt atgtgtcttc tagtaaat catgttaatt gccagttact ttctgactac 1080
 aaagaaaaac agatgtctaa atctctctct gaataacgca atccagaaaa tgtctcaaga 1140
 accagaataa aatcaa 1155
 <210> 302
 <211> 2000
 <212> DNA
 <213> Homo sapien

WO 01/00152

PCT/US01/16776

89

<400> 302
 atggtggttg aggttgatgc catgccggct gctctctctg tgaagaagcc atttggcttc 60
 agggagcaaga tgggcaagtg gtgtgcgcgt tcttccctct gctgcaggga gaggggcaag 120
 agcaacgttg gcaactcttg agaccacgac gaactctgta tgaagacact caggagcaag 180
 atgggcaagt ggtgcgcga ctgcttccct tctgcaggg ggaagtggca gagcaactg 240
 ggctgctctg gagaccaga cgaactctgt atgaagacac tcaagacaa gatgggcaag 300
 tgggtctctg actgcttccg ctgctgcagg gggagcgga agagcaaggt gggcgtctg 360
 ggaagctacg atgcagctgc ctctcaggag cccaggtacc acgtccgttg acaagatctg 420
 gacagctccg acagagctgc ctgctggggg aaagtcccca gaaaggtct catctctatg 480
 ctcaaggaca ctgacgtgaa caagaaggac agcaaaaaga ggaatgctct acatctggcc 540
 tctgccaatg ggaattccga agtgaataaa ctctcgttg agagagatg tcaacttaai 600
 tctcttgaca acaaaaagag gaagctctga ataaagcccg taacttgca ggaagatgaa 660
 tgtgcgttaa tgttgcgtga acatggcaat gctccaaata ttccagatga glatggaaat 720
 acaactctgc actacgtat ctataatgaa gataaattaa tggccaagc actgctctta 780
 tatggctgtg atatcgatc aaaaaaacag catggcctca caccactgtt acttgggtga 840
 catgagcaaa aacagcaagt cgtgaattt ttaatacaga aaaaagcga tttaaatgca 900
 ctggatagat atggaaggac tgcctcatal ctgctgtat gttgtggat agcaagtata 960
 gtcagccttc tacttgagca aaatattgat gtatctcttc agatctcttc tggacagacg 1020
 gcaagaaagt atgctgttct tagtcatcat catgtaattt gccagttact tctgactac 1080
 aaagaaaac agatgctaaa aatctctctt gaaaacagca atccagaca agacttaag 1140
 ctgacatcag aggaagagtc acaaaagttc aaaggcagtg aaatagcca gccagagaaa 1200
 atgtctcag atccagaact aaataaggat ggtgatagag aggttgaaag agaaatgaa 1260
 agcatgaaa gtaataatg gggatctcta gaacactga caactggtgt cactgctggc 1320
 aatggtgata atgattcaat tctccaaagg aagagcgaat cactgcaatt 1380
 cctgacacgc aaagtgaaga gtatcacaga atttgcaat tagtttctga ctacaaagaa 1440
 aaacagatgc caaaatactc ttctgaacac agcaaacccg aacagactt aaagtctgca 1500
 tcagaggaag agtcccaag gcttgaggc agtgaataat gccagccaga gctagaatat 1560
 tttatggcta ccgaagaat gaagagcac ggaagtact atgtcggat cccagaaaac 1620
 ctgactaatg gtgccactgc tggcaatgt gatgatgat taattctctc aaggaagagc 1680
 agaacacctg aagccagca atttctctg actgagaaat aagatatac cagtgaagaa 1740
 caaanatgata ctacagaagca atttgtgaa gaacagaaca ctggactatt acagatgag 1800
 attctgtatc atgaagaaaa gcagatgaaa gtggtgaaa aaatgaattc tgaagttctc 1860
 cttaattgta agaaagaaaa agacatcttg catgaaaaa gtacgttgag ggaagaaatt 1920
 gccatgttaa gactggagct agcaccaatg aaacatcaga gccagotaaa aaaaaaaa 1980
 aaaaaaaa aaaaaaaa 2000

<210> 303
 <211> 2040
 <212> DNA
 <213> Homo sapien

<400> 303
 atggtggttg aggttgatgc catgccggct gctctctctg tgaagaagcc atttggcttc 60
 agggagcaaga tgggcaagtg gtgtgcgcgt tcttccctct gctgcaggga gaggggcaag 120
 agcaacgttg gcaactcttg agaccacgac gaactctgta tgaagacact caggagcaag 180
 atgggcaagt ggtgcgcga ctgcttccct tctgcaggg ggaagtggca gagcaactg 240
 ggctgctctg gagaccaga cgaactctgt atgaagacac tcaagacaa gatgggcaag 300
 tgggtctctg actgcttccg ctgctgcagg gggagcgga agagcaaggt gggcgtctg 360
 ggaagctacg atgcagctgc ctctcaggag cccaggtacc acgtccgttg acaagatctg 420
 gacagctccg acagagctgc ctgctggggg aaagtcccca gaaaggtct acatctggcc 540
 ctcaaggaca ctgacgtgaa caagaaggac agcaaaaaga ggaatgctct acatctggcc 600
 tctgccaatg ggaattccga agtgaataaa ctctcgttg acagcagatg tcaacttaai 660
 tgtgcgttaa tgttgcgtga acatggcaat gctccaaata ttccagatga glatggaaat 720
 acaactctgc actacgtat ctataatgaa gataaattaa tggccaagc actgctctta 780
 tatggctgtg atatcgatc aaaaaaacag catggcctca caccactgtt acttgggtga 840
 catgagcaaa aacagcaagt cgtgaattt ttaatacaga aaaaagcga tttaaatgca 900
 ctggatagat atggaaggac tgcctcatal ctgctgtat gttgtggat agcaagtata 960

WO 01/90152

PCT/US01/16776

90

```

gtcagccctt tacttgagca aaatattgat gtatctcttc aagatctatc tggcagagag 1020
gcccagagag atcctgtttc tagtctctat catgtatatt gcaggttact ttctgactac 1080
aaagaaaaac agatgctaaa aatctcttct gaaaacagca atccagagca agacttnaag 1140
ctgacatcag aggaagagto ecaaggttc aaagggcagtg aaattagcca gccagagaaa 1200
atgtctcagag aaccagaaat aaataaggtt ggtatagag aggttcagag aaaaatgaag 1260
aagcatgaaa gtaataatgt gggattacta gaaaacctga ctattggtgt caatgctggo 1320
aatggtgata atggattaat tcttcaagag aagagagaaa caactgaaa bcagcaattt 1380
cctgcaacag aagtgaaaga gtatcagaga atttgogaat tagttcttga ctacaagaaa 1440
aaacagatgc caaataactc ttctgaaaac agcaacccag acaagactt aaagctgaca 1500
tcagagggag agtcacaaag gcttgagggc agtgaaatag gccagccaga gaaaagatct 1560
caagagccag aaataaataa ggtatgtgat agagagctag aaatttttat ggctatogaa 1620
gaatgagaga agcagggagag tactcatgtc ggattccag aaacctgac taalggtgcc 1680
aotgolgua atgtgtgatga tggattaat ctccagaga agagcagaa acctgaaagc 1740
cagcaatttc ctgacactga gaatgaagag tatcagagtg acgaaacaaa tgatcctcag 1800
aagcaatttt gtgaagacaa gaacctgga atattacac atgagattct gattcatgaa 1860
gaaaagcaga tagaagtgtg tgaataaatg aattctgag ttctcttag ttgtaagaaa 1920
gaaaagacaa tcttgcatga aataglaag ttgcgggaag aaattgcat gctaaagctg 1980
gagctagaca caatgaacaa tcagagccag ctaaaaaaa aaaaaaana aaaaaaana 2040

```

<210> 304
 <211> 384
 <212> PRT
 <213> Homo sapien

<400> 304

```

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

```

WO 01/90152

PCT/US01/16776

91

260 265 270
 Leu Thr Pro Leu Leu Leu Gly Val His Glu Gln Lys Gln Gln Val Val
 275 280 285
 Lys Phe Leu Ile Lys Lys Lys Ala Asn Leu Asn Ala Leu Asp Arg Tyr
 290 295 300
 Gly Arg Thr Ala Leu Ile Leu Ala Val Cys Cys Gly Ser Ala Ser Ile
 305 310 315 320
 Val Ser Leu Leu Leu Glu Gln Asn Ile Asp Val Ser Ser Gln Asp Leu
 325 330 335
 Ser Gly Gln Thr Ala Arg Glu Tyr Ala Val Ser Ser His His Val
 340 345 350
 Ile Cys Gln Leu Leu Ser Asp Tyr Lys Glu Lys Gln Met Leu Lys Ile
 355 360 365
 Ser Ser Glu Asn Ser Asn Pro Glu Asn Val Ser Arg Thr Arg Asn Lys
 370 375 380

