

(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2004-505611

(P2004-505611A)

(43) 公表日 平成16年2月26日(2004.2.26)

(51) Int. Cl.⁷

C 1 2 N 15/09

C 1 2 Q 1/68

G O 1 N 33/53

G O 1 N 33/566

G O 1 N 33/574

F I

C 1 2 N 15/00

C 1 2 Q 1/68

G O 1 N 33/53

G O 1 N 33/566

G O 1 N 33/574

Z N A A

A

M

Z

テーマコード (参考)

2 G O 4 5

4 B O 2 4

4 B O 6 3

審査請求 未請求 予備審査請求 有 (全 185 頁) 最終頁に続く

(21) 出願番号 特願2001-573043 (P2001-573043)
 (86) (22) 出願日 平成13年4月2日 (2001.4.2)
 (85) 翻訳文提出日 平成14年10月3日 (2002.10.3)
 (86) 国際出願番号 PCT/US2001/010631
 (87) 国際公開番号 W02001/075171
 (87) 国際公開日 平成13年10月11日 (2001.10.11)
 (31) 優先権主張番号 60/194, 241
 (32) 優先日 平成12年4月3日 (2000.4.3)
 (33) 優先権主張国 米国 (US)
 (31) 優先権主張番号 60/219, 862
 (32) 優先日 平成12年7月20日 (2000.7.20)
 (33) 優先権主張国 米国 (US)
 (31) 優先権主張番号 60/221, 300
 (32) 優先日 平成12年7月27日 (2000.7.27)
 (33) 優先権主張国 米国 (US)

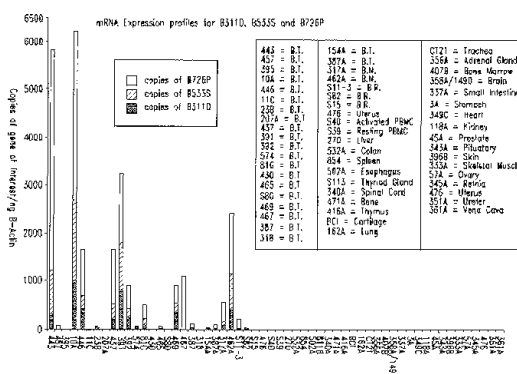
(71) 出願人 397069329
 コリクサ コーポレーション
 アメリカ合衆国 ワシントン 98104
 , シアトル, コロンビア ストリート
 1124, スイート 200
 (74) 代理人 100078282
 弁理士 山本 秀策
 (74) 代理人 100062409
 弁理士 安村 高明
 (74) 代理人 100113413
 弁理士 森下 夏樹
 (72) 発明者 ヒュートン, レイモンド エル.
 アメリカ合衆国 ワシントン 98021
 , ボセル, 242 エヌディー プレイ
 ス エス. イー. 2636

最終頁に続く

(54) 【発明の名称】 乳癌の検出およびモニタリングのための方法、組成物、およびキット

(57) 【要約】

癌 (例えば、乳癌) の治療および診断のための組成物および方法を開示する。組成物は、1つ以上の胸部腫瘍タンパク質、その免疫原性部分、またはこのような部分をコードするポリヌクレオチドを含み得る。あるいは、治療組成物は、胸部腫瘍タンパク質を発現する抗原提示細胞、またはこのようなタンパク質を発現する細胞に特異的なT細胞を含み得る。このような組成物は、例えば、乳癌のような疾患の予防および処置のために用いられ得る。サンプル中の胸部腫瘍タンパク質またはこのようなタンパク質をコードするmRNAを検出することに基づく診断方法もまた提供される。



【特許請求の範囲】

【請求項 1】

患者における癌細胞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) 該生物学的サンプルを、第 1 のポリヌクレオチドにハイブリダイズする第 1 のオリゴヌクレオチドと接触させる工程であって、該第 1 のポリヌクレオチドは、配列番号 73、配列番号 74、および配列番号 76 に示されるポリヌクレオチドからなる群より選択される、工程；
 - (c) 該生物学的サンプルを、第 2 のオリゴヌクレオチドと接触させる工程であって、該第 2 のオリゴヌクレオチドは、配列番号 1、3、5～7、11、13、15、17、19～24、30、32、および 75 からなる群より選択される、第 2 のポリヌクレオチドにハイブリダイズする、工程；
 - (d) 該サンプルにおいて、該オリゴヌクレオチドの少なくとも 1 つに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
 - (e) 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、そこから、該患者における癌の存在または非存在を決定する、工程、
- を包含する、方法。

【請求項 2】

患者における癌細胞の存在または非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) 該生物学的サンプルを、第 1 のオリゴヌクレオチドと接触させる工程であって、該第 1 のオリゴヌクレオチドは、配列番号 73、配列番号 74、および配列番号 76 に示されるポリヌクレオチドからなる群より選択される第 1 のポリヌクレオチドにハイブリダイズする、工程；
 - (c) 該生物学的サンプルを、第 2 のオリゴヌクレオチドと接触させる工程であって、該第 2 のオリゴヌクレオチドは、配列番号 75 に示される、第 2 のポリヌクレオチドにハイブリダイズする、工程；
 - (d) 該生物学的サンプルを、第 3 のオリゴヌクレオチドと接触させる工程であって、該第 3 のオリゴヌクレオチドは、配列番号 5、配列番号 6、および配列番号 7 に示されるポリヌクレオチドからなる群より選択される、第 3 のポリヌクレオチドにハイブリダイズする、工程；
 - (e) 該生物学的サンプルを、第 4 のオリゴヌクレオチドと接触させる工程であって、該第 4 のオリゴヌクレオチドは、配列番号 13、配列番号 15、配列番号 17、配列番号 19、配列番号 20、配列番号 21、配列番号 22、配列番号 23、および配列番号 24 に示されるポリヌクレオチドからなる群より選択される、第 4 のポリヌクレオチドにハイブリダイズする、工程；
 - (f) 該生物学的サンプルにおいて、該オリゴヌクレオチドの少なくとも 1 つに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
 - (g) 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、そこから、該患者における癌の存在または非存在を決定する、工程、
- を包含する、方法。

【請求項 3】

患者における癌細胞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) 該生物学的サンプルを、第 1 のオリゴヌクレオチドと接触させる工程であって、該第 1 のオリゴヌクレオチドは、マンマグロビンおよびリポフィリン B からなる群より選択される第 1 のポリヌクレオチドにハイブリダイズする、工程；

(c) 該生物学的サンプルを、第2のオリゴヌクレオチドと接触させる工程であって、該第2のオリゴヌクレオチドは、G A B A (B 8 9 9 P)、B 7 2 6 P、B 5 1 1 S、B 5 3 3 S、B 3 0 5 D、およびB 3 1 1 Dからなる群より選択される、第2のポリヌクレオチドにハイブリダイズする、工程；

(d) 該サンプルにおいて、該オリゴヌクレオチドの少なくとも1つに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

(e) 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、そこから、該患者における癌の存在または非存在を決定する、工程、

を包含する、方法。

10

【請求項4】

患者における癌細胞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) 該生物学的サンプルを、第1のオリゴヌクレオチドと接触させる工程であって、該第1のオリゴヌクレオチドは、配列番号73および配列番号74に示されるポリヌクレオチドまたはその相補鎖からなる群より選択される第1のポリヌクレオチドにハイブリダイズする、工程；

(c) 該生物学的サンプルを、第2のオリゴヌクレオチドと接触させる工程であって、該第2のオリゴヌクレオチドは、配列番号75に示される第2のポリヌクレオチドまたはその相補鎖にハイブリダイズする、工程；

20

(d) 該生物学的サンプルを、第3のオリゴヌクレオチドと接触させる工程であって、該第3のオリゴヌクレオチドは、配列番号1、配列番号3、配列番号5、配列番号6、および配列番号7に示されるポリヌクレオチド、またはその相補鎖からなる群より選択される、第3のポリヌクレオチドにハイブリダイズする、工程；

(e) 該生物学的サンプルを、第4のオリゴヌクレオチドと接触させる工程であって、該第4のオリゴヌクレオチドは、配列番号11に示されるポリヌクレオチド、またはその相補鎖からなる群より選択される、第4のポリヌクレオチドにハイブリダイズする、工程；

(f) 該生物学的サンプルを、第5のオリゴヌクレオチドと接触させる工程であって、該第5のオリゴヌクレオチドは、配列番号13、配列番号15、および配列番号17に示されるポリヌクレオチド、またはその相補鎖からなる群より選択される、第5のポリヌクレオチドにハイブリダイズする、工程；

30

(g) 該生物学的サンプルを、第6のオリゴヌクレオチドと接触させる、工程であって、該第6のオリゴヌクレオチドは、配列番号19、配列番号20、配列番号21、配列番号22、配列番号23、および配列番号24に示されるポリヌクレオチド、またはその相補鎖からなる群より選択される、第6のポリヌクレオチドにハイブリダイズする、工程；

(h) 該生物学的サンプルを、第7のオリゴヌクレオチドと接触させる工程であって、該第7のオリゴヌクレオチドは、配列番号30に示される第7のポリヌクレオチド、またはその相補鎖にハイブリダイズする、工程；

(i) 該生物学的サンプルを、第8のオリゴヌクレオチドと接触させる工程であって、該第8のオリゴヌクレオチドは、配列番号32に示される第8のポリヌクレオチド、またはその相補鎖にハイブリダイズする、工程；

40

(j) 該生物学的サンプルを、第9のオリゴヌクレオチドと接触させる工程であって、該第9のオリゴヌクレオチドは、配列番号76に示される第9のポリヌクレオチド、またはその相補鎖にハイブリダイズする、工程；

(k) 該生物学的サンプルにおいて、工程(b)～(j)のいずれか1つのハイブリダイズしたオリゴヌクレオチドを検出し、そして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する、工程、

を包含し、ここで、所定のカットオフ値を超える、工程(b)～(j)のいずれか1つにおけるハイブリダイズしたオリゴヌクレオチドの存在は、該患者の生物学的サンプルにお

50

ける癌細胞の存在を示す、方法。

【請求項 5】

患者における癌細胞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) 該生物学的サンプルを、第 1 のオリゴヌクレオチドおよび第 2 のオリゴヌクレオチドと接触させる工程であって；
 - i. ここで、該第 1 のオリゴヌクレオチドおよび該第 2 のオリゴヌクレオチドは、それぞれ、第 1 のポリヌクレオチドおよび第 2 のポリヌクレオチドにハイブリダイズし；
 - i i. ここで、該第 1 のポリヌクレオチドおよび第 2 のポリヌクレオチドは、配列番号 7 3、配列番号 7 4、配列番号 7 5、配列番号 1、配列番号 3、配列番号 5、配列番号 6、配列番号 7、配列番号 1 1、配列番号 1 3、配列番号 1 5、配列番号 1 7、配列番号 1 9、配列番号 2 0、配列番号 2 1、配列番号 2 2、配列番号 2 3、配列番号 2 4、配列番号 3 0、配列番号 3 2、および配列番号 7 6 に示されるポリヌクレオチドからなる群より選択され；ならびに
 - i i i. ここで、該第 1 のポリヌクレオチドは、該第 2 のポリヌクレオチドに対するヌクレオチド配列に下線が付される、工程；
 - (c) 生物学的サンプルにおいて、該ハイブリダイズした第 1 のオリゴヌクレオチドおよび該ハイブリダイズした第 2 のオリゴヌクレオチドを検出する工程；ならびに
 - (d) 該ハイブリダイズした第 1 のオリゴヌクレオチドおよび該ハイブリダイズした第 2 のオリゴヌクレオチドの量を、所定のカットオフ値と比較する工程；
- を包含し、ここで、該所定のカットオフ値を超える、該ハイブリダイズした第 1 のオリゴヌクレオチドまたは該ハイブリダイズした第 2 のオリゴヌクレオチドの量は、該患者の生物学的サンプルにおける癌細胞の存在を示す、方法。

【請求項 6】

患者における癌細胞の存在または非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) 該生物学的サンプルを、第 1 のオリゴヌクレオチドおよび第 2 のオリゴヌクレオチドと接触させる工程であって；
 - i. ここで、該第 1 のオリゴヌクレオチドおよび該第 2 のオリゴヌクレオチドは、それぞれ、第 1 のポリヌクレオチドおよび第 2 のポリヌクレオチドにハイブリダイズし；
 - i i. ここで、該第 1 のポリヌクレオチドおよび該第 2 のポリヌクレオチドは、共に、検出される癌細胞の組織特異的ポリヌクレオチドであり；ならびに
 - i i i. ここで、該第 1 のポリヌクレオチドは、該第 2 のポリヌクレオチドに対するヌクレオチド配列に下線が付される、工程；
 - (c) 該生物学的サンプルにおいて、該第 1 のハイブリダイズしたオリゴヌクレオチドおよび該第 2 のハイブリダイズしたオリゴヌクレオチドを検出する工程；ならびに
 - (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する工程；
- を包含し、ここで、該所定のカットオフ値を超えるハイブリダイズした第 1 のオリゴヌクレオチドまたはハイブリダイズした第 2 のオリゴヌクレオチドの存在は、該患者の生物学的サンプルにおける癌細胞の存在を示す、方法。

【請求項 7】

患者における癌細胞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) 該生物学的サンプルを第 1 の対と比較する工程であって、該第 1 の対は第 1 のオリゴヌクレオチドおよび第 2 のオリゴヌクレオチドを含み、ここで、該第 1 のオリゴヌクレオチドおよび該第 2 のオリゴヌクレオチドは、それぞれ、第 1 のポリヌクレオチドおよびその相補鎖にハイブリダイズする工程；

(c) 該生物学的サンプルを第2のオリゴヌクレオチド対と比較する工程であって、該第2の対は第3のオリゴヌクレオチドおよび第4のオリゴヌクレオチドを含み、ここで、該第3のオリゴヌクレオチドおよび第4のオリゴヌクレオチドは、それぞれ、第2のポリヌクレオチドおよびその相補鎖にハイブリダイズし、そして、ここで、該第1のポリヌクレオチドは、該第2のポリヌクレオチドに対するヌクレオチド配列に無関係である、工程；

(d) 該第1のポリヌクレオチドおよび該第2のポリヌクレオチドを増幅する工程；ならびに

(e) 第1のポリヌクレオチドおよび該増幅した第2のポリヌクレオチドを検出する工程；

を包含し、ここで、該増幅した第1のポリヌクレオチドまたは該増幅した第2のポリヌクレオチドの存在は、該患者における癌細胞の存在を示す、方法。

10

【請求項8】

前記生物学的サンプルが、血液、血清、リンパ節、骨髓、痰、尿、および腫瘍生検サンプルからなる群より選択される、請求項1～7のいずれか1項に記載の方法。

【請求項9】

前記生物学的サンプルが、血液、リンパ節、および骨髓からなる群より選択される、請求項8に記載の方法。

【請求項10】

前記リンパ節が前哨リンパ節である、請求項9に記載の方法。

20

【請求項11】

前記癌が、前立腺癌、乳癌、結腸癌、卵巣癌、肺癌、頭部および頸部の癌、リンパ腫、白血病、黒色腫、肝癌、胃癌、腎癌、膀胱癌、脾癌、および子宮内膜癌からなる群より選択される、請求項1～7のいずれか1項に記載の方法。

【請求項12】

前記第1のポリヌクレオチドおよび前記第2のポリヌクレオチドが、マンマグロビン、リポフィリンB、GABA (B899P)、B726P、B511S、B533S、B305D、およびB311Dからなる群より選択される、請求項6または7に記載の方法。

【請求項13】

前記第一のポリヌクレオチドおよび前記第二のポリヌクレオチドが、以下：
配列番号73、配列番号74、配列番号75、配列番号1、配列番号3、配列番号5、配列番号6、配列番号7、配列番号11、配列番号13、配列番号15、配列番号17、配列番号19、配列番号20、配列番号21、配列番号22、配列番号23、配列番号24、配列番号30、配列番号32、および配列番号76、
に記載されるポリヌクレオチドからなる群から選択される、請求項6および7に記載の方法。

30

【請求項14】

前記オリゴヌクレオチドが、配列番号33～35および63～72に記載されるオリゴヌクレオチドからなる群から選択される、請求項6または7に記載の方法。

【請求項15】

前記第一の増幅されたポリヌクレオチドおよび前記第二のポリヌクレオチドの検出工程が、放射標識を検出する工程および発蛍光団を検出する工程からなる群から選択される工程を包含する、請求項6または7に記載の方法。

40

【請求項16】

前記検出工程が、分画化工程を包含する、請求項1～7のいずれか1項に記載の方法。

【請求項17】

前記第一および第二のオリゴヌクレオチドが、イントロンに広がるオリゴヌクレオチドである、請求項6または7に記載の方法。

【請求項18】

前記イントロンに広がるオリゴヌクレオチドが、配列番号36～62に記載されるオリゴ

50

ヌクレオチドからなる群から選択される、請求項 17 に記載の方法。

【請求項 19】

前記増幅された第一または第二のポリヌクレオチドの前記検出が、該増幅された第一または第二のポリヌクレオチドを、標識したオリゴヌクレオチドプローブと接触させる工程を包含し、ここで該標識されたオリゴヌクレオチドプローブが、中程度にストリンジェントな条件下で、前記第一または第二のポリヌクレオチドにハイブリダイズする、請求項 7 に記載の方法。

【請求項 20】

前記標識されたオリゴヌクレオチドプローブが、放射標識および発蛍光団からなる群から選択される検出可能な部分を含む、請求項 7 に記載の方法。

10

【請求項 21】

該オリゴヌクレオチドプライマーをハイブリダイズする前に、前記生物学的サンプルから、前記癌細胞を富化させる工程をさらに包含する、請求項 1 ~ 7 に記載の方法。

【請求項 22】

前記生物学的サンプルから前記癌細胞を富化する工程が、細胞捕獲および細胞枯渇からなる群から選択される方法論によって達成される、請求項 1 ~ 7 に記載の方法。

【請求項 23】

前記細胞捕獲が、免疫捕獲によって達成され、該免疫捕獲は、以下の工程：

(a) 抗体を、該癌細胞の表面に吸着させる工程；および

(b) 該抗体吸着癌細胞を、該生物学的サンプルの残りから分離する工程、を包含する、方法。

20

【請求項 24】

前記抗体が、以下：

CD2、CD3、CD4、CD5、CD8、CD10、CD11b、CD14、CD15、CD16、CD19、CD20、CD24、CD25、CD29、CD33、CD34、CD36、CD38、CD41、CD45、CD45RA、CD45RO、CD56、CD66B、CD66e、HLA-DR、IgE および TCR からなる群から選択される抗原に指向される、請求項 23 に記載の方法。

【請求項 25】

前記抗体が、胸部腫瘍抗原に指向される、請求項 23 に記載の方法。

30

【請求項 26】

前記抗体が、モノクローナル抗体である、請求項 23 ~ 25 に記載の方法。

【請求項 27】

前記抗体が、磁気ビーズに結合された、請求項 23 に記載の方法。

【請求項 28】

前記抗体が、四量体抗体複合体で処方される、請求項 23 に記載の方法。

【請求項 29】

細胞枯渇が、以下：

(a) 赤血球と白血球とを架橋する工程；および

(b) 該架橋された赤血球と白血球とを、前記生物学的サンプルの残りから分画化する工程、

40

を包含する方法によって達成される、請求項 23 に記載の方法。

【請求項 30】

前記精製工程が、以下：逆転写ポリメラーゼ連鎖反応 (RT-PCR)、逆 PCR、RACE、リガーゼ連鎖反応 (LCR)、Qbeta レプリカーゼ、等温増幅、鎖移動増幅 (SDA)、ローリング連鎖反応 (RCP)、環状プローブ反応 (CPR)、転写ベースの増幅系 (TAS)、核酸配列ベースの増幅 (NASBA) および 3SR からなる群から選択されるポリヌクレオチド増幅方法論によって達成される、請求項 7 に記載の方法。

【請求項 31】

患者の生物学的サンプル中の癌細胞を検出するための組成物であって、該組成物は、以下

50

:

(a) 第一のオリゴヌクレオチド; および

(b) 第二のオリゴヌクレオチド、

を含み、ここで、該第一のオリゴヌクレオチドおよび該第二のオリゴヌクレオチドは、それぞれ、第一のポリヌクレオチドおよび第二のポリヌクレオチドにハイブリダイズし; ここで、該第一のポリヌクレオチドは、ヌクレオチド配列において、該第二のポリヌクレオチドとは関係せず; ここで、第一のポリヌクレオチドおよび該第二のポリヌクレオチドは、検出される癌細胞の組織特異的ポリヌクレオチドであり; ここで、該第一のポリヌクレオチドおよび該第二のポリヌクレオチドは、該癌細胞の組織型の相補的組織特異的ポリヌクレオチドであり; そしてここで、該第一のポリヌクレオチドおよび該第二のポリヌクレオチドが、以下: 配列番号 73、配列番号 74、配列番号 75、配列番号 1、配列番号 3、配列番号 5、配列番号 6、配列番号 7、配列番号 11、配列番号 13、配列番号 15、配列番号 17、配列番号 19、配列番号 20、配列番号 21、配列番号 22、配列番号 23、配列番号 24、配列番号 30、配列番号 32、および配列番号 76、に記載のポリヌクレオチドからなる群から選択される、組成物。

10

【請求項 32】

前記オリゴヌクレオチドが、配列番号 33 ~ 72 に開示されるようなオリゴヌクレオチドからなる群から選択される、請求項 31 に記載の組成物。

【請求項 33】

患者の生物学的サンプル中の癌細胞を検出するための組成物であって、該組成物は、以下:

20

(a) 第一のオリゴヌクレオチド対; および

(b) 第二のオリゴヌクレオチド対、

を含み、ここで、該第一のオリゴヌクレオチド対および該第二のオリゴヌクレオチド対は、それぞれ、第一のポリヌクレオチド(または、その相補体)および該第二のポリヌクレオチド(または、その相補体)にハイブリダイズし; ここで、該第一のポリヌクレオチドは、ヌクレオチド配列において、該第二のポリヌクレオチドとは関係せず; ここで、該第一のポリヌクレオチドおよび該第二のポリヌクレオチドは、検出される癌細胞の組織特異的ポリヌクレオチドであり; ここで、該第一のポリヌクレオチドおよび該第二のポリヌクレオチドは、該癌細胞の組織型の相補的組織特異的ポリヌクレオチドであり; そしてここで、該第一のポリヌクレオチドおよび該第二のポリヌクレオチドが、以下:

30

配列番号 73、配列番号 74、配列番号 75、配列番号 1、配列番号 3、配列番号 5、配列番号 6、配列番号 7、配列番号 11、配列番号 13、配列番号 15、配列番号 17、配列番号 19、配列番号 20、配列番号 21、配列番号 22、配列番号 23、配列番号 24、配列番号 30、配列番号 32、および配列番号 76、に記載のポリヌクレオチドからなる群から選択される、組成物。