 <210> 305
 <211> 656
 <212> FRT
 <213> Homo sapien

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

WO 01/90152

PCT/US01/16776

92

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

 <210> 306
 <211> 671
 <212> PRT
 <213> Homo sapien

 <400> 306
 Met Val Val Glu Val Asp Ser Met Pro Ala Ala Ser Ser Val Lys Lys
 1 5 10 15
 Pro Phe Gly Leu Arg Ser Lys Met Gly Lys Trp Cys Cys Arg Cys Phe
 20 25 30
 Pro Cys Cys Arg Glu Ser Gly Lys Ser Asn Val Gly Thr Ser Gly Asp

WO 01/00152

PCT/US01/6776

93

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

WO 01/90152

PCT/US01/16776

94

500 505 510
 Asn Gly Gln Pro Glu Lys Arg Ser Gln Glu Pro Glu Ile Asn Lys Asp
 515 520 525
 Gly Asp Arg Glu Leu Glu Asn Phe Met Ala Ile Glu Glu Met Lys Lys
 530 535 540
 His Gly Ser Thr His Val Gly Phe Pro Glu Asn Leu Thr Asn Gly Ala
 545 550 555 560
 Thr Ala Gly Asn Gly Asp Asp Gly Leu Ile Pro Pro Arg Lys Ser Arg
 565 570 575
 Thr Pro Glu Ser Gln Gln Phe Pro Asp Thr Glu Asn Glu Glu Tyr His
 580 585 590
 Ser Asp Glu Gln Asn Asp Thr Gln Lys Gln Phe Cys Glu Glu Gln Asn
 595 600 605
 Thr Gly Ile Leu His Asp Glu Ile Leu Ile His Glu Glu Lys Gln Ile
 610 615 620
 Glu Val Val Glu Lys Met Asn Ser Glu Leu Ser Leu Ser Cys Lys Lys
 625 630 635 640
 Glu Lys Asp Ile Leu His Glu Asn Ser Thr Leu Arg Glu Glu Ile Ala
 645 650 655
 Met Leu Arg Leu Glu Leu Asp Thr Met Lys His Gln Ser Gln Leu
 660 665 670

<210> 307
 <211> 800
 <212> DNA
 <213> Homo sapien

<400> 307
 atkagottcc gcttctgaca acactagaga tccctccctt cctctagggt atggccctcc 60
 aottcatttt tggatcataa catctttata ggcaggggt annatcccaa tactaacagg 120
 aqaatgotta ggaactotnac aggtttttga gaagtgtttg gtaagggccaa ctcaatccaa 180
 tttttcttgg tctctottgt ggtctaggag gacagggcaag ggtgcagatt ttcaagaaag 240
 catcagtaag ggcactataa tccgaccttc ctggttcttc cttgtggtct gggaggaaaa 300
 ctagtgtttc tgttctgtgt tcaatgagca caactattcc gatcagcagg gtccagggac 360
 cactgcaggt tcttgggcag ggggagaaa aaacaaaacc aasaccatgg gorgttttgt 420
 ctltccagatg ggaacacact aggcataaac aggtccacct ttgaaatgca tctaaagcca 480
 atggagccaa ttgacccac aaaccttga aaagctgtg ctcattttt ttgacactat 540
 ggtttggccc caacattctc tctctgaggg ggaatatgg ccaactgagg gaagtaaga 600
 ttacaaact atctcgaagc ttgaccttt ctgtaagagg gaagcaaat gaagtgaat 660
 aacttatgtc caagttttt ttcatgaa ggaatata ctatgcaag ctgaaattt 720
 aactcccaaa ggaagacctc tcaagttacc cccatctct agctcccta tagctccct 780
 tctattagt gataagctc 800

<210> 308
 <211> 162
 <212> PPT
 <213> Homo sapien

<220>
 <221> VAR[AMT]
 <222> (1)...(102)
 <223> Xaa = Any Amino Acid
 <400> 308
 Met Gly Xaa Phe Val Phe Gln Met Gly Asn Thr Gln Ala Ser Thr Gly
 1 5 10 15
 Ser Pro Leu Lys Cys Ile Leu Ser Gln Trp Asp Lys Phe Asp Pro Gln
 20 25 30

WO 01/90152

PCT/US01/16776

95

Thr Leu Glu Lys Glu Val Ala His Phe Phe Cys Thr Met Ala Trp Pro
 35 40 45
 Gln His Ser Leu Ser Asp Gly Glu Lys Trp Pro Pro Glu Gly Ser Thr
 50 55 60
 Asp Tyr Asn Thr Ile Leu Gln Leu Asp Leu Phe Cys Lys Arg Glu Gly
 65 70 75 80
 Lys Trp Ser Glu Ile Pro Tyr Val Gln Ala Phe Phe Ser Leu Lys Glu
 85 90 95
 Asn Thr Leu Cys Lys Ala
 100

<210> 309
 <211> 9
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Made in the lab

<400> 309
 Leu Met Ala Glu Glu Tyr Thr Ile Val
 1 5

<210> 310
 <211> 9
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Made in the lab

<400> 310
 Lys Leu Met Ala Lys Ala Leu Leu Leu
 1 5

<210> 311
 <211> 9
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Made in the lab

<400> 311
 Gly Leu Thr Pro Leu Leu Leu Gly Ile
 1 5

<210> 312
 <211> 10
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Made in the lab

<400> 312
 Lys Leu Val Leu Asp Arg Arg Cys Gln Leu
 1 5 10

WO 01/90152

PCT/US01/16776

96

<210> 313
 <211> 1852
 <212> DNA
 <213> Homo sapiens

<400> 313
 ggcacgagaa ttaaaaccc cagcaaaaca ggcatagaag ggcataacct taaagtaata 60
 aaaaacaccc atgacaagcc caccgcccaac ataatactaa atgggggaaaa gttagaagca 120
 ttccctctga gaactgcac aataaatata aggatgctgg attttgcata atgccttttc 180
 tctgctctgt gagatgctta tggactcttg ctttcaatc tgtttatgtg attatcaaat 240
 ttattgaact gctctgttta gacggaaaga gctggggtgt ttctcaaggc caacgctgtg 300
 ctccggcagc ttccggataa ctggaggctg cctcaactgg gaagaacac acytctctcc 360
 gtggcgctga tggctgagga cagagcttca gtgtggcttc tctgcgactg gcttctctgg 420
 ggcattcttc cttcatagtt catccatag gtccagagg aaaaattat tatittgtta 480
 tggatgaaga gtattacgtt gtccagatat actgagtg cttcatctc tgaatgtga 540
 ttgggtagtg tccaccatgt tcccgagat gacatgatt cagtaacctg gctctgctga 600
 aaagtgtttg tttgtgaat gatattgtgg ttctctgata tcatctctg tgggtggaca 660
 gttttctcua ccttgctgga agtgaacctg tctccagaag ttctgtggct gaggagata 720
 ccactgtcca tgcactcttc aittctgca ttctctctc cctgagtgga caggggagc 780
 ggcacagaca agtggggcac ttctgagaa cacaacgact cctctgtgaa gacgcttggg 840
 agcaagaggt gaaagtgtgt ctgcacatgc ttccctctgt gcaaggggag cgcacagagc 900
 acgtgtgtgt ctgtgggaga ctacagtagc agcgcctcca tggatccag gtaccagtc 960
 catggaag atctggacaa gctccacaga gctgctgtg ggggtaagt ccccgagag 1020
 gatctcctg tcatgctcag gacacaggat gtgaacaga ggcacacaga aaaggagct 1080
 gctctacatc ttgctctgca caatgggaat tcaagagtg laaaactcgt gctggcaga 1140
 cgaatgaac ttaattctct tgacacacaa aagaggagc ccttgacaaa ggcgtacaa 1200
 tgcacagaa atgaattgtg gttaatgttg ctggacatg gcaatgctc aaatttcca 1260
 gatgagtatg gaaataccac tctacactat gctgtctaca atgaagata ataatggc 1320
 aaagcactgc tcttatacgg tctgtatc ataatcaaaa acaagcatgg cctcacacca 1380
 ctgctacttg gtatacctga gcaaaaacag caagtgtgga aatttttaat caagaaaaa 1440
 ggcattttaa atggcgctga tagatatgga agaactgtc tcatacttgc tgtatgtgt 1500
 ggcacagcaa gtatgtcag cctctactt gacacaaat ttgatgtatc ttctcaaat 1560
 ctggaaagac ggcacagag tatgtctgtt ctagtcaatc tcatgtatc tgcagctac 1620
 ttctcacta caaagaaaaa cagatgttaa aattctctc tgaacacag aatccagaa 1680
 aagactttaa gctacacata gagggaagt caaagaggt taaaggaggt gaaacagc 1740
 agcagagct agaagattta tggctactga aagagatga agaacagga agtactcat 1800
 tgggtctccc agaaaacctg actaacggcg cgcgtgctg caatggtgat ga 1852