【請求項 34】

前記オリゴヌクレオチドが、配列番号 33 ~ 72 に開示されるようなオリゴヌクレオチドからなる群から選択される、請求項 33 に記載の組成物。

【請求項 35】

組成物であって、該組成物は、配列番号 33 ~ 72 に記載されるオリゴヌクレオチドからなる群から選択されるオリゴヌクレオチドを含む、15ヌクレオチドと100ヌクレオチドとの間のオリゴヌクレオチドプライマーまたはプローブを含む、組成物。

40

【請求項 36】

配列番号 33 ~ 72 に記載されるオリゴヌクレオチドからなる群から選択されるオリゴヌクレオチドプライマーまたはプローブを含む、請求項 35 に記載の組成物。

【発明の詳細な説明】

【0001】

(発明の技術分野)

本発明は、一般に、癌診断の分野に関する。より詳細には、本発明は、癌の検出のための

50

方法、組成物およびキットに関し、この検出は、オリゴヌクレオチドのハイブリダイゼーションおよび/または増幅を使用し、癌細胞を含むことが疑われる生物学的サンプルにおける2以上の組織特異的ポリヌクレオチドを同時に検出する。

【0002】

(発明の背景)

癌は、依然として、世界中の重要な健康上の問題である。従来の癌処置レジメンの失敗は、一般に、疾患診断の遅延に一部起因し得る。有意な進歩が癌診断の分野でなされてきたが、依然として、癌細胞の存在の早期で信頼性の高かつ高感度な決定を可能にする、改善された検出方法の必要性が存在する。

【0003】

乳癌は、米国の女性の間の死亡率が肺癌に次いで2番目であり、毎年180,000人より多くの女性が罹患し、そして年に約40,000~50,000人の死亡を生じる。北米の女性について、生涯で乳癌を発症する可能性は、8分の1である。

【0004】

この疾患の管理は、現在、早期の診断(慣用的なスクリーニング手段を介する)および積極的な処置の組み合わせに依存し、その処置は、種々の処置の1以上(例えば、手術、放射線療法、化学療法およびホルモン療法)を含み得る。特定の乳癌についての処置過程は、しばしば、種々の予後パラメーター(特異的な腫瘍マーカーの分析を含む)に基づいて選択される。例えば、Porter-Jordanら、Breast Cancer 8:73-100(1994)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、理解が困難な結果を導き;そして乳癌患者に観察される高い死亡率は、この疾患の診断において改善が必要とされることを示す。

【0005】

Her2/neuに標的化される乳癌処置に対する免疫療法アプローチの近年の導入は、治療的な抗体およびT細胞ワクチンならびにこの疾患の診断のための標的としての、さらなる乳癌特異的遺伝子を同定する有意な動機付けを提供した。この目的のために、マンマグロビン(mammaglobin)が、現在までに発見されている最も乳癌特異的な遺伝子の1つとして同定されており、乳癌の約70~80%において発現される。この高度に組織特異的な分布に起因して、マンマグロビン遺伝子発現の検出は、リンパ節組織における微小転移病変を同定するため、およびより近年では、既知の一次病変および転移性病変を有する乳癌患者の末梢血中の循環乳癌細胞を検出するために使用されている。

【0006】

マンマグロビンは、ウサギグテログロビン(uteroglobin)およびラットステロイド結合タンパク質サブユニットC3のホモログであり、そして高度にグリコシル化される低分子量タンパク質である。Watsonら、Cancer Res. 56:860-5(1996); Watsonら、Cancer Res. 59:3028-3031(1999); Watsonら、Oncogene 16:817-24(1998)。そのホモログとは対照的に、マンマグロビンは、乳癌特異的であることが報告されており、そして過剰発現が、胸部腫瘍生検(23%)、一次および転移性胸部腫瘍(約75%)において記載されており、それと共に、転移性乳癌患者由来の91%のリンパ節においてマンマグロビンのmRNA発現の検出が報告されている。Leygueら、J. Pathol. 189:28-33(1999)およびMinら、Cancer Res. 58:4581-4584(1998)。

【0007】

マンマグロビン発現は乳癌の一般的特徴ではないので、この遺伝子単独の検出は、全ての乳癌の信頼性の高い検出を可能にするには不十分であり得る。従って、当該分野で必要とされていることは、例えば、リンパ節および骨髄における微小転移の検出の感度および信頼性を改善するための、ならびに/または末梢循環中の足場非依存性細胞の認識のための、2以上の乳癌特異的遺伝子の検出を使用する方法である。

【0008】

10

20

30

40

50

本発明は、組織特異的ポリヌクレオチド（特に、腫瘍特異的ポリヌクレオチド）の同定に有用な方法、ならびに疾患に罹患している患者における癌細胞の検出およびモニタリングのための方法、組成物およびキットを提供することによって、これらの目的および他の関連する目的を達成する。

【0009】

（発明の要旨）

特定の実施形態によって、本発明は、1以上の組織特異的ポリヌクレオチドを同定するための方法を提供し、これらの方法は、以下の工程を包含する：（a）遺伝子サブトラクション（genetic subtraction）を行って、目的の組織由来のポリヌクレオチドのプールを同定する工程；（b）DNAマイクロアレイ分析を行って、この目的のポリヌクレオチドのプールの第1のサブセットを同定する工程であって、ここで、この第1のサブセットの各メンバーポリヌクレオチドは、コントロール組織と比較した場合に、この目的の組織において少なくとも2倍過剰発現される、工程；および（c）この第1のサブセット中のポリヌクレオチドに対して定量的ポリメラーゼ連鎖反応分析を行って、ポリヌクレオチドの第2のサブセットを同定する工程であって、このポリヌクレオチドは、コントロール組織と比較して少なくとも2倍過剰発現される、工程。好ましい遺伝子サブトラクションは、ディファレンシャルディスプレイおよびcDNAサブトラクションからなる群より選択され、そして本明細書中以下にさらに詳細に記載する。

10

【0010】

本発明の代替的な実施形態は、目的の組織における一致する組織特異的発現プロファイルおよび/または相補的な組織特異的発現プロファイルを示す、ポリヌクレオチドのサブセットを同定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）DNAマイクロアレイおよび定量的PCRからなる群より選択される発現分析を行って、第1のポリヌクレオチドを同定する工程であって、この第1のポリヌクレオチドは、コントロール組織と比較した場合に、目的の組織において少なくとも2倍過剰発現される、工程；および（b）DNAマイクロアレイおよび定量的PCRからなる群より選択される発現分析を行って、第1のポリヌクレオチドを同定する工程であって、この第1のポリヌクレオチドは、コントロール組織と比較した場合に、目的の組織において少なくとも2倍過剰発現される、工程。

20

【0011】

本発明のさらなる実施形態は、患者における癌細胞の存在を検出するための方法を提供する。このような方法は、以下の工程を包含する：（a）この患者から生物学的サンプルを得る工程；（b）この生物学的サンプルに第1のオリゴヌクレオチド対を接触させる工程であって、ここで、第1のオリゴヌクレオチド対のメンバーは、中程度にストリンジェントな条件下で、それぞれ、第1のポリヌクレオチドおよびその相補体にハイブリダイズする、工程；（c）この生物学的サンプルに第2のオリゴヌクレオチド対を接触させる工程であって、ここで、第2のオリゴヌクレオチド対のメンバーは、中程度にストリンジェントな条件下で、それぞれ、第2のポリヌクレオチドおよびその相補体にハイブリダイズし、そしてこの第1のポリヌクレオチドは、ヌクレオチド配列において第2のポリヌクレオチドに対して無関係である、工程；（d）第1のポリヌクレオチドおよび第2のポリヌクレオチドを増幅する工程；および（e）増幅された第1のポリヌクレオチドおよび増幅された第2のポリヌクレオチドを検出する工程であって；ここで、増幅された第1のポリヌクレオチドまたは増幅された第2のポリヌクレオチドの存在は、患者中の癌細胞の存在を示す、工程。

30

40

【0012】

いくつかの実施形態によって、この増幅された第1のポリヌクレオチドおよび/または第2のポリヌクレオチドの検出には、分画工程（例えば、ゲル電気泳動のような）が先行し得る。あるいは、またはさらに、この増幅された第1のポリヌクレオチドおよび/または第2のポリヌクレオチドの検出は、標識されたオリゴヌクレオチドプローブのハイブリダイゼーションによって達成され得、このプローブは、中程度にストリンジェントな条件下

50

で、第1または第2の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標識は、放射標識したヌクレオチドの組み込みまたは蛍光標識の組み込みによって達成され得る。

【0013】

特定の好ましい実施形態において、特定の組織型の細胞が、検出工程の前の生物学的サンプルから富化され得る。富化は、細胞捕捉および細胞枯渇からなる群より選択される方法によって達成され得る。例示的な細胞捕捉方法としては、免疫捕捉 (i m m u n o c a p t u r e) が挙げられ、そして以下の工程を含む：(a) 組織特異的細胞の表面に対する抗体を、その生物学的サンプル中の細胞に吸着させる工程；(b) 抗体が吸着した組織特異的細胞を、その生物学的サンプルの残りの部分から分離する工程。例示的な細胞枯渇は、赤血球および白血球のクロスリンク (c r o s s - l i n k i n g)、それに続く、クロスリンクした細胞を除去するための分画工程によって達成され得る。

10

【0014】

本発明の代替的な実施形態は、患者における癌の存在または非存在を決定するための方法を提供し、この方法は、以下の工程を包含する：(a) この患者から得られた生物学的サンプルに、胸部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オリゴヌクレオチドを接触させる工程；(b) その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 (例えば、mRNAのような) のレベルを、そのサンプル中で検出する工程；および(c) その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レベルを所定のカットオフ値と比較し、そこからその患者における癌の存在または非存在を決定する工程。特定の実施形態において、mRNAの量は、例えば、上記に言及したポリペプチドをコードするポリヌクレオチド (または、このようなポリヌクレオチドの相補体) に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を介して検出される。他の実施形態において、mRNAの量は、上記に言及したポリペプチドをコードするポリヌクレオチド (または、このようなポリヌクレオチドの相補体) に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。

20

【0015】

関連の局面において、患者における癌の進行をモニターするための方法が提供され、この方法は、(a) 患者から得た生物学的サンプルを、胸部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オリゴヌクレオチドと接触させる工程；(b) このサンプルにおいて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検出する工程；(c) 後の時点でこの患者から得られた生物学的サンプルを使用して工程 (a) および (b) を繰り返す工程；ならびに (d) 工程 (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、それからこの患者における癌の進行をモニターする工程を包含する。

30

【0016】

本発明の特定の実施形態は、この第1のポリヌクレオチドおよび第2のポリヌクレオチドを増幅する工程がポリメラーゼ連鎖反応 (P C R) によって達成されることを提供する。

【0017】

特定の実施形態において、検出される癌細胞は、前立腺癌、乳癌、結腸癌、卵巣癌、肺癌、頭および首の癌、リンパ腫、白血病、黒色腫、肝臓癌、胃癌、腎臓癌、膀胱癌、膵臓癌および子宮内膜癌からなる群より選択され得る。本発明のなおさらなる実施形態は、生物学的サンプルが血液、リンパ節および骨髄からなる群より選択されることを提供する。リンパ節は、前哨リンパ節であり得る。

40

【0018】

本発明の特定の実施形態において、第1のポリヌクレオチドが、マンマグロビン (m a m m a g l o b i n)、リポフィリン B (l i p o p h i l i n B)、G A B A (B 8 9 9 P)、B 7 2 6 P、B 5 1 1 S、B 5 3 3 S、B 3 0 5 D および B 3 1 1 D からなる群より選択される。他の実施形態は、第2のポリヌクレオチドがマンマグロビン、リポフ

50

イリン B、G A B A (B 8 9 9 P)、B 7 2 6 P、B 5 1 1 S、B 5 3 3 S、B 3 0 5 D および B 3 1 1 D からなる群より選択されることを提供する。

【 0 0 1 9 】

本発明の代替の実施形態は、患者における癌の存在または非存在を検出するための方法を提供し、この方法は、(a) 患者から得た生物学的サンプルを、マンマグロビンおよびリポフィリン B からなる群より選択されるポリヌクレオチドにハイブリダイズする第 1 のポリヌクレオチドと接触させる工程；(b) この生物学的サンプルを、G A B A (B 8 9 9 P)、B 7 2 6 P、B 5 1 1 S、B 5 3 3 S、B 3 0 5 D および B 3 1 1 D からなる群より選択されるポリヌクレオチド配列にハイブリダイズする第 2 のオリゴヌクレオチドと接触させる工程；(c) このサンプルにおいて、これらのオリゴヌクレオチドの少なくとも 1 つに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(d) この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て、それからこの患者における癌の存在または非存在を決定する工程を包含する。

10

【 0 0 2 0 】

特定の実施形態に従って、オリゴヌクレオチドは、配列番号 3 3 ~ 7 2 に提示されるオリゴヌクレオチドのような、本明細書中に開示されるオリゴヌクレオチドから選択され得る。他の実施形態によって、この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は、ポリメラーゼ連鎖反応を使用して決定される。あるいは、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は、ハイブリダイゼーションアッセイを使用して決定され得る。

20

【 0 0 2 1 】

本発明のなお他の実施形態は、患者における癌細胞の存在または非存在を決定するための方法を提供し、この方法は、(a) 患者から得た生物学的サンプルを、配列番号 7 3 および配列番号 7 4 に示されるポリヌクレオチドまたはそれらの相補体からなる群より選択されるポリヌクレオチドにハイブリダイズする第 1 のポリヌクレオチドと接触させる工程；(b) この生物学的サンプルを、配列番号 7 5 に示されるポリヌクレオチドまたはその相補体にハイブリダイズする第 2 のオリゴヌクレオチドと接触させる工程；(c) この生物学的サンプルを、配列番号 1、配列番号 3、配列番号 5、配列番号 6、および配列番号 7 に示されるポリヌクレオチドまたはそれらの相補体からなる群より選択されるポリヌクレオチドにハイブリダイズする第 3 のオリゴヌクレオチドと接触させる工程；(d) この生物学的サンプルを、配列番号 1 1 に示されるポリヌクレオチドまたはその相補体からなる群より選択されるポリヌクレオチドにハイブリダイズする第 4 のオリゴヌクレオチドと接触させる工程；(e) この生物学的サンプルを、配列番号 1 3、配列番号 1 5 および配列番号 1 7 に示されるポリヌクレオチドまたはその相補体からなる群より選択されるポリヌクレオチドにハイブリダイズする第 5 のオリゴヌクレオチドと接触させる工程；(f) この生物学的サンプルを、配列番号 1 9、配列番号 2 0、配列番号 2 1、配列番号 2 2、配列番号 2 3、および配列番号 2 4 に示されるポリヌクレオチドまたはそれらの相補体からなる群より選択されるポリヌクレオチドにハイブリダイズする第 6 のオリゴヌクレオチドと接触させる工程；(g) この生物学的サンプルを、配列番号 3 0 に示されるポリヌクレオチドまたはその相補体にハイブリダイズする第 7 のオリゴヌクレオチドと接触させる工程；(h) この生物学的サンプルを、配列番号 3 2 に示されるポリヌクレオチドまたはその相補体にハイブリダイズする第 8 のオリゴヌクレオチドと接触させる工程；(i) この生物学的サンプルを、配列番号 7 6 に示されるポリヌクレオチドまたはその相補体にハイブリダイズする第 9 のオリゴヌクレオチドと接触させる工程；(j) このサンプルにおいて、工程 (a) ~ (i) のいずれか 1 つのハイブリダイズされたオリゴヌクレオチドを検出する工程；ならびに(j) この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する工程であって、ここで、所定のカットオフ値を越える、工程 (a) ~ (i) のいずれか 1 つにおいてハイブリダイズされたオリゴヌクレオチドの存在が、この患者の生物学的サンプルにおける癌細胞の存在を示す、工程を包含する。

30

40

50

【 0 0 2 2 】

本発明の他の関連する実施形態は、患者における癌細胞の存在または非存在を決定するための方法を提供し、この方法は、(a)患者から得られた生物学的サンプルを第1のオリゴヌクレオチドおよび第2のオリゴヌクレオチドと接触させる工程であって、ここで、この第1のオリゴヌクレオチドおよび第2のオリゴヌクレオチドが、中程度にストリンジェントな条件下で、配列番号73、配列番号74、配列番号75、配列番号1、配列番号3、配列番号5、配列番号6、配列番号7、配列番号11、配列番号13、配列番号15、配列番号17、配列番号19、配列番号20、配列番号21、配列番号22、配列番号23、配列番号24、配列番号30、配列番号32、および配列番号76からなる群より選択される第1のオリゴヌクレオチドおよび第2のオリゴヌクレオチドにハイブリダイズし、ここで、この第1のオリゴヌクレオチドがこの第2のオリゴヌクレオチドに構造的に関連しない、工程；(b)サンプルにおいて、この第1のハイブリダイズされたオリゴヌクレオチドおよび第2のハイブリダイズされたオリゴヌクレオチドを検出する工程；および(c)このオリゴ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の量を所定のカットオフ値と比較する工程であって、ここで、所定のカットオフ値を越える、ハイブリダイズされた第1のオリゴヌクレオチドまたはハイブリダイズされた第2のオリゴヌクレオチドの存在は、この患者の生物学的サンプルにおける癌細胞の存在を示す、工程を包含する。

10

【 0 0 2 3 】

本発明の他の関連の実施形態は、患者の癌細胞の存在または非存在を決定するための方法を提供し、この方法は、(a)患者から得た生物学的サンプルを、第1のオリゴヌクレオチドおよび第2のオリゴヌクレオチドと接触する工程であって、ここで、この第1のオリゴヌクレオチドおよび第2のオリゴヌクレオチドが、中程度のストリンジェントな条件下で、第1のオリゴヌクレオチドおよび第2のオリゴヌクレオチドにハイブリダイズし、両方が検出される癌の組織特異的オリゴヌクレオチドであり、ここで、この第1のオリゴヌクレオチドが、この第2のオリゴヌクレオチドに構造的に関連しない、工程；(b)サンプルにおいてこの第1のハイブリダイズしたオリゴヌクレオチドおよび第2のハイブリダイズしたオリゴヌクレオチドを検出する工程；および(c)このオリゴ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の量を所定のカットオフ値と比較する工程であって、ここで、所定のカットオフ値を越える、ハイブリダイズされた第1のオリゴヌクレオチドまたはハイブリダイズされた第2のオリゴヌクレオチドの存在は、この患者の生物学的サンプルにおける癌細胞の存在を示す、工程を包含する。

20

30

【 0 0 2 4 】

他の関連する局面において、本発明は、さらに、本明細書中に開示される方法において有用な組成物を提供する。例示的な組成物は、2つ以上のオリゴヌクレオチドプライマー対(これらのそれぞれが、異なるオリゴヌクレオチドに特異的にハイブリダイズする)を含む。本発明の組成物に適切な例示的なオリゴヌクレオチドプライマーは、配列番号33~71によって本明細書中で開示される。本発明の組成物に適切な例示的なオリゴヌクレオチドは、配列番号73、配列番号74、配列番号75、配列番号1、配列番号3、配列番号5、配列番号6、配列番号7、配列番号11、配列番号13、配列番号15、配列番号17、配列番号19、配列番号20、配列番号21、配列番号22、配列番号23、配列番号24、配列番号30、配列番号32、および配列番号76に開示される。

40

【 0 0 2 5 】

本発明はまた、本発明の検出方法を実行するのに適切であるキットを提供する。例示的なキットとしては、オリゴヌクレオチドプライマー対(それぞれが、異なるオリゴヌクレオチドに特異的にハイブリダイズする)を含む。特定の実施形態において、本発明に従うキットはまた、核酸ポリメラーゼおよび適切な緩衝液を含み得る。本発明のキットに適切な例示的なオリゴヌクレオチドプライマーは、配列番号33~71によって本明細書中で開示される。本発明のキットに適切な例示的なオリゴヌクレオチドは、配列番号73、配列番号74、配列番号75、配列番号1、配列番号3、配列番号5、配列番号6、配列番号7、配列番号11、配列番号13、配列番号15、配列番号17、配列番号19、配列番号2

50

0、配列番号21、配列番号22、配列番号23、配列番号24、配列番号30、配列番号32、および配列番号76に開示される。

【0026】

本発明のこれらおよび他の局面は、以下の詳細な説明および添付の図面を参照して明らかとなる。本明細書中に開示されるすべての参考文献は、本明細書中にそれぞれが個々に組み込まれているように、その全体が参考として援用される。