<210> 314
 <211> 879
 <212> DNA
 <213> Homo sapiens

<400> 314
 atgcattctt catttctgc atttcttct cctggatgg acagggggag cggcaagagc 60
 aacgtgggca ctctggaga ccacacgac tctctgtga agagcttgg gagcaagagg 120
 tgaagtggt gctgacatc cttccctgc tgcagggga gggcaagag caacgtggc 180
 gcttggggag actacagca cagcgcttc atgatatca ggtacacgt ccatgggaa 240
 gctcgggaca agctccacag agctcctgtg tgggttaag tcccgagaa ggtctctc 300
 gtctgtctca ggcacagga tgtgaacag agggacacg aaaaaggagc tgcctcatc 360
 ctggcctctg ccaatgggaa ttcagagta gtaaacctg tgcgtgacg acgatgtca 420
 ctttaagtgc ttgacacaa aagaggacg gctctgaca aggcgtaca atgccaggaa 480
 gatgaatgtg cgttaattt gctggacat ggcactgac caaatcttc agatgagat 540
 ggaatacca ctctacacta tgcgtctc aatgaagata aattaatgg caaagcactg 600
 cttattacg gtgctgat ataatcaaaa aacaagcatg gctctacac actgtactt 660
 gctatacatg agcaaaaaca gcaagtgtg aaattttta tcaagaaaa agcgatcta 720
 aatggctg atagatatg aagaactgt ctcatactg ctgtatgtg tggatcaga 780

WO 01/00152

PCT/US01/16776

97

agtatatgctcc gccctctact tgcgcacaat gttgatgtat ctctcaaga tctggaaaga 940
 cggccagaga gtaigtgtgtt tctagtcttc atcatgtaa 879

<210> 315

<211> 292

<212> PRI

<213> Homo sapiens

<400> 315

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

WO 01/00152

PCT/US01/16776

98

260 265 270

Val Ser Ser Gln Asp Leu Glc Arg Arg Pro Gln Ser Met Leu Phe Leu
275 280 285

Val Ile Ile Met
290

<210> 316
<211> 584
<212> DNA
<213> Homo sapiens

<400> 316
agttgggcca aattccocct cccctacagc ttgaagggga cataaccaat agcctggggt 60
ttttttgtgg tcccttgagg atttctttgc ttattttctt ctgggtgggg gtgattagay 120
gaggcttacc actaatagga agggagagct tagggaggct aggatatggg ggtaagctga 180
gaggtctctcc tgtgggatgt aaattccagc cttgcatagc tgtattctcc ttcaatgaaa 240
agaaagctty gacataaggt atttccatcc atttgccttc cctcttaccg aaaggtcaaa 300
gctgcaggat agtatgttaa tctgtacttc cctcaggtgg ccatttttcc ccatcagaga 360
gggaattgtt ggcccaagcc atagtgcaga aaaaaaattg agccacctct ttttcagggg 420
tttggggttc aaattctgct cctttgggca ggatgcaatt caaaggtgag cctgttgatg 480
cctgaagtgt tcccatctga aagacaaaac tgcctatggt ttgttttgt ttgtttcttc 540
ccctgcacca agaatatna aactctyag caacaaacta aaaa 584

<210> 317
<211> 829
<212> DNA
<213> Homo sapiens

<400> 317
attagcttcc gcttcigaca acactagaga tccctccocct ccttcagggt atggccctcc 60
acttcatatt tggtaacata catcttata ggcaggggtt aaatcccaa tactaacagg 120
agaaatgctta ggaactaac aggtttttga gaattgtttg gtaaggggca ctaaatcaaa 180
ttttctctgg tccctcttgt ggtttggagc gacaggcaag ggtgcagatt ttaagaattg 240
catcagtaag ggcacataaa tccgacatcc ctgcttcttc cttgtgtctt gggagggaaa 300
ctagtgcttc tgttctgtg tcaatgagca caactattcc gctcagcagg gtccaggagc 360
cactcaggtt tcttgggcaz ggggagaaac aaacaaacc aaacaaatg gcxgttttgt 420
ctttcagctg ggaacacatc aggcataaac aggcataact ttgaatgca tcttaagcca 480
atgggacaaa tttagaccac aaacccctgga aaagagggtg gctcattttt ttgtcactat 540
ggcttggccc caacatttct tctctgatgg ggaataatgg cccactgagg gaagtacaga 600
ttacaatact atcctgcaga ttgacctttt ctgtagaggg gaaggcaaat ggaagtgaat 660
aacttatgtc caagctttct ttctattgaa ggagaataca ctatgcaag cttgaattt 720
acatccaca ggaagacctc tcaagctaac cccatatact agcctcccta tagctccctt 780
tctattagt gataagcttc ctctaatcac cccacccag aagaaaata 829

<210> 318
<211> 30
<212> PRT
<213> Homo sapien

<400> 318
Thr Ala Ala Ser Asp Asn Phe Gln Leu Ser Gln Gly Gly Gln Gly Phe
1 5 10 15
Ala Ile Pro Ile Gly Gln Ala Met Ala Ile Ala Gly Gln Ile
20 25 30

WO 01/40152

PCT/US01/16776

99

<210> 319
<211> 41
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR primer

<400> 319

ggctctggc aatgggaact cagaagtagt aaactcctg c

41

<210> 320
<211> 41
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR primer

<400> 320

gcaggagtgt tactacttct gaggttccat tggcagaggg c

41

<210> 321
<211> 60
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR primer

<400> 321

ggggaattcc cgtctgtgcc gcgcggcagc cctatggtgg ttgaggttga
ttccatgcgc59
60

<210> 322
<211> 42
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR primer

<400> 322

ccgaattct tatttatttc tggttcttga gacattttct gg

42

<210> 323
<211> 1590
<212> DNA
<213> Homo sapiens

WO 01/90152

PCT/US01/16776

100

```

<400> 323
atgcattacc atcaccatca cagggccgag tccgataact tccagctgtc ccagggtggg 60
caggggtttc cacttcgat cgggcagcag atggcgatcg cgggcccgat caagcttccc 120
acggttcata tggggctaac cgccttcctc ggcttggttg ttgtcgaaaa caacggcaac 180
ggcgacgtag tcaaccggtt ggcgggagc gctccggcgg caagtctcgg cacttccacc 240
ggcgacgtag tcaaccggtt ggcgggagc cagataact cggccaccgc gatggcggac 300
ggccttaaac ggcattatcc cggtagcgtc atctcggtag cctggcaaac caagtggggc 360
ggcagcgcta cagggaaagt gacattggcc gagggaaccc cggcgaatt cctgtggg 420
cggcggggca ggcctatggt ggtttgaggt gattccatgc cgggtgcttc ttctgttaag 480
aagccatttg gtctcaggag caagatgggc aagtgggtgt gcyttgctt cccctgtgc 540
agggagagcg gcaagagcna cgtgggcact tctgagacc accacgactc tcttatgaag 600
acattcagga gcaagatggg caagtggggt cgcactgtct tccctgtgt caggggaggt 660
ggcaagagca acgtgggcgc ttctggagac caagacgact ctgttatgaa gacactcagg 720
aacaagatgg gcaagtgggt ctgcacatgc ttcccctgtt caggggggag cggcaagagc 780
aaggtggggt cttggggaga ctacgatgac agygccttca tggagccag gtaccacgtc 840
cgtggagag atctggacna gclccacaga gctgcctggt ggggttaagt cccagaaaag 900
galotcatgt tcatgtcag ggaactgac gtgaacaaga agcaaacgca aagaggaact 960
galotcatgt tggcctctgc caatgggaat taagaagtag taactctct gctgacaga 1020
cgaatgcacac ttaattctct tgaacaacaa aagaggacag ctctgataaa ggcggtacaa 1080
tgcaggagag atgaatgtgc gtttaatttg ctgggaacatg gcaatgatcc aaatatcca 1140
gatgagatg gaaataccac tctgcaatc gctatctata atgaagataa attaatggcc 1200
aagcaactgc tttatattcg tctgtatgc gnatcaaaaa naagcattgg cttcacacca 1260
ctgttaattg gtgtacatga gcaaaaacag caagtctgta aatttttcat caagaaaaaa 1320
gcgaatttaa atgcactgga tagatacggg aggaactgct tcatacttgc tgtatgtgt 1380
ggatcagcaa gtatagtcag ccttctactt gacaaataa ttgatgtatc ttctcagat 1440
ctatctggac agacggccag agagtatgct gtttctatgc atctcatgtt aatttgcag 1500
ttatcttctg actacaaaga aaaaacagat cttaaatctt ctcttgaaa cagcaatcca 1560
gaaaatgtct caagaccagg aataaataa 1590

```

```

<210> 324
<211> 529
<212> PRT
<213> Homo sapiens

```

```

<400> 324
Met His His His His His Thr Ala Ala Ser Asp Asn Thr Gln Leu
      5              10              15

Ser Gln Gly Gly Gln Gly Phe Ala Ile Pro Ile Gly Gln Ala Met Ala
      20              25              30

Ile Ala Gly Gln Ile Lys Leu Pro Thr Val His Ile Gly Pro Thr Ala
      35              40              45

Phe Leu Gly Leu Gly Val Val Asp Asn Asn Gly Asn Gly Ala Arg Val
      50              55              60

Gln Arg Val Val Gly Ser Ala Pro Ala Ala Ser Leu Gly Ile Ser Thr
      65              70              75              80

Gly Asp Val Ile Thr Ala Val Asp Gly Ala Pro Ile Asn Ser Ala Thr
      85              90              95

Ala Met Ala Asp Ala Leu Asn Gly His His Pro Gly Asp Val Ile Ser
      100             105             110

```

WO 01/90152

PCT/US01/16776

101

Val Thr Trp Gln Thr Lys Ser Gly Gly Thr Arg Thr Gly Asn Val Thr
 115 120 125
 Leu Ala Glu Gly Pro Pro Ala Glu Phe Pro Leu Val Pro Arg Gly Ser
 130 135 140
 Pro Met Val Val Glu Val Asp Ser Met Pro Ala Ala Ser Ser Val Lys
 145 150 155 160
 Lys Pro Phe Gly Leu Arg Ser Lys Met Gly Lys Trp Cys Cys Arg Cys
 165 170 175
 Phe Pro Cys Cys Arg Glu Ser Gly Lys Ser Asn Val Gly Thr Ser Gly
 180 185 190
 Asp His Asp Asp Ser Ala Met Lys Thr Leu Arg Ser Lys Met Gly Lys
 195 200 205
 Trp Cys Arg His Cys Phe Pro Cys Cys Arg Gly Ser Gly Lys Ser Asn
 210 215 220
 Val Gly Ala Ser Gly Asp His Asp Asp Ser Ala Met Lys Thr Leu Arg
 225 230 235 240
 Asn Lys Met Gly Lys Trp Cys Cys His Cys Phe Pro Cys Cys Arg Gly
 245 250 255
 Ser Gly Lys Ser Lys Val Gly Ala Trp Gly Asp Tyr Asp Asp Ser Ala
 260 265 270
 Phe Met Glu Pro Arg Tyr His Val Arg Gly Glu Asp Leu Asp Lys Leu
 275 280 285
 His Arg Ala Ala Trp Trp Gly Lys Val Pro Arg Lys Asp Leu Ile Val
 290 295 300
 Met Leu Arg Asp Thr Asp Val Asn Lys Lys Asp Lys Gln Lys Arg Thr
 305 310 315 320
 Ala Leu His Leu Ala Ser Ala Asn Gly Asn Ser Glu Val Val Lys Leu
 325 330 335
 Leu Leu Asp Arg Arg Cys Gln Leu Asn Val Leu Asp Asn Lys Lys Arg
 340 345 350
 Thr Ala Leu Ile Lys Ala Val Gln Cys Gln Glu Asp Glu Cys Ala Leu
 355 360 365
 Met Leu Leu Glu His Gly Thr Asp Pro Asn Ile Pro Asp Glu Tyr Gly
 370 375 380
 Asn Thr Thr Leu His Tyr Ala Ile Tyr Asn Glu Asp Lys Leu Met Ala
 385 390 395 400
 Lys Ala Leu Leu Leu Tyr Gly Ala Asp Ile Glu Ser Lys Asn Lys His
 405 410 415
 Gly Leu Thr Pro Leu Leu Leu Gly Val His Glu Gln Lys Gln Gln Val

WO 01/90152

PCT/US01/16776

102

420 425 430
 Val Lys Phe Leu Ile Lys Lys Lys Ala Asn Leu Asn Ala Leu Asp Arg
 435 440 445
 Tyr Gly Arg Thr Ala Leu Ile Leu Ala Val Cys Cys Gly Ser Ala Ser
 450 455 460
 Ile Val Ser Leu Leu Leu Glu Gln Asn Ile Asp Val Ser Ser Gln Asp
 465 470 475 480
 Leu Ser Gly Gln Thr Ala Arg Glu Tyr Ala Val Ser Ser His His His
 485 490 495
 Val Ile Cys Gln Leu Leu Ser Asp Tyr Lys Glu Lys Gln Met Leu Lys
 500 505 510
 Ile Ser Ser Glu Asn Ser Asn Pro Glu Asn Val Ser Arg Thr Arg Asn
 515 520 525
 Lys
 <210> 325
 <211> 1155
 <212> DNA
 <213> Homo sapiens
 <400> 325
 atggtggcctg aggtttgttc aatgcccaact gccctctactg tgaagaagcc atttgatctc 60
 agggcaaga tgggcaagtgt gtgcaaacac cgtttccctct gctgcagggg gaggcgcaag 120
 agcaacatgg gcactctctgg agacaaacac gactctcttta tgaagatgct caggagcaag 180
 atgggcaagt gtgcccgcga ctgcttccccc tgctgcaggg ggagcgccac gaggcaactg 240
 ggcaattctg gaggccatga aaactccttt atgaagatgc tgaaggccaa gctggcgcaag 300
 tgggtctgtc actgcttccc ctgctgcagg gggaggggga agagcaactt gggcgcttgg 360
 gggagctacg accaagcggc ctctatggag ccgagglacc acatccgtcg agagactctg 420
 gacagctccc acagagctgc ctggtggggt aaagtcacca gaaagatctc catcgctatg 480
 ctacgggaca ctgcaatgaa caagagggac aaggaaga gaactgctct acatttgcc 540
 tctgcgaatg gaattcaga agtagtacaa ctctgctctg acagagctg tcaacttaat 600
 gtccttgaca acaaaaaag gacagctctg ataaaggcca tacaatgcca ggaagatgaa 660
 tgggtgttaa tgttgctgga acatggcgct gatcgaaata ttccagatga gtagggaaat 720
 accgctctac actatgctat ctacaatgaa gataaattaa tggcacaagc actgctotta 780
 tatgtgtctg atattgaac aaaaaaag gttggcctca caccactttt gcttggcgta 840
 catgaacaaa aacagcaagt ggtgaatttt ttaataaaga aaaaagctaa tttaaatgta 900
 cttagatgat atggaaggac tgcctcataa ctgctgtgat gtttggatc agcaagtata 960
 gtcactcttc tacttgagca aaatgttgat gtatcttctc aagatctatc tggcagaagc 1020
 gccagaaagt atgctgttcc tagtctatcat catgtaattt gtgaattact ttctgactat 1080
 aaggaataac agatgctaaa aatctcttct gaaaacagca atccagaata tgcctcaaga 1140
 accagaaata aataa 1155
 <210> 326
 <211> 384
 <212> PRT
 <213> Homo sapiens
 <400> 326
 Met Val Ala Glu Val Cys Ser Met Pro Thr Ala Ser Thr Val Lys Lys
 5 10 15