【0027】

配列番号1は、B305DアイソフォームAの第1のスプライス改変体についての決定されたcDNA配列である。

【0028】

配列番号2は、配列番号1の配列によりコードされるアミノ酸配列である。

【0029】

配列番号3は、B305DアイソフォームAの第2のスプライス改変体についての決定されたcDNA配列である。

【0030】

配列番号4は、配列番号3の配列によりコードされるアミノ酸配列である。

【0031】

配列番号5～7は、B305DアイソフォームCの3つのスプライス改変体についての決定されたcDNA配列である。

【0032】

配列番号8～10は、それぞれ配列番号5～7の配列によりコードされるアミノ酸配列である。

【0033】

配列番号11は、B311Dについての決定されたcDNA配列である。

【0034】

配列番号12は、配列番号11の配列によりコードされるアミノ酸配列である。

【0035】

配列番号13は、B726Pの第1のスプライス改変体の決定されたcDNA配列である。

【0036】

配列番号14は、配列番号13の配列によりコードされるアミノ酸配列である。

【0037】

配列番号15は、B726Pの第2のスプライス改変体の決定されたcDNA配列である。

【0038】

配列番号16は、配列番号15の配列によりコードされるアミノ酸配列である。

【0039】

配列番号17は、B726Pの第3のスプライス改変体の決定されたcDNA配列である。

【0040】

配列番号18は、配列番号17の配列によりコードされるアミノ酸配列である。

【0041】

配列番号19～24は、B726Pのさらなるスプライス改変体の決定されたcDNA配列である。

【0042】

配列番号25～29は、それぞれ配列番号19～24の配列によりコードされるアミノ酸配列である。

【0043】

配列番号30は、B511Sについての決定されたcDNA配列である。

【0044】

10

20

30

40

50

配列番号 31 は、配列番号 30 によりコードされたアミノ酸配列である。

【 0 0 4 5 】

配列番号 32 は、B 5 3 3 S についての決定された c D N A 配列である。

【 0 0 4 6 】

配列番号 33 は、リポフィリン B 正方向プライマーの D N A 配列である。

【 0 0 4 7 】

配列番号 34 は、リポフィリン B 逆方向プライマーの D N A 配列である。

【 0 0 4 8 】

配列番号 35 は、リポフィリン B プローブの D N A 配列である。

【 0 0 4 9 】

配列番号 36 は、G A B A (B 8 9 9 P) 正方向プライマーの D N A 配列である。

【 0 0 5 0 】

配列番号 37 は、G A B A (B 8 9 9 P) 逆方向プライマーの D N A 配列である。

【 0 0 5 1 】

配列番号 38 は、G A B A (B 8 9 9 P) プローブの D N A 配列である。

【 0 0 5 2 】

配列番号 39 は、B 3 0 5 D (C 形態) 正方向プライマーの D N A 配列である。

【 0 0 5 3 】

配列番号 40 は、B 3 0 5 D (C 形態) 逆方向プライマーの D N A 配列である。

【 0 0 5 4 】

配列番号 41 は、B 3 0 5 D (C 形態) プローブの D N A 配列である。

【 0 0 5 5 】

配列番号 42 は、B 7 2 6 P 正方向プライマーの D N A 配列である。

【 0 0 5 6 】

配列番号 43 は、B 7 2 6 P 逆方向プライマーの D N A 配列である。

【 0 0 5 7 】

配列番号 44 は、B 7 2 6 P プローブの D N A 配列である。

【 0 0 5 8 】

配列番号 45 は、アクチン正方向プライマーの D N A 配列である。

【 0 0 5 9 】

配列番号 46 は、アクチン逆方向プライマーの D N A 配列である。

【 0 0 6 0 】

配列番号 47 は、アクチンプローブの D N A 配列である。

【 0 0 6 1 】

配列番号 48 は、マンマグロビン正方向プライマーの D N A 配列である。

【 0 0 6 2 】

配列番号 49 は、マンマグロビン逆方向プライマーの D N A 配列である。

【 0 0 6 3 】

配列番号 50 は、マンマグロビンプローブの D N A 配列である。

【 0 0 6 4 】

配列番号 51 は、第 2 の G A B A (B 8 9 9 P) 逆方向プライマーの D N A 配列である。

【 0 0 6 5 】

配列番号 52 は、第 2 の B 7 2 6 P 正方向プライマーの D N A 配列である。

【 0 0 6 6 】

配列番号 53 は、G A B A B 8 9 9 P - I N T 正方向プライマーの D N A 配列である。

【 0 0 6 7 】

配列番号 54 は、G A B A B 8 9 9 P - I N T 逆方向プライマーの D N A 配列である。

【 0 0 6 8 】

配列番号 55 は、G A B A B 8 9 9 P - I N T T a q m a n プローブの D N A 配列である。

10

20

30

40

50

【 0 0 6 9 】

配列番号 5 6 は、 B 3 0 5 D - I N T 正方向プライマーの D N A 配列である。

【 0 0 7 0 】

配列番号 5 7 は、 B 3 0 5 D - I N T 逆方向プライマーの D N A 配列である。

【 0 0 7 1 】

配列番号 5 8 は、 B 3 0 5 D - I N T T a q m a n プローブの D N A 配列である。

【 0 0 7 2 】

配列番号 5 9 は、 B 7 2 6 - I N T 正方向プライマーの D N A 配列である。

【 0 0 7 3 】

配列番号 6 0 は、 B 7 2 6 - I N T 逆方向プライマーの D N A 配列である。

10

【 0 0 7 4 】

配列番号 6 1 は、 B 7 2 6 - I N T T a q m a n プローブの D N A 配列である。

【 0 0 7 5 】

配列番号 6 2 は、 G A B A B 8 9 9 P T a q m a n プローブの D N A 配列である。

【 0 0 7 6 】

配列番号 6 3 は、 B 3 1 1 D 正方向プライマーの D N A 配列である。

【 0 0 7 7 】

配列番号 6 4 は、 B 3 1 1 D 逆方向プライマーの D N A 配列である。

【 0 0 7 8 】

配列番号 6 5 は、 B 3 1 1 D T a q m a n プローブの D N A 配列である。

20

【 0 0 7 9 】

配列番号 6 6 は、 B 5 3 3 S 正方向プライマーの D N A 配列である。

【 0 0 8 0 】

配列番号 6 7 は、 B 5 3 3 S 逆方向プライマーの D N A 配列である。

【 0 0 8 1 】

配列番号 6 8 は、 B 5 3 3 S T a q m a n プローブの D N A 配列である。

【 0 0 8 2 】

配列番号 6 9 は、 B 5 1 1 S 正方向プライマーの D N A 配列である。

【 0 0 8 3 】

配列番号 7 0 は、 B 5 1 1 S 逆方向プライマーの D N A 配列である。

30

【 0 0 8 4 】

配列番号 7 1 は、 B 5 1 1 S T a q m a n プローブの D N A 配列である。

【 0 0 8 5 】

配列番号 7 2 は、 G A B A 逆方向プライマーの D N A 配列である。

【 0 0 8 6 】

配列番号 7 3 は、マンマグロビンについての全長 c D N A 配列である。

【 0 0 8 7 】

配列番号 7 4 は、 E . c o l i において発現されるマンマグロビン組換えポリペプチドをコードするオープンリーディングフレームの決定された c D N A 配列である。

【 0 0 8 8 】

配列番号 7 5 は、 G A B A についての全長 c D N A 配列である。

40

【 0 0 8 9 】

配列番号 7 6 は、リポフィリン B についての全長 c D N A 配列である。

【 0 0 9 0 】

配列番号 7 7 は、配列番号 7 6 の配列によりコードされるアミノ酸配列である。

【 0 0 9 1 】

(発明の詳細な説明)

上記のように、本発明は、一般に、組織特異的なポリヌクレオチドの同定に適切な方法、ならびに癌の診断およびモニタリングに適切な方法、組成物およびキットに関する。本明細書中に開示された特定の例証された方法、組成物およびキットは乳癌、特に乳癌特異的

50

ポリヌクレオチドの同定、検出およびモニタリングに関するが、本発明は一般に広範な種々の癌および関連の過剰発現したポリヌクレオチド（例えば、前立腺癌、乳癌、結腸癌、卵巣癌、肺癌、頭頸癌、リンパ腫、白血病、黒色腫、肝癌、胃癌、腎癌、膀胱癌、膵臓癌および子宮内癌を含む）の同定、検出およびモニタリングに適用可能であることが当業者によって理解される。従って、本発明が乳癌特異的ポリヌクレオチドの同定または乳癌の検出およびモニタリングにのみ限定されないことは、明白である。

【0092】

（組織特異的ポリヌクレオチドの同定）

本発明の特定の実施形態は、生物学的サンプル内の癌細胞の検出についての方法、組成物およびキットを提供する。これらの方法は、患者由来の生物学的サンプルから1以上の組織特異的ポリヌクレオチドを検出する工程を包含する。このポリヌクレオチドの過剰発現は、その患者の生物学的サンプル内の癌細胞の存在を示す。従って、本発明はまた、組織特異的ポリヌクレオチドの同定に適切である方法を提供する。本明細書中で使用される場合、句「組織特異的ポリヌクレオチド」または「腫瘍特異的ポリヌクレオチド」は、1以上のコントロール組織と比較して少なくとも2倍の過剰発現であるポリヌクレオチド全てを含むことが意味される。本明細書中以下にさらに詳細に議論されるように、所定のポリヌクレオチドの過剰発現は、例えば、マイクロアレイおよび/または定量的リアルタイムポリメラーゼ連鎖反応（Real-time PCRTM）方法論によって評価され得る。

10

【0093】

組織特異的ポリヌクレオチドを検出するための例示的方法は、以下の工程を包含する：（a）目的の組織由来のポリヌクレオチドのプールを同定するために遺伝子サブトラクションを実施する工程；（b）目的のポリヌクレオチドの上記プールの第一のセットを同定するためにDNAマイクロアレイ分析を実施する工程であって、ここで、上記第一のサブセットの各メンバーポリヌクレオチドは、上記目的の組織においてコントロール組織と比較して少なくとも2倍過剰発現される工程；および（c）上記第一のサブセット中のポリヌクレオチドに定量的ポリメラーゼ連鎖反応分析をして、上記コントロール組織と比較して少なくとも2倍過剰発現されるポリヌクレオチドの第二のサブセットを同定する工程。

20

【0094】

（一般的ポリヌクレオチド）

本明細書中で使用される場合、用語「ポリヌクレオチド」は一般に、DNA分子またはRNA分子のいずれかをいう。ポリヌクレオチドは、生物学的サンプル（例えば、血液サンプル、血清サンプル、リンパ節サンプル、骨髄サンプル、痰サンプル、尿サンプルおよび腫瘍生検サンプル）中に正常に見出されるように天然に存在し得る。あるいは、ポリヌクレオチドは、例えば、核酸重合反応によって合成的に誘導され得る。当業者によって認識されるように、ポリヌクレオチドは、一本鎖（コーディングまたはアンチセンス）あるいは二本鎖であり得、そしてDNA分子（ゲノム、cDNAまたは合成）あるいはRNA分子であり得る。RNA分子は、HnRNA分子を含む。これは、イントロンを含み、そして1対1の様式でDNA分子に対応する。そしてRNA分子は、イントロンを含まない。さらなるコード配列または非コード配列は、本発明のポリペプチド内に存在し得るが、必要ではない。そしてポリヌクレオチドは、他の分子および/または支持体物質に結合し得るが、必要ではない。

30

40

【0095】

ポリヌクレオチドは、ネイティブ配列（すなわち、腫瘍タンパク質をコードする内因性配列（例えば、胸部腫瘍またはその一部））またはこのような配列の改変体、または生物学的もしくは抗原的に機能的な等価物を含み得る。ポリヌクレオチド改変体は、以下にさらに記載されるように1以上の置換、付加、欠失および/または挿入を含み得る。用語「改変体」はまた、異種性器官の相同遺伝子を含む。

【0096】

ポリヌクレオチド配列またはポリペプチド配列を比較する場合、最大一致について整列さ

50

れる場合にその２つの配列中のヌクレオチドまたはアミノ酸の配列が同一であるならば、以下に記載されるように２つの配列は、「同一」であるといわれる。２つの配列の間の比較は、配列類似性の局所領域を同定および比較するために比較ウインドウを超えて配列を比較することによって代表的に実施される。本明細書中で使用される場合、「比較ウインドウ」は、少なくとも約２０、通常は３０～約７５、４０～約５０連続する位置のセグメントをいう。ここで、２つの配列が最適に整列される後に、配列は、同数の連続する位置の参照配列と比較され得る。

【００９７】

比較のための配列の最適整列は、バイオインフォマティクスソフトウェアの *Laser gene suite* 中の *Megalign* プログラム (*DNA STAR, Inc., Madison, WI*) を、初期設定の変数を使用して実施され得る。このプログラムは、以下の参考文献に記載されるいくつかの整列図式を具体化する。Dayhoff, M. O. (1978) *A model of evolutionary change in proteins - Matrices for detecting distant relationships*. Dayhoff, M. O. (編) *Atlas of Protein Sequence and Structure*, National Biomedical Research Foundation, Washington DC 第５巻、第３章、３４５ - ３４８頁; Hein J. (1990) *Unified Approach to Alignment and Phylogenesis* 626 - 645頁 *Methods in Enzymology*, 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: 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。

【００９８】

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 の局所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 の類似性についての探索方法によって、これらのアルゴリズム (*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*) のコンピューター化改良によって実施され得る。

【００９９】

配列同一性パーセントおよび配列類似性を決定するために適切なアルゴリズムの１つの好ましい例示は、*BLAST* アルゴリズム 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* を、例えば、本明細書中に記載される変数を用いて使用して、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パーセントを決定し得る。*BLAST* 分析を実施するためのソフトウェアは、National Center for Biotechnology