WO 01/90152

PCT/US01/16776

103

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

WO 01/90152

PCT/US01/16776

104

Val Asn Leu Leu Ieu Glu Gln Asn Val Asp Val Ser Ser Gln Asp Leu
 525 330 335

Ser Gly Gln Thr Ala Arg Glu Tyr Ala Val Ser Ser His His Val
 340 345 350

Ile Cys Glu Leu Leu Ser Asp Tyr Iys Glu Lys Gln Met Leu Lys Ile
 355 360 365

Ser Ser Glu Asn Ser Asn Pro Glu Asn Val Ser Arg Thr Arg Asn Lys
 370 375 380

<210> 327

<211> 634

<212> DNA

<213> Homo sapiens

<400> 327
 gactgctcta catctggcct ctgccaatgg aaattoagaa gtagtaaaac tctctctgga 60
 cagaagatgt caacttaata tctttgacaa caaaaagagg acagctctga caaaggcagt 120
 acactgcaag gaagatgaat gtggtttaat gttgttgaa catggcactg atccagaat 180
 tccagtgag tatggaaata cgtcttaca ctatgtctac tacaatgaag ataatat 240
 ggcacaagca ctgctcttat acgtgtctga tatgtatca aaaaacagc atggctcag 300
 accactgita ctggtgtgac atgagcaaaa aacgaagtg gtgaaatttt taatcaagaa 360
 aaaaacaat ttaaatgcac tggatagata tggagaagt gctctctac ttgctgtatg 420
 ttgtggatcg gcaagtatag tccgctttct acttgagcaa aacattgat tatcttctca 480
 agatctatct ggcagacgg ccagagagta tgcgtttct agtcgtcata atgtaattg 540
 ccagttaatt tctgactaca aagaaaaa gatactaaaa ctctctctcg aaaaacgcaa 600
 tccaggaat gtctcaagaa ccagaaataa ataa 634

<210> 328

<211> 1155

<212> DNA

<213> Homo sapiens

<400> 328
 atgtgtgttg aggttgatc catgcggcct gctctctctg tgaagaagcc atttggtctc 60
 agcagcaaga tggcgaagt gtgctgocgt tcttccctt gctgcaggga gagcgccag 120
 agcaagctgg gacttctg agacacagac gactctgcta tgaagacact caggagcag 180
 atgggaagt ggtgcagcca ctgcttcccc tctgtcaggg ggaagtgccg gacgaactg 240
 ggcgttctg gagaccacga cgaactctct atgaagacac tcaagaaaca gatggcnaa 300
 tgggtctgcc actgcttccc ctgctgcagc gggagcagca agagcaaggt ggcgcttgg 360
 ggaactacg atgcagctgc ctccatggag ccaggtacc acgtccgtgg agaagatctg 420
 gacaactcc acagagctgc ctggtgggtt aaagtcccc gaaaggtct catctctatg 480
 ctccaggaca ctgactgaa cagcaggac aagcaasga ggaactgctt acatctggcc 540
 tctgcaatg ggaattcaga agtaglaaaa ctctgtctgg acagacgatg tcaactaat 600
 gtctctgaca acaaaaagag gcagctctg acaagggcg tacaatgcca ggaagatgaa 660
 tgtgctttaa tttgtctgga acatggcact gatcaataa tccagatga gtatgaaat 720
 accaactctg actacgctat ctataatgaa gataaattaa tggcacaagc actgcttta 780
 tatgtgtctg atatcgaatc aaaaacaaga catgycctca cccactggt actgtgtgta 840
 catggcaca aacagcaagt cgtgaaattt ttaattaaga aaaaagcgaa ttttaattgca 900
 ctggatagat atggaaggac tgcctccta ctgtgttat gttgtggatc agcaagtata 960
 gtcagcttct tacttgagca aatatattgt gtatcttctc aagatctctc tggcagacg 1020
 gccagagagt atgctgttct tagtcatcat catgtatttt gccagttact tctgcaatg 1080
 aaagaaaaac agatgctaaa aatctcttct gaaacagcga atccagaaaa tgtctcaaga 1140
 accagaata aataa 1155

<210> 329

WO 01/90152

PCT/US01/16776

105

```

<211> 1155
<212> DNA
<213> Homo sapiens

<400> 329
atggtggctg aggtttgttc aatgcccctt gctctctctg tgaagaagcc atttgatctc 60
aggggcaaga tgggcaagtg gtgccaacac cgtctccctt gctgcagggy gaggcgcaag 120
agcaacaatg gcaactctcg agaccaagac gactccttta tgaagagcgt caggagcaay 180
atgggcaagt gttgccaaca ctgctctccc tctgtccaggy ggaagcgca caggcaatgtg 240
ggcaactctg gaggacatga caactccttt atgaagacac tcaggagcaa gttgggcaay 300
tgggtgtgtc actgctctcc ctgctgcagg gggagcgcaa agagcaactg ggcacttag 360
gggaactaag acgacagcgc ctloattgag ccgaggtaac aagtccgtcg agaagatctg 420
gacaagctcc acagagctgc ctgttggggt aagttcccca gaaggatctt catcgtcatg 480
ctcagggaca ctgacatgaa caagagggac aagcaaaaga ggaactgctt acatttggcc 540
tctgccaatg gaatttcaga agtagtaca ctcctgctgg acagacgatg tcaacttaac 600
gtccttgaca acaaaaaaag gacagctctg ataaaggcgg taacatgcca ggaagatgaa 660
tgtgtgttaa tgttgttga acatggcgct gatgaaata ttcaagatga gtatggaaat 720
acgcctctac actatgctat ctacaatgaa gataaattaa tggccaaagc actgctotta 780
tatgtgtgtg atattgaatc aaaaaacaag tctggcctca caccactttt gcttggccta 840
catgaacana aacagcaagt ggtgaatttt ttaataaaga aaaaaatcaa tttaaatgca 900
cttgatagat atggaaagac tgcctctata ctgtctgtat gttgtggato agcaagtata 960
gtcaactctc taactgagca aatgttlyat gtatctcttc aagcttctc tggcagagcg 1020
gcaagagagt atgctgtttc tagtcatcat catgtaattt gtgaattact ttctgactat 1080
aaagaaaaac agatgctaaa aatctctctt gaaaacagca atccagaaaa tgcctcaaga 1140
accagaataa aataa 1155

<210> 330
<211> 1155
<212> DNA
<213> Homo sapiens

<400> 333
atggtggctg aggtttgttc aatgcccctt gctctctctg tgaagaagcc atttgatctc 60
aggggcaaga tgggcaagtg gtgccaacac cgtctccctt gctgcagggy gaggcgcaag 120
agcaacaatg gcaactctcg agaccaagac gactccttta tgaagagcgt caggagcaay 180
atgggcaagt gttgccaaca ctgctctccc tctgtccaggy ggaagcgca caggcaatgtg 240
ggcaactctg gaggacatga caactccttt atgaagatgc tcaggagcaa gatggcgca 300
tgggtgtgtc actgctctcc ctgctgcagg gggagcgcaa agagcaactg ggcacttag 360
gggaactaag acgacagcgc ctloattgag ccgaggtaac aagtccgtcg agaagatctg 420
gacaagctcc acagagctgc ctgttggggt aagttcccca gaaggatctt catcgtcatg 480
ctcagggaca ctgacatgaa caagagggac aagcaaaaga ggaactgctt acatttggcc 540
tctgccaatg gaatttcaga agtagtaca ctcctgctgg acagacgatg tcaacttaac 600
gtccttgaca acaaaaaaag gacagctctg ataaaggcgg taacatgcca ggaagatgaa 660
tgtgtgttaa tgttgttga acatggcgct gatgaataa ttcaagatga gtatggaaat 720
acgcctctac actatgctat ctacaatgaa gataaattaa tggccaaagc actgctotta 780
tatgtgtgtg atattgaatc aaaaaacaag tctggcctca caccactttt gcttggccta 840
catgaacana aacagcaagt ggtgaatttt ttaataaaga aaaaaatcaa tttaaatgta 900
cttgatagat atggaaagac tgcctctata ctgtctgtat gttgtggato agcaagtata 960
gtcaactctc taactgagca aatgttlyat gtatctcttc aagcttctc tggcagagcg 1020
gcaagagagt atgctgtttc tagtcatcat catgtaattt gtgaattact ttctgactat 1080
aaagaaaaac agatgctaaa aatctctctt gaaaacagca atccagaaaa tgcctcaaga 1140
accagaataa aataa 1155

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

```

WO 01/90152

PCT/US01/6776

106

<400> 331
 Thr Ala Leu His Leu Ala Ser Ala Asn Gly Asn Ser Glu Val Val Lys
 3 10 15
 Leu Leu Leu Asp Arg Arg Cys Gln Leu Asn Ile Leu Asp Asn Lys Lys
 20 25 30
 Arg Thr Ala Leu Thr Lys Ala Val Gln Cys Gln Glu Asp Glu Cys Ala
 35 40 45
 Leu Met Leu Leu Glu His Gly Thr Asp Pro Asn Ile Pro Asp Glu Tyr
 50 55 60
 Gly Asn Thr Ala Leu His Tyr Ala Ile Tyr Asn Glu Asp Lys Leu Met
 65 70 75 80
 Ala Lys Ala Leu Leu Leu Tyr Gly Ala Asp Ile Glu Ser Lys Asn Lys
 85 90 95
 His Gly Leu Thr Pro Leu Leu Leu Gly Val His Glu Gln Lys Gln Gln
 100 105 110
 Val Val Lys Phe Leu Ile Lys Lys Lys Ala Asn Leu Asn Ala Leu Asp
 115 120 125
 Arg Tyr Gly Arg Thr Ala Leu Ile Leu Ala Val Cys Cys Gly Ser Ala
 130 135 140
 Ser Ile Val Ser Leu Leu Leu Glu Gln Asn Ile Asp Val Ser Ser Gln
 145 150 155 160
 Asp Leu Ser Gly Gln Thr Ala Arg Glu Tyr Ala Val Ser Ser Arg His
 165 170 175
 Asn Val Ile Cys Gln Leu Leu Ser Asp Tyr Lys Glu Lys Gln Ile Leu
 180 185 190
 Lys Val Ser Ser Glu Asn Ser Asn Pro Gly Asn Val Ser Arg Thr Arg
 195 200 205
 Asn Lys
 210
 <210> 332
 <211> 364
 <212> PRT
 <213> Homo sapiens
 <400> 332
 Met Val Ala Glu Val Cys Ser Met Pro Thr Ala Ser Thr Val Lys Lys
 5 10 15
 Pro Phe Asp Leu Arg Ser Lys Met Gly Lys Trp Cys His His Arg Phe
 20 25 30
 Pro Cys Cys Arg Gly Ser Gly Lys Ser Asn Met Gly Thr Ser Gly Asp
 35 40 45

WO 01/90152

PCT/US01/16776

107

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

WO 01/90152

PCT/US01/16776

108

Ile Cys Glu Leu Leu Ser Asp Tyr Lys Glu Lys Gln Met Leu Lys Ile
 335 360 365
 Ser Ser Glu Asn Ser Asn Pro Glu Asn Val Ser Arg Thr Arg Asn Lys
 370 375 380
 <210> 333
 <211> 384
 <212> PRT
 <213> Homo sapiens
 <400> 333
 Met Val Ala Glu Val Cys Ser Met Pro Ala Ala Ser Ala Val Lys Lys
 5 10 15
 Pro Phe Asp Leu Arg Ser Lys Met Gly Lys Trp Cys His His Arg Phe
 20 25 30
 Pro Cys Cys Arg Gly Ser Gly Lys Ser Asn Met Gly Thr Ser Gly Asp
 35 40 45
 His Asp Asp Ser Phe Met Lys Thr Leu Arg Ser Lys Met Gly Lys Cys
 50 55 60
 Cys His His Cys Phe Pro Cys Cys Arg Gly Ser Gly Thr Ser Asn Val
 65 70 75 80
 Gly Thr Ser Gly Asp His Asp Asn Ser Phe Met Lys Thr Leu Arg Ser
 85 90 95
 Lys Met Gly Lys Trp Cys Cys His Cys Phe Pro Cys Cys Arg Gly Ser
 100 105 110
 Gly Lys Ser Asn Val Gly Thr Trp Gly Asp Tyr Asp Asp Ser Ala Phe
 115 120 125
 Met Glu Pro Arg Tyr His Val Arg Arg Glu Asp Leu Asp Lys Leu His
 130 135 140
 Arg Ala Ala Trp Trp Gly Lys Val Pro Arg Lys Asp Leu Ile Val Met
 145 150 155 160
 Leu Arg Asp Thr Asp Met Asn Lys Arg Asp Lys Gln Lys Arg Thr Ala
 165 170 175
 Leu His Leu Ala Ser Ala Asn Gly Asn Ser Glu Val Val Gln Leu Leu
 180 185 190
 Leu Asp Arg Arg Cys Gln Leu Asn Val Leu Asp Asn Lys Lys Arg Thr
 195 200 205
 Ala Leu Ile Lys Ala Val Gln Cys Gln Glu Asp Glu Cys Val Leu Met
 210 215 220
 Leu Leu Glu His Gly Ala Asp Gly Asn Ile Gln Asp Glu Tyr Gly Asn
 225 230 235 240

WO 01/90152

PCT/US01/16776

109

Thr Ala Leu His Tyr Ala Ile Tyr Asn Glu Asp Lys Leu Met Ala Lys
 245 250 255
 Ala Leu Leu Leu Tyr Gly Ala Asp Ile Glu Ser Lys Asn Lys Cys Gly
 260 265 270
 Leu Thr Pro Leu Leu Leu Gly Val His Glu Gln Lys Gln Gln Val Val
 275 280 285
 Lys Phe Leu Ile Lys Lys Lys Ala Asn Leu Asn Ala Leu Asp Arg Tyr
 290 295 300
 Gly Arg Thr Ala Leu Ile Leu Ala Val Cys Cys Gly Ser Ala Ser Ile
 305 310 315 320
 Val Asn Leu Leu Leu Glu Gln Asn Val Asp Val Ser Ser Gln Asp Leu
 325 330 335
 Ser Gly Gln Thr Ala Arg Glu Tyr Ala Val Ser Ser His His His Val
 340 345 350
 Ile Cys Glu Leu Leu Ser Asp Tyr Lys Glu Lys Gln Met Leu Lys Ile
 355 360 365
 Ser Ser Glu Asn Ser Asn Pro Glu Asn Val Ser Arg Thr Arg Asn Lys
 370 375 380

 <210> 334
 <211> 384
 <212> PRT
 <213> Homo sapiens

 <400> 334
 Met Val Val Glu Val Asp Ser Met Pro Ala Ala Ser Ser Val Lys Lys
 5 10 15
 Pro Phe Gly Leu Arg Ser Lys Met Gly Lys Trp Cys Cys Arg Cys Phe
 20 25 30
 Pro Cys Cys Arg Glu Ser Gly Lys Ser Asn Val Gly Thr Ser Gly Asp
 35 40 45
 His Asp Asp Ser Ala Met Lys Thr Leu Arg Ser Lys Met Gly Lys Trp
 50 55 60
 Cys Arg His Cys Phe Pro Cys Cys Arg Gly Ser Gly Lys Ser Asn Val
 65 70 75 80
 Gly Ala Ser Gly Asp His Asp Asp Ser Ala Met Lys Thr Leu Arg Asn
 85 90 95
 Lys Met Gly Lys Trp Cys Cys His Cys Phe Pro Cys Cys Arg Gly Ser
 100 105 110
 Ser Lys Ser Lys Val Gly Ala Trp Gly Asp Tyr Asp Asp Ser Ala Phe
 115 120 125

WO 01/90152

PCT/US01/16776

110

Met Glu Pro Arg Tyr His Val Arg Gly Glu Asp Leu Asp Lys Leu His
 130 135 140
 Arg Ala Ala Trp Trp Gly Lys Val Pro Arg Lys Asp Leu Ile Val Met
 145 150 155 160
 Leu Arg Asp Thr Asp Val Asn Lys Gln Asp Lys Gln Lys Arg Thr Ala
 165 170 175
 Leu His Leu Ala Ser Ala Asn Gly Asn Ser Glu Val Val Lys Leu Leu
 180 185 190
 Leu Asp Arg Arg Cys Gln Leu Asn Val Leu Asp Asn Lys Lys Arg Thr
 195 200 205
 Ala Leu Ile Lys Ala Val Gln Cys Gln Glu Asp Glu Cys Ala Leu Met
 210 215 220
 Leu Leu Glu His Gly Thr Asp Pro Asn Ile Pro Asp Glu Tyr Gly Asn
 225 230 235 240
 Thr Thr Leu His Tyr Ala Ile Tyr Asn Glu Asp Lys Leu Met Ala Lys
 245 250 255
 Ala Leu Leu Leu Tyr Gly Ala Asp Ile Glu Ser Lys Asn Lys His Gly
 260 265 270
 Leu Thr Pro Leu Leu Leu Gly Val His Glu Gln Lys Gln Val Val
 275 280 285
 Lys Phe Leu Ile Lys Lys Lys Ala Asn Leu Asn Ala Leu Asp Arg Tyr
 290 295 300
 Gly Arg Thr Ala Leu Ile Leu Ala Val Cys Cys Gly Ser Ala Ser Ile
 305 310 315 320
 Val Ser Leu Leu Leu Gln Gln Asn Ile Asp Val Ser Ser Gln Asp Leu
 325 330 335
 Ser Gly Gln Thr Ala Arg Glu Tyr Ala Val Ser Ser His His His Val
 340 345 350
 Ile Cys Gln Leu Leu Ser Asp Tyr Lys Glu Lys Gln Met Leu Lys Ile
 355 360 365
 Ser Ser Glu Asn Ser Asn Pro Glu Asn Val Ser Arg Thr Arg Asn Lys
 370 375 380

【国際公開パンフレット（コレクトバージョン）】

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

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
29 November 2001 (29.11.2001)