Informationを介して公的に利用可能である。1つの実例としての例示において、蓄積スコアは、ヌクレオチド配列について変数M（一致する残基の対についての報酬スコア；常に >0 ）およびN（一致しない残基についてのペナルティスコア；常に <0 ）を使用して計算され得る。アミノ酸配列について、スコア付けマトリックスを使用して、蓄積スコアを計算し得る。各方法におけるワードヒットの伸長は、以下の場合に停止される：数量Xによって蓄積整列スコアがその最大達成値より減少する；1以上のネガティブスコア付け残基整列の蓄積に起因して蓄積スコアが0以下になる；または、どちらかの配列の末端が到達される。BLASTアルゴリズム変数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（W）11、および期待値（E）10を初期設定として使用し、そしてBLOSUM 62スコア付けマトリックス（HenikoffおよびHenikoff（1989）Proc. Natl. Acad. Sci. USA 89:10915を参照のこと）整列は、（B）50、期待値（E）10、M = 5、N = 4および両鎖の比較を使用する。

10

【0100】

好ましくは、「配列同一性のパーセント」は、少なくとも20位の比較のウィンドウを越える2つの最適な整列された配列を比較することによって決定される。ここで、比較ウィンドウ中のポリヌクレオチド配列またはポリペプチド配列の一部は、この2つの配列の最適な整列についての参照配列（付加または欠失を含まない）と比較して20%以下、通常は5~15%、または10~12%の付加または欠失（すなわち、ギャップ）を含み得る。パーセントは、位置の数を決定する工程（ここで、同一の核酸塩基またはアミノ酸残基は、一致した位置の数を得るために両方の配列中に生じる）、一致した位置の数を参照配列（すなわち、ウィンドウサイズ）中の位置の総数で除する工程、およびこれらの結果に100を掛けて配列の同一性のパーセントを得る工程によって算出される。

20

【0101】

従って、本発明は、本明細書中に開示される配列に対して実質的な同一性を有するポリヌクレオチド配列およびポリペプチド配列を含む。例えば、これらは、本明細書中に記載される方法（例えば、以下に記載されるような標準変数を使用するBLAST分析）を使用して、本発明のポリヌクレオチド配列またはポリペプチド配列を比較して、少なくとも50%配列同一性、好ましくは少なくとも55%、60%、65%、70%、75%、80%、85%、90%、95%、96%、97%、98%または99%以上の配列同一性を 30 含む。コドン縮重、アミノ酸類似性、リーディングフレーム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よってコードされるタンパク質の対応する同一性を決定するためにこれらの値が適切に調整され得ることを、当業者は認識する。

30

【0102】

さらなる実施形態において、本発明は、本明細書に開示される1つ以上の配列に対して配列同一であるか、または相補性である種々の長さの連続するストレッチを含む単離されたポリヌクレオチドおよびポリペプチドを提供する。例えば、本発明によってポリヌクレオチドが、提供され、このポリヌクレオチドは、本明細書に開示される1つ以上の配列の少なくとも約15、20、30、40、50、75、100、150、200、300、400、500または1000以上連続するヌクレオチド、ならびにその間の中間の長さの 40 全てのヌクレオチドを含む。この文脈において、「中間の長さ」とは、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）のような、述べられた値の間の任意の長さを意味することが容易に理解される。

40

【0103】

本発明のポリヌクレオチド、またはそのフラグメントは、コード配列それ自体の長さにもかかわらず、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、マルチクロニング部位、他のコードセグメントなど）と組み合わせられ得、その結果その長さ全体にわたってかなり変化し得る。従って、任意の長さのほとん 50

50

どの核酸フラグメントが使用され得ることが考えられる。この全長は、好ましくは調製の容易さおよび意図される組み換えDNAプロトコールにおける用途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）の全長を有する代表的なDNAセグメントが、本発明の多くの実施において有用であることが考えられる。

【0104】

他の実施形態において、本発明は、本明細書において提供されるポリヌクレオチド配列に対して、中程度にストリンジェントな条件下でハイブリダイズし得るポリヌクレオチド、またはそのフラグメント、またはその相補的な配列に関する。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の適切な中程度にストリンジェントな条件としては、 $5 \times \text{SSC}$ 、 $0.5\% \text{ SDS}$ 、 1.0 mM EDTA ($\text{pH } 8.0$) の溶液中の事前洗浄； $50 \sim 65^\circ\text{C}$ 、 $5 \times \text{SSC}$ での、一晩のハイブリダイゼーション；続いて、 $0.1\% \text{ SDS}$ を含有する $2 \times$ 、 $0.5 \times$ および $0.2 \times \text{SSC}$ のそれぞれを用いた20分間、 65°C での2回の洗浄が挙げられる。

10

【0105】

さらに、遺伝子コードの縮重性の結果として、本明細書において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いくつかは、任意のネイティブ遺伝子のヌクレオチド配列に対して最小限の相同性を保有する。それにもかかわらず、コドン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が、本発明によって特に意図される。さらに、本明細書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つ以上の変異（例えば、ヌクレオチドの欠失、付加、および/または置換）の結果として変更される内因性遺伝子である。得られたmRNAおよびタンパク質は、おそらく、ただし必ずしもではないが、変化した構造または機能を有する。対立遺伝子は、標準的技術（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。

20

【0106】

（マイクロアレイ分析）

本発明の方法による検出に適切であるポリヌクレオチドは、以下にさらに詳細に記載されるように、組織および/または腫瘍関連発現（例えば、本明細書において提供される代表的なアッセイを用いて決定される場合、正常組織よりも腫瘍において少なくとも2倍大きい発現）についてcDNAのマイクロアレイを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、製造業者の指示に従って（そして、本質的には、Schenら、Proc. Natl. Acad. Sci. USA 93:10614~10619 (1996) および Hellerら、Proc. Natl. Acad. Sci. USA 94:2150~2155 (1997) に記載のように）、Syntheni microarray (Palo Alto, CA) を用いて、実施され得る。

30

【0107】

マイクロアレイは、多数の遺伝子を評価するための有効な方法であるが、その感度の限界に起因して、量の少ない遺伝子の絶対的な組織分布を正確に決定し得ないか、またはシグナル飽和に起因して、より豊富な遺伝子の過剰発現の程度を過小評価し得る。マイクロアレイ発現プロファイリングによる過剰発現を示すこれらの遺伝子について、TaqqmanTM プローブ検出に基づいた定量的RT-PCRを用いてさらに分析を実施した。この分析はさらに大きいダイナミックレンジの感度を有する。この目的のために、正常および腫瘍組織、遠位転移、および細胞株のいくつかの異なるパネルを用いた。

40

【0108】

（定量的リアルタイムポリメラーゼ連鎖反応）

本発明による適切なポリヌクレオチドは、定量的PCR方法（例えば、リアルタイムPCR

50

R 方法 (Real-time PCR methodology) など) を使用することによって、さらに特徴づけされるか、あるいは本来的に同定され得る。この方法論によって、組織および/または腫瘍のサンプル (例えば、転移性腫瘍サンプルなど) が、対応する正常組織サンプル、および/または関連しない正常組織サンプルのパネルと並べて試験され得る。

【0109】

リアルタイムPCR (Gibsonら、Genome Research 6:995~1001, 1996; Heidら、Genome Research 6:986~994, 1996を参照のこと) は、増幅の間PCR産物蓄積のレベルを評価する技術である。この技術は、多数のサンプルにおけるmRNAのレベルの定量的評価を可能にする。要するに、mRNAは、腫瘍および正常組織から抽出され、そしてcDNAが標準的技術を用いて調製される。

10

【0110】

リアルタイムPCRは、例えば、ABI 7700 Prismか、またはGeneAmp (登録商標) 5700配列決定システム (PE Biosystems, Foster City, CA) のいずれかで実施され得る。この7700システムは、一端で5' 蛍光レポーター色素を、そしてもう一方の端に3' クエンチャー色素を有する特異的プローブ (TaqmanTM) と組み合わせて、フォワードプライマーおよびリバースプライマーを使用する。5' - 3'ヌクレアーゼ活性を有するTaq DNAポリメラーゼを用いてリアルタイムPCRを実施する場合、このプローブは切断され、そして蛍光の増大によって反応がモニターされることを可能にする蛍光を発光しはじめる (リアルタイム)。この5700システムは、蛍光色素であるSYBR (登録商標) グリーンを使用する。この蛍光色素は、7700装置と同様、二本鎖DNA、ならびに同じフォワードプライマーおよびリバースプローブにのみ結合する。マッチするプライマーおよび蛍光プローブは、プライマー発現プログラム (PE Biosystems, Foster City, CA) に従って設計され得る。プライマーおよびプローブの至適濃度は、当業者によって最初に決定される。コントロール (例えば、 β -アクトリン) プライマーおよびプローブは、例えば、Perkin Elmer / Applied Biosystems (Foster City, CA) から市販され得る。

20

【0111】

サンプルにおける特異的RNAの量を定量するため、目的の遺伝子を含有するプラスミドを用いて標準曲線を作成する。標準曲線は、リアルタイムPCRにおいて決定したCt値を用いて作成する。この値は、このアッセイで使用した最初のcDNA濃度に関係する。目的の遺伝子の $10^1 \sim 10^6$ コピーにわたる標準希釈が一般に十分である。さらに、コントロール配列について標準曲線を作成する。これによって、比較目的で、コントロールの量に対して、組織サンプルの最初のRNA含量の標準化が可能になる。

30

【0112】

上記に従って、そしてさらに以下に記載されるように、本発明は、配列番号1、3、5~7、11、13、15、17、19~24、30、32、および73~76に記載の配列を有する、代表的な乳房組織特異的、および/または胸部腫瘍特異的なポリヌクレオチドである、マンマグロビン (mamaglobin)、リポフィリンB、GABA (B899P)、B726P、B511S、B533S、B305D、およびB311D、癌、より詳細には、乳癌の検出において適切に使用され得る、配列番号2、4、8~10、12、14、16、18、25~29、および31および77に記載のアミノ酸配列を有する、上記のポリヌクレオチドによってコードされる代表的ポリペプチド、を提供する。

40

【0113】

本明細書において開示される方法はまた、前立腺癌、乳癌、結腸癌、卵巣癌、肺癌、頭頸部癌、リンパ腫、白血病、黒色腫、肝癌、胃癌、腎癌、膀胱癌、膵臓癌、および子宮内膜癌が挙げられるがこれらに限定されない広範な癌の検出に適切であるさらなるポリヌクレオチドおよび/または代替的ポリヌクレオチドの同定を可能にする。

50

【 0 1 1 4 】

(癌の検出方法)

一般的に、癌細胞は、患者から得られた生物学的サンプル（例えば、血液、リンパ節、骨髓、血清、精液、尿、および／または腫瘍生検）の細胞内における１つ以上の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。言い換えれば、このようなポリヌクレオチドは、癌（例えば、乳癌など）の存在または非存在を示すマーカーとして用いられ得る。

【 0 1 1 5 】

本明細書中にさらに詳細に議論されるように、本発明は、これらおよび他の関連する目的を、１つより多くのポリヌクレオチドの同時検出のための方法論を提供することによって達成し、このポリヌクレオチドの存在は、生物学的サンプルにおける癌細胞の存在の診断である。本明細書中に開示される種々の癌検出方法の各々は、共通して、１つ以上のオリゴヌクレオチドプライマーおよび／またはプローブをハイブリダイズする工程を有し、このハイブリダイゼーションは、腫瘍特異的および／または組織特異的なポリヌクレオチドの存在の実証である。意図される正確な用途に依存して、ゲノムDNAのポリヌクレオチドに対して効力がない、イントロンにまたがる１つ以上のオリゴヌクレオチドを利用することが好ましく、従って、これらのオリゴヌクレオチドは、生物学的サンプル中のゲノムDNAの検出を実質的に減少および／または排除する際に有効である。

10

【 0 1 1 6 】

さらに、試験されるべき生物学的サンプルを、１つ以上の細胞捕獲方法および／または細胞枯渇方法に供することによって、これらの検出方法の感度を増大するための方法が、本明細書中に開示される。