PCT

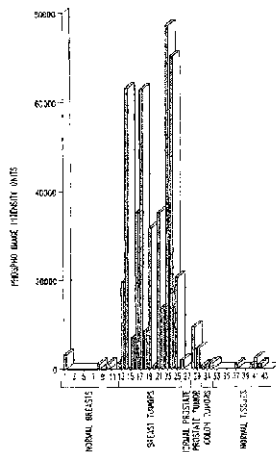
(10) International Publication Number
WO 01/90152 A3

- (51) International Patent Classification: C12N 15/12, 15/62, 510, C12Q 1/68, C07K 14/47, 16/18, A61K 38/00, G01N 33/59
- (21) International Application Number: PCT/US01/16776
- (22) International Filing Date: 22 May 2001 (22.05.2001)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
06/577,505 24 May 2000 (24.05.2000) US
09/590,583 8 June 2000 (08.06.2000) US
09/099,295 26 October 2000 (26.10.2000) US
09/810,936 16 March 2001 (16.03.2001) US
- (71) Applicant (for all designated States except US): CORIXA CORPORATION (US/US), 1124 Columbia Street, Suite 200, Seattle, WA 98104 (US).
- (72) Inventors and
(75) Inventors/Applicants (for US only): FRIEDAKIS, Tony, N. [US/US], 7937 Broadview Pines Blvd., Sarasota, FL 34243 (US); REED, Steven, G. [US/US], 2843 122nd Place N.E., Bellevue, WA 98005 (US); SMITH, John, M. [US/US], 1101 Pierce Ter N.E., Columbia Heights, MN 55421 (US); MISHER, Lynda, E. [US/US], 5043 16th Avenue N.E., Seattle, WA 98105 (US); DILLON, Devin, C. [US/US], 18112 N.W. Muntrous Drive, Issaquah, WA 98027 (US); RETTER, Marc, W. [US/US], 33402 N.E. 4th Place, Carnation, WA 98014 (US); WANG, Aijun [CN/US], 3106 213th Place S.E., Issaquah, WA 98029 (US); SHENK, Yashir, A., W. [IL/US], 15106 S.E. 47th Place, Bellevue, WA 98006 (US); HARLOCKER, Susan, L. [US/US], 7522 13th Avenue W., Seattle, WA 98117 (US); DAY, Craig, H. [US/US], 11501 Sunne Avenue N., C122, Seattle, WA 98135-8317 (US).

[Continued on next page]

(54) Title: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS OF BREAST CANCER

WO 01/90152 A3



(57) Abstract: Compositions and methods for the therapy and diagnosis of cancer, particularly breast cancer, are disclosed. Inactive compositions comprise one or more breast tumor polypeptides, immunogenic portions thereof, polynucleotides that encode such polypeptides, antigen presenting cell that expresses such polypeptides, and T cells that are specific for cells expressing such polypeptides. The disclosed compositions are useful, for example, in the diagnosis, prevention and/or treatment of diseases, particularly breast cancer.

WO 01/90152 A3



(74) Agents: POTTER, Jane, K., R.; Seed Intellectual Property Law Group PLLC, Suite 6300, 701 5th Avenue, Seattle, WA 98104 7092 et al. (US)

patent (AU, BF, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, JI, MC, NL, PT, SE, TR), OAPI patent (BF, BI, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG)

(81) Designated States (national): AF, 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, GL, GR, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

Published:

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

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

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European

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

【国際調査報告】

INT. NATIONAL SEARCH REPORT		International Application No. PCT/US 01/16776
A. CLASSIFICATION OF SUBJECT MATTER IPC 7 C12N15/12 C12N15/62 C12N5/10 C12Q1/68 C07K14/47 C07K16/18 A61K38/00 G01N33/50		
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 C12N C12Q A61K G01N		
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, WPI Data, PAJ, BIOSIS, SEQUENCE SEARCH		
C. DOCUMENTS CONSIDERED TO BE RELEVANT Category * 1 Citation of document, with indication, where appropriate, of the relevant passages Relevant to claim No.		
X	WO 97 25431 A (CORIXA CORP) 17 July 1997 (1997-07-17) see SEQ ID NO: the whole document	1-17
X	WO 97 25426 A (CORIXA CORP) 17 July 1997 (1997-07-17) see SEQ ID NO: 1 and 2 (pp. 34-37) page 1 - page 23; claims 1-42; figures 1,4,6; examples 1-4	1-17
X	WO 98 45328 A (CORIXA CORP) 15 October 1998 (1998-10-15) see SEQ ID NO: 1 and 2 (pp. 34-36) page 1 - page 33; claims 1-49; figures 1,4,6; examples 1-4	1-17
-/-		
☒ Further documents are listed in the continuation of box C. ☒ Patent family members are listed in annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document but published on or after the international filing date "L" document which may throw doubt 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 obvious in combination with the document or documents cited "Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is compared with one or more other such documents, such combination being obvious to a person skilled in the art "Z" document member of the same patent family		
Date of the actual completion of the international search 24 January 2002		Date of mailing of the international search report 25. 04. 2002
Name and mailing address of the ISA European Patent Office, P.O. Box 5818 Postfach 2 NL 2280 HV Rijswijk Tel: (+31-70) 340-2060, Telex: 5211 epo nl, Fax: (+31-70) 340-3018		Authorized officer Oderwald, H

* see PCTISA220 (second sheet) (July 1999)

INT NATIONAL SEARCH REPORT

International Application No.
PCT/US 01/16776

C2(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P, X, L	WO 00 61753 A (CORIXA CORP) 19 October 2000 (2000-10-19) L: priority page 1 -page 33; claims 1-60; examples 1-4 -----	1-17

INTERNATIONAL SEARCH REPORT	International application No. PCT/US 01/16776
Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)	
This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:	
1. ☒ Claims Nos.: because they relate to subject matter not required to be searched by this Authority, namely: Although claims 9 (in so far as in vivo methods are concerned), 12, 13, 17 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.	
2. ☒ Claims Nos.: 6 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.: 1-17 (all partially)	
Remark on Protest ☐ The additional search fees were accompanied by the applicant's protest. ☐ No protest accompanied the payment of additional search fees.	

International Application No. PCT/US 01/16776

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 6

Present claim 6 relates to a product/compound defined by reference to a desirable characteristic or property, namely a binding agent that binds to a polypeptide. The claims cover all products/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 only a very limited number of such products/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 product/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. Consequently, the search has been carried out for those parts of the claims which appear to be clear, supported and disclosed, namely those parts relating to the products/compounds antibodies as mentioned in the description at page 6 first paragraph.

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 65.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

International Application No. PCT/US 01/16776

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

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

1. Claims: Invention 1: claims 1-17 partially

An isolated polynucleotide comprising SEQ ID NO: 1. An isolated polypeptide encoded by said polynucleotide (SEQ ID NO: 2). An expression vector, a host cell, an antibody, a fusion protein, an oligonucleotide. A method for detecting the presence of a cancer in a patient, a method for stimulating or expanding T cells, an isolated T cell population, a method for stimulating an immune response in a patient, a method for the treatment of a cancer patient, a diagnostic kit, a method for inhibiting the development of a cancer in a patient using said polynucleotide, polypeptide, antibody, fusion protein and oligonucleotide.

2. Claims: Inventions 2-255: claims 1-17 partially

same as invention 1 but comprising the nucleotide sequence in the order as given in claim 1 (invention 2 is limited to SEQ ID NO: 3 and invention 255 is limited to SEQ ID NO: 330).