20

【 0 1 1 7 】

本発明の特定の実施形態において、以下の工程を用いることによって、患者における癌細胞の存在が決定され得る：（a）この患者から生物学的サンプルを得る工程；（b）この生物学的サンプルを、第１のポリヌクレオチドにハイブリダイズする第１のオリゴヌクレオチドと接触させる工程であって、この第１のポリヌクレオチドは、配列番号７３および配列番号７４において示されるポリヌクレオチドからなる群から選択される、工程；（c）この生物学的サンプルを、配列番号１、３、５～７、１１、１３、１５、１７、１９～２４、３０、３２、および７５からなる群から選択される第２のポリヌクレオチドにハイブリダイズする第２のオリゴヌクレオチドと接触させる工程；（d）このサンプル中で、このオリゴヌクレオチドの少なくとも１つにハイブリダイズする一定量のポリヌクレオチドを検出する工程；ならびに（e）この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、予め決定したカットオフ値と比較し、そしてそれから、この患者における癌の存在または非存在を決定する工程。

30

【 0 1 1 8 】

本発明の代替的实施形態は、以下の工程を利用して、患者における癌細胞の存在を決定する方法を提供する：（a）この患者から生物学的サンプルを得る工程；（b）この生物学的サンプルを、第１のポリヌクレオチドにハイブリダイズする第１のオリゴヌクレオチドと接触させる工程であって、この第１のポリヌクレオチドは、配列番号７６に示される、工程；（c）この生物学的サンプルを、配列番号１、３、５～７、１１、１３、１５、１７、１９～２４、３０、３２、および７５からなる群から選択される第２のポリヌクレオチドにハイブリダイズする第２のオリゴヌクレオチドと接触させる工程；（d）このサンプル中で、このオリゴヌクレオチドの少なくとも１つにハイブリダイズする一定量のポリヌクレオチドを検出する工程；ならびに（e）この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、予め決定したカットオフ値と比較し、そしてそれから、この患者における癌の存在または非存在を決定する工程。

40

【 0 1 1 9 】

本発明の他の実施形態は、患者における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。このような方法は、以下の工程を包含する：（a）この患者から生物学的サン

50

ルを得る工程；(b)患者から得られたこの生物学的サンプルを、配列番号73、配列番号74および配列番号76に示されるポリヌクレオチドからなる群から選択されるポリヌクレオチド配列にハイブリダイズする第1のオリゴヌクレオチドと接触させる工程；(c)この生物学的サンプルを、配列番号75に示されるポリヌクレオチドにハイブリダイズする第2のオリゴヌクレオチドと接触させる工程；(d)この生物学的サンプルを、配列番号5、配列番号6および配列番号7に示されるポリヌクレオチドからなる群から選択されるポリヌクレオチドにハイブリダイズする第3のオリゴヌクレオチドと接触させる工程；(e)この生物学的サンプルを、配列番号13、配列番号15、配列番号17、配列番号19、配列番号20、配列番号21、配列番号22、配列番号23および配列番号24に示されるポリヌクレオチドからなる群から選択されるポリヌクレオチドにハイブリダイズする第4のオリゴヌクレオチドと接触させる工程；(f)この生物学的サンプル中で、このオリゴヌクレオチドの少なくとも1つにハイブリダイズする一定量のポリヌクレオチドを検出する工程；ならびに(g)この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、予め決定したカットオフ値と比較し、そしてそれから、この患者における癌の存在または非存在を決定する工程。

10

【0120】

アッセイ条件下で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ヌクレオチド長、および好ましくは少なくとも20ヌクレオチド長の胸部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%、および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は、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、上記で規定されるような中程度のストリンジェント条件下で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効に利用され得るオリゴヌクレオチドプライマーは、好ましくは、少なくとも10~40ヌクレオチド長である、好ましい実施形態において、このオリゴヌクレオチドプライマーは、配列番号1、3、5~7、11、13、15、17、19~24、30、32および73~76に列挙される配列を有するDNA分子のうち、少なくとも10個連続したヌクレオチド、より好ましくは少なくとも15個連続したヌクレオチドを含む。PCRベースの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)。

20

30

【0121】

本発明はまた、患者における癌細胞の存在を検出するための、増幅ベースの方法を提供する。例示的な方法は、以下の工程を包含する：(a)生物学的サンプルを患者から得る工程；(b)この生物学的サンプルを、第1のオリゴヌクレオチド対と接触させる工程であって、この第1の対は、第1のオリゴヌクレオチドおよび第2のオリゴヌクレオチドを含み、ここで、この第1のオリゴヌクレオチドおよび第2のオリゴヌクレオチドは、それぞれ、第1のポリヌクレオチドおよびその相補体にハイブリダイズする、工程；(c)この生物学的サンプルを、第2のオリゴヌクレオチド対と接触させる工程であって、この第2の対は、第3のオリゴヌクレオチドおよび第4のオリゴヌクレオチドを含み、ここで、この第3および第4のオリゴヌクレオチドは、それぞれ、第2のポリヌクレオチドおよびその相補体にハイブリダイズし、そしてここで、この第1のポリヌクレオチドは、ヌクレオチド配列においてこの第2のポリヌクレオチドに対して無関係である、工程；(d)第1のポリヌクレオチドおよび第2のポリヌクレオチドを増幅する工程；ならびに(e)増幅された第1のポリヌクレオチドおよび増幅された第2のポリヌクレオチドを検出する工程；ここで、増幅された第1のポリヌクレオチドまたは増幅された第2のポリヌクレオチドの存在が、この患者における癌細胞の存在を示す。

40

【0122】

50

本発明に従う方法は、広範な種々の生物学的サンプル（例えば、血液、血清、リンパ節、骨髓、痰、尿および腫瘍生検サンプル）から得られるポリヌクレオチドの同定のために適切である。特定の好ましい実施形態において、生物学的サンプルは、血液、リンパ節または骨髓のいずれかである。本発明の他の実施形態において、リンパ節は、歩哨リンパ節であり得る。

【0123】

本発明の方法は、広範な種々の癌の検出において利用され得ることが明らかである。例示的な癌としては、以下が挙げられるがこれらに限定されない：前立腺癌、乳癌、結腸癌、卵巣癌、肺癌、頭部および頸部の癌、リンパ腫、白血病、黒色腫、肝臓癌、胃癌、腎癌、膀胱癌、膵臓癌および子宮内膜癌。

10

【0124】

本発明の特定の例示的な実施形態は、検出されるべきポリヌクレオチドが、マンマグロブリン（mammaglobin）、リポフィリンB、GABA（B899P）、B726P、B511S、B533S、B305DおよびB311Dからなる群から選択される方法を提供する。あるいは、および/またはさらに、検出されるべきポリヌクレオチドは、配列番号73、配列番号74、配列番号75、配列番号1、配列番号3、配列番号5、配列番号6、配列番号7、配列番号11、配列番号13、配列番号15、配列番号17、配列番号19、配列番号20、配列番号21、配列番号22、配列番号23、配列番号24、配列番号30、配列番号32、および配列番号76に記載されるポリヌクレオチドからなる群から選択され得る。

20

【0125】

本発明の方法に従って使用され得る、適切な例示的なオリゴヌクレオチドプローブおよび/またはプライマーは、配列番号33～35および63～72によって本明細書中に開示される。ゲノムDNAのバックグラウンドの検出を排除する、特定の好ましい実施形態において、このオリゴヌクレオチドは、イントロンにまたがるオリゴヌクレオチドであり得る。本明細書中に開示される種々のポリヌクレオチドの検出に適切な、例示的なイントロンにまたがるオリゴヌクレオチドは、配列番号36～62に示される。

【0126】

意図される正確な適用に依存して、当業者は、放射性標識を検出し、そして蛍光体を検出することによって、組織特異的および/または腫瘍特異的なポリヌクレオチドの検出を選び得る。より詳細には、このオリゴヌクレオチドプローブおよび/またはプライマーは、例えば、放射性標識および/または発蛍光団のような検出可能な部分を含み得る。

30

【0127】

あるいは、またはさらに、本発明の方法はまた、組織特異的および/または腫瘍特異的ポリヌクレオチドの検出の前に、分画する工程（例えば、電気泳動によって）を包含し得る。

【0128】

他の実施形態において、本明細書中に記載の方法は、癌の進行のモニターに関して用いられ得る。これらの実施形態により、癌の診断に提供されるアッセイは、経時的に行われ得、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が評価され得る。例えば、このアッセイは、6ヶ月～1年の期間にわたり、24～72時間ごとに行われ得、それ以降は、必要時に行われ得る。一般に、癌は、検出されるポリペプチドまたはポリヌクレオチドのレベルが時間がたつにつれて増大する患者において進行している。対照的に、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か、または時間とともに減少しているかのいずれかである場合は、癌は進行していない。

40

【0129】

特定のインビボ診断アッセイは、腫瘍に対して直接行われ得る。1つのこのようなアッセイは、腫瘍細胞を結合薬剤と接触させる工程を包含する。次いで、この結合した結合薬剤は、レポーター基を介して直接的または間接的に検出され得る。このような結合薬剤はまた、組織学的適用において用いられ得る。あるいは、ポリヌクレオチドプローブがこのよ

50

うな適用内で用いられ得る。

【0130】

上記のように、感度を改善するために、複数の胸部腫瘍タンパク質マーカーが所定のサンプル内でアッセイされ得る。本明細書中で提供される異なるタンパク質に特異的な結合薬剤が1つのアッセイ内で組み合わせられ得ることは明らかである。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、最適感度を生じる組み合わせを決定するための慣用的な実験に基づき得る。さらに、または代わりに、本明細書中に提供される腫瘍タンパク質についてのアッセイは、他の公知の腫瘍抗原についてのアッセイと組み合わせられ得る。

【0131】

(細胞富化)

本発明の他の局面において、細胞捕捉技術は、本明細書中に開示される種々の検出方法論の感度を改善するために、ポリヌクレオチド検出の前に用いられ得る。

【0132】

例示的な細胞富化方法論は、細胞表面マーカーに対する特異的モノクローナル抗体もしくは4量体抗体複合体でコーティングされた免疫磁性ビーズを用い、この方法論は、サンプル中の癌細胞を最初に富化するか、もしくはサンプル中の癌細胞をポジティブに選択するために用いられ得る。種々の市販のキットが用いられ得る。これらのキットとしては、Dynabeads (登録商標) Epithelial Enrich (Dynal Biotech, Oslo, Norway)、StemSepTM (Stem Cell Technologies, Inc., Vancouver, BC)、およびRosetteSep (Stem Cell Technologies) が挙げられる。当業者は、他の容易に利用可能な方法論およびキットが適切に用いられて、所望の細胞集団が富化またはポジティブに選択され得ることを認識する。

【0133】

Dynabeads (登録商標) Epithelial Enrichは、正常上皮組織および新生物上皮組織上で発現された2つの糖タンパク質膜抗原に特異的なmAbでコーティングされた磁性ビーズを含む。このコーティングされたビーズをサンプルに添加し得、次いで、このサンプルを磁石に接触させ得、それによりこのビーズに結合した細胞を捕捉し得る。所望でない細胞は洗い流され、磁石により単離された細胞は、ビーズから溶出され、さらなる分析において用いられる。

【0134】

RosetteSepは、血液サンプルから直接細胞を富化するために用いられ得、そして種々の所望でない細胞を標的とし、そしてサンプル中に存在する赤血球(RBC)上のグリコホリンAにそれらを架橋して、ロゼットを形成する4量体の抗体のカクテルからなる。Ficoll上で遠心分離すると、標的とした細胞は、遊離のRBCとともにペレットを形成する。