INTERNATIONAL SEARCH REPORT
 Information on patent family members

 International Application No.
 PCT/US 01/16776

Patent document cited in search report	Publication date	Patent family members	Publication date
WO 9725431	A 17-07-1997	AU 1575697 A WO 9725431 A1	01-08-1997 17-07-1997
WO 9725426	A 17-07-1997	AU 728777 B2 AU 1697497 A BR 9707125 A CA 2242340 A1 CN 1211279 A EP 0874902 A2 JP 2001501447 T NO 983183 A US 6344550 B1 WO 9725426 A2 US 6225054 B1	18-01-2001 01-08-1997 20-07-1999 17-07-1997 17-03-1999 04-11-1998 06-02-2001 10-09-1998 05-02-2002 17-07-1997 01-05-2001
WO 9845328	A 15-10-1998	US 6225054 B1 AU 6956098 A BR 9808509 A EP 1127093 A2 EP 0975666 A2 JP 2001521384 T NO 994932 A PL 336349 A1 TR 9903154 T2 WO 9845328 A2 ZA 9802968 A HU 0001270 A2 US 6344550 B1	01-05-2001 30-10-1998 18-09-2001 29-08-2001 02-02-2000 05-11-2001 07-12-1999 19-06-2000 21-08-2000 15-10-1998 27-10-1998 28-07-2000 05-02-2002
WO 0061753	A 19-10-2000	AU 4213000 A EP 1169444 A2 NO 20014009 A WO 0061753 A2	14-11-2000 09-01-2002 06-12-2001 19-10-2000

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 K 48/00	A 6 1 K 48/00	4 C 0 8 6
A 6 1 P 35/00	A 6 1 P 35/00	4 C 0 8 7
C 0 7 K 14/82	C 0 7 K 14/82	4 H 0 4 5
C 0 7 K 16/32	C 0 7 K 16/32	
C 0 7 K 19/00	C 0 7 K 19/00	
C 1 2 N 1/15	C 1 2 N 1/15	
C 1 2 N 1/19	C 1 2 N 1/19	
C 1 2 N 1/21	C 1 2 N 1/21	
C 1 2 N 5/10	C 1 2 Q 1/68	A
C 1 2 Q 1/68	G 0 1 N 33/53	D
G 0 1 N 33/53	G 0 1 N 33/53	M
G 0 1 N 33/566	G 0 1 N 33/566	
G 0 1 N 33/574	G 0 1 N 33/574	A
	C 1 2 N 5/00	A
	C 1 2 N 5/00	B
	A 6 1 K 37/02	

(31)優先権主張番号 09/810,936

(32)優先日 平成13年3月16日(2001.3.16)

(33)優先権主張国 米国(US)

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

(72)発明者 リード, スティーブン ジー.

アメリカ合衆国 ワシントン 98005, ベルビュー, 122エヌディー プレイス エヌ
. イー. 2843

(72)発明者 スミス, ジョン エム.

アメリカ合衆国 ミネソタ 55421, ピアース ター エヌ. イー. 1101, コロン
ビア ハイッ

(72)発明者 ミシャー, リンダ イー.

アメリカ合衆国 ワシントン 98105, シアトル, 16ティーエイチ アベニュー エヌ
. イー. 5043

(72)発明者 ディロン, ダビン シー.

アメリカ合衆国 ワシントン 98027, イザクア, エヌ. ダブリュー. モントリュー
ドライブ 18112

(72)発明者 レッター, マーク ダブリュー.

アメリカ合衆国 ワシントン 98014, カーネイション, エヌ. イー. 43アールディ
ー プレイス 33402

(72)発明者 ワン, アイジュン

アメリカ合衆国 ワシントン 98029, イザクア, 213ティーエイチ プレイス エス
. イー. 3106

(72)発明者 スカイキー, ヤシル エイ. ダブリュー.

アメリカ合衆国 ワシントン 98006, ベルビュー, エス. イー. 47ティーエイチ

プレイス 15106

(72)発明者 ハーロッカー, スーザン エル.

アメリカ合衆国 ワシントン 98117, シアトル, 13ティーエイチ アベニュー ダブ
リュー. 7522

(72)発明者 デイ, クレイグ エイチ.

アメリカ合衆国 ワシントン 98133-8317, シアトル, シー122, ストーン
アベニュー エヌ. 11501

F ターム(参考) 4B024 AA01 AA12 BA36 CA04 CA07 DA02 DA05 DA06 DA11 DA12
EA04 GA11 HA12 HA15

4B063 QA19 QQ03 QQ08 QQ43 QQ58 QR08 QR42 QR56 QS25 QS34
QX02

4B065 AA01X AA57X AA72X AA90X AA93Y AB01 BA02 CA24 CA44 CA46

4C084 AA02 AA06 AA07 AA13 BA01 BA08 BA22 BA23 MA01 NA14
ZB262

4C085 AA13 AA14 CC32 EE01 GG01

4C086 AA01 AA02 AA03 AA04 EA16 MA01 MA04 NA14 ZB26

4C087 BB37 BB63 BB65 CA12 MA01 NA14 ZB26

4H045 AA10 AA11 AA20 AA30 BA10 CA41 DA75 DA86 EA28 EA51

FA72 FA74

专利名称(译)	用于治疗 and 诊断乳腺癌的组合物和方法		
公开(公告)号	JP2004508812A	公开(公告)日	2004-03-25
申请号	JP2001586963	申请日	2001-05-22
[标]申请(专利权)人(译)	科里克萨有限公司		
申请(专利权)人(译)	Corixa公司公司		
[标]发明人	フルダキストニーエヌ リードスティーブンジー スミスジョンエム ミシャーリンダイー デイロンダビンシー レッターマークダブリュー ワンアイジュン スカイキーヤシルエイダブリュー ハーロッカースーザンエル デイクレイグエイチ		
发明人	フルダキス, トニー エヌ. リード, スティーブン ジー. スミス, ジョン エム. ミシャー, リンダ イー. デイロン, ダビン シー. レッター, マーク ダブリュー. ワン, アイジュン スカイキー, ヤシル エイ. ダブリュー. ハーロッカー, スーザン エル. デイ, クレイグ エイチ.		
IPC分类号	G01N33/53 A61K31/7088 A61K35/12 A61K35/14 A61K38/00 A61K39/00 A61K39/395 A61K48/00 A61P35/00 C07K14/15 C07K14/47 C07K14/82 C07K16/32 C07K19/00 C12N1/15 C12N1/19 C12N1/21 C12N5/10 C12N15/09 C12N15/12 C12Q1/68 G01N33/566 G01N33/574		
CPC分类号	C07K14/005 A61K35/12 A61K38/00 A61K39/00 A61K48/00 A61K2039/505 A61K2039/5154 A61K2039/53 A61K2039/57 C07K14/47 C07K14/82 C07K2319/00 C12N2740/10022 C12Q1/6886 C12Q2600/158		
FI分类号	C12N15/00.ZNA.A A61K31/7088 A61K35/14.Z A61K39/395.E A61K39/395.T A61K48/00 A61P35/00 C07K14/82 C07K16/32 C07K19/00 C12N1/15 C12N1/19 C12N1/21 C12Q1/68.A G01N33/53.D G01N33/53.M G01N33/566 G01N33/574.A C12N5/00.A C12N5/00.B A61K37/02		
F-TERM分类号	4B024/AA01 4B024/AA12 4B024/BA36 4B024/CA04 4B024/CA07 4B024/DA02 4B024/DA05 4B024/DA06 4B024/DA11 4B024/DA12 4B024/EA04 4B024/GA11 4B024/HA12 4B024/HA15 4B063/QA19 4B063/QQ03 4B063/QQ08 4B063/QQ43 4B063/QQ58 4B063/QR08 4B063/QR42 4B063/QR56 4B063/QS25 4B063/QS34 4B063/QX02 4B065/AA01X 4B065/AA57X 4B065/AA72X 4B065/AA90X 4B065/AA93Y 4B065/AB01 4B065/BA02 4B065/CA24 4B065/CA44 4B065/CA46 4C084/AA02 4C084/AA06 4C084/AA07 4C084/AA13 4C084/BA01 4C084/BA08 4C084/BA22 4C084/BA23 4C084/MA01 4C084/NA14 4C084/ZB262 4C085/AA13 4C085/AA14 4C085/CC32 4C085/EE01 4C085/GG01 4C086/AA01 4C086/AA02 4C086/AA03 4C086/AA04 4C086/EA16 4C086/MA01 4C086/MA04 4C086/NA14 4C086/ZB26 4C087/BB37 4C087/BB63 4C087/BB65 4C087/CA12 4C087/MA01 4C087/NA14 4C087/ZB26 4H045/AA10 4H045/AA11 4H045/AA20 4H045/AA30 4H045/BA10 4H045/CA41 4H045/DA75 4H045/DA86 4H045/EA28 4H045/EA51 4H045/FA72 4H045/FA74		
代理人(译)	夏木森下		

优先权	09/577505 2000-05-24 US
	09/590583 2000-06-08 US
	09/699295 2000-10-26 US
	09/810936 2001-03-16 US

外部链接 [Espacenet](#)

摘要(译)

癌症（特别是乳腺癌）的组合物和用于治疗和诊断的方法被公开。示例性组合物可以包括一种或多种乳房癌的多肽，或其免疫原性部分，编码这些多肽的多核苷酸，表达这些多肽的抗原呈递细胞，和这些多肽细胞中表达，包括特异性T细胞。所公开的组合物，例如，一种疾病（特别是乳腺癌）的诊断，是在预防和/或治疗是有用的。在某些优选的实施方案中，本发明的多肽和/或多核苷酸是免疫原性的。也就是说，它们可以诱导免疫应答（特别是体液和/或细胞免疫应答）。