【0135】

枯渇カクテルにおける抗体の組み合わせは、どの細胞が除去され、そして同時にどの細胞が回収されるかを決定する。利用可能な抗体としては、以下が挙げられるが、これらに限定されない: CD2、CD3、CD4、CD5、CD8、CD10、CD11b、CD14、CD15、CD16、CD19、CD20、CD24、CD25、CD29、CD33、CD34、CD36、CD38、CD41、CD45、CD45RA、CD45RO、CD56、CD66B、CD66e、HLA-DR、IgE、およびTCR。さらに、本発明において、胸部腫瘍抗原に特異的なmAbが開発され得、そして同様な様式で用いられ得ることが意図される。例えば、腫瘍特異的細胞表面抗原に結合するmAbは、磁性ビーズに結合体化され得るか、または4量体抗体複合体に処方され得、そしてサンプルから転移性の胸部腫瘍細胞を富化するか、またはポジティブに選択するために用いられ得る。

【0136】

10

20

30

40

50

一旦サンプルが富化されるか、またはポジティブに選択されると、細胞はさらにアッセイされ得る。例えば、これらの細胞は、溶解され得、そしてRNAが単離され得る。次いで、RNAは、本明細書中に記載のようにリアルタイムPCRアッセイにおいて胸部腫瘍特異的多重プライマーを用いてRT-PCR分析に供され得る。

【0137】

本発明の別の局面において、細胞捕捉技術は、リアルタイムPCRとともに用いられて、胸部腫瘍抗原を発現する転移性細胞の検出のためのより感度の高いツールを提供し得る。骨髓サンプル、末梢血サンプル、および少量の針吸引サンプル中の乳癌細胞の検出は、乳癌患者における診断および予後のために望ましい。

【0138】

10

(プローブおよびプライマー)

上記のように、および本明細書中でさらに詳細に記載されるように、本発明に従う特定の方法、組成物およびキットは、癌の検出のために2以上のオリゴヌクレオチドプライマー対を利用する。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により、これらの核酸プローブは生物学的サンプル中の相補的配列の存在を検出する際に有用になる。

【0139】

あるいは、他の実施形態において、本発明のプローブおよび/またはプライマーは、核酸ハイブリダイゼーションを介した検出のために用いられ得る。このようにして、検出されるポリヌクレオチドの15ヌクレオチドの長さの連続する配列と同じ配列を有するか、またはこの配列に相補的な、少なくとも約15ヌクレオチドの長さの連続する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。より長い連続する同一の配列または相補的な配列、例えば、約20、30、40、50、100、200、500、1000の配列(全ての中間の長さを含む)および全長までの長さの配列すらもまた、特定の実施形態において有用である。

20

【0140】

検出されるポリヌクレオチドに同一またはこのポリヌクレオチドに相補的な、10~14、15~20、30、50、または100~200ヌクレオチド程度(中間の長さも同様に含む)の連続するヌクレオチドのストレッチからなる配列領域を有するオリゴヌクレオチドプライマーは、例えば、ポリメラーゼ連鎖反応(PCRTM)のような増幅反応において使用するためのプライマーとして特に意図される。これは、例えば、血液、リンパ節および骨髓のような異なる生物学的サンプルにおいて、ポリヌクレオチドを分析することを可能にする。

30

【0141】

約15~25ヌクレオチド長のプライマーを使用することにより、安定かつ選択的な二重鎖分子の形成が可能になる。15塩基を超える長さのストレッチにわたって連続する相補的配列を有する分子が一般に好ましいが、ハイブリッドの安定性および選択性を増大するために、そしてこれにより得られた特定のハイブリッド分子の質および程度を改善する。当業者は、一般に、15~25の連続するヌクレオチド、または所望される場合よりさらに長い遺伝子相補的ストレッチを有するプライマーを設計することを好む。

40

【0142】

検出されるポリヌクレオチドの任意の部分から、プライマーが選択され得る。当業者がプライマーとして利用しようとする、配列番号1、3、5~7、11、13、15、17、19~24、30、32、73~75(それぞれ、図3~6)および配列番号76(リポフィリンB)に示される例示的ポリヌクレオチド、またはこれらの配列の任意の連続する部分、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)のような配列を検討することのみが必要とされる。プライマー配列の選択は、種々の因子により支配され得る。例えば、当業者は、完全な配列の末端に対するプライマーを用いようとし得る。本明細書中に開示される例示的プライマーは、必要に応じて、目的のポリヌクレオチド全体(例えば、配列番号1、3、5~7、11、13、15、17、19~24、3

50

0、32、73～75（それぞれ、図3～6）および配列番号76（リポフィリンB）に示される配列）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。

【0143】

本発明はさらに、配列番号33～71に示される、癌の検出のための本発明の方法に従って使用され得る（本明細書中でさらに詳細に記載される）種々の例示的オリゴヌクレオチドプライマーおよびプローブのヌクレオチド配列を提供する。

【0144】

本発明に従うオリゴヌクレオチドプライマーは、当業者に一般に利用可能な方法（例えば、自動化オリゴヌクレオチド合成機を用いて一般に行われるように、化学的手段によりプライマーを直接合成することを含む）により慣用的に容易に調製され得る。想定される適用に依存して、当業者は、代表的には、種々の条件のハイブリダイゼーションを用いて、標的配列に対するプローブの種々の程度の選択性を達成することを望む。高い選択性を要する適用に関して、当業者は、代表的には、ハイブリッドを形成するために比較的ストリンジェントな条件を用いることを望む（例えば、当業者は、約50～70の温度で、約0.02M～約0.15Mの塩の塩濃度により提供されるような、比較的低い塩条件および/または高い温度条件を選択する）。このような選択条件は、存在するとすれば、プローブとテンプレートまたは標的鎖との間のわずかなミスマッチを許容し、そして関連する配列を単離するために特に適切である。

【0145】

（ポリヌクレオチド増幅技術）

ガンの検出のために本明細書に概説される具体的な実施形態の各々は、1つ以上のオリゴヌクレオチドプライマーおよび/またはプローブのハイブリダイゼーションを介する、組織および/または腫瘍特異的ポリヌクレオチドの検出において共通する。生物学的サンプル中に存在するガン細胞の相対数および/または各ガン細胞内でのポリヌクレオチドの発現レベルのような因子に依存して、検出の工程を行う前に、増幅工程を行うことが好ましい。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が、生物学的サンプルに由来する胸部腫瘍cDNAの一部を増幅するためにポリメラーゼ連鎖反応（PCR）に基づくアッセイにおいて利用され得る。ここで、少なくとも1つのオリゴヌクレオチドプライマーは、胸部腫瘍タンパク質をコードするポリヌクレオチドに対して特異的である（すなわち、ハイブリダイズする）。その増幅されたcDNAは、必要に応じて、例えば、ゲル電気泳動のような分画工程に供され得る。

【0146】

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。最も良く知られている増幅方法の1つは、米国特許第4,683,195号、同第4,683,202号、および同第4,800,159号（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）に詳述されるポリメラーゼ連鎖反応（PCRTM）である。簡単には、PCRTMにおいて、標的配列の反対の相補鎖上の領域に相補的である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トリホスフェートが、DNAポリメラーゼ（例えば、Taqポリメラーゼ）とともに反応混合物中に添加される。その標的配列がサンプル中に存在する場合、そのプライマーは、その標的に結合し、そしてそのポリメラーゼにより、プライマーがその標的配列に沿って、ヌクレオチドを付加することによって伸長する。反応混合物の温度を上げることおよび下げることによって、その伸長プライマーは、標的から解離して反応産物を形成し、過剰のプライマーは、標的およびその反応産物に結合し、そしてそのプロセスが繰り返される。好ましくは、逆転写およびPCRTM増幅手順が、増幅されたmRNAの量を定量するために行われ得る。ポリメラーゼ連鎖反応法は、当該分野で周知である。

【0147】

ポリヌクレオチド増幅のための1つの好ましい方法は、RT-PCRを利用し、これにおいてPCRが、逆転写と組み合わせて適用される。代表的には、RNAが生物学的サン

10

20

30

40

50

ル（例えば、血液、血清、リンパ節、骨髄、痰、尿、および腫瘍生検サンプル）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使用するPCR増幅は、cDNA分子を生成し、これは、例えば、ゲル電気泳導を使用して分離されて、可視化され得る。増幅は、患者およびガンに罹患していない個体由来の生物学的サンプル上で行われ得る。その増幅反応は、2オーダーの大きさにわたるcDNAの数倍希釈に対して行われ得る。数倍希釈の試験患者サンプルにおいて、非ガンサンプルの同じ希釈に比較して、発現の2倍以上の増大は、代表的には陽性と考えられる。

【0148】

種々の商業的に利用可能なキットのいずれもが、増幅工程を行うために使用され得る。1つのこのような増幅技術は、制限酵素を使用してその遺伝子の既知の領域においてフラグメントを生成する逆PCR (Trigliaら、Nuc l . A c i d s . R e s . 16 : 8186, 1988) である。次いで、そのフラグメントは、分子内ライゲーションによって環状化され、そして既知の領域に由来する広範なプライマーを使用するPCRのためのテンプレートとして使用される。代替のアプローチにおいて、部分配列に隣接する配列は、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用いて増幅することによって回収され得る。その増幅された配列は、代表的には、同じリンカープライマーおよびその既知の領域に特異的な二次プライマーを用いて二回目の増幅に供される。この手順に対する改変（その既知の配列から反対方向に伸長を開始する2つのプライマーを利用する）は、WO96/38591に記載される。別のこのような技術は、「cDNA末端の迅速増幅」または「RACE」として公知である。この技術は、既知の配列の5'および3'である配列を同定するために、poly A領域またはベクター配列にハイブリダイズする内部プライマーおよび外部プライマーの使用を含む。さらなる技術には、捕捉PCR (Lagerstrom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) が含まれる。増幅を利用する他の方法は、全長cDNA配列を得るためにも利用され得る。

【0149】

増幅のための別の方法は、Eur . Pat . Appl . Publ . No . 320,308（その全体が本明細書において特に参考として援用される）に開示されるリガーゼ連鎖反応 (LCR と呼ばれる) である。LCRにおいて、2つの相補的なプローブ対が調製され、そして標的配列の存在下で、各対は、それらが隣接するように、標的の反対側の相補鎖に結合する。リガーゼの存在下において、その2つのプローブ対は、単一のユニットを形成するために連結する。PCR^{T M}と同様に、温度サイクリングによって、結合したリガンドユニットは、その標的から解離し、次いで、過剰のプローブ対のライゲーションのための「標的配列」として作用する。米国特許第4,883,750号（その全体が参考として本明細書に援用される）は、標的配列に対してプローブ対を結合するためのLCRに類似する代替の増幅方法を記載する。

【0150】

PCT国際特許出願公開PCT/US87/00880（その全体が参考として本明細書に援用される）に記載されるQbeta Replicaseもまた、本発明においてなお別の増幅方法として使用され得る。この方法において、標的の複製的配列に相補的な領域を有するRNAの複製的配列が、RNAポリメラーゼの存在下でサンプルに添加される。そのポリメラーゼは、次いで検出され得る複製的配列を複写する。

【0151】

制限エンドヌクレアーゼおよびリガーゼが、制限部位の1つの鎖のヌクレオチド5'-[- チオ] トリホスフェートを含む標的分子の増幅を達成するために使用される等温増幅法 (Walkerら、1992、その全体が本明細書中に参考として援用される) もまた、本発明における核酸の増幅において有用であり得る。

【0152】

鎖置換増幅 (SDA) は、多数回の鎖置換および合成（すなわち、ニクトランスレーシ

10

20

30

40

50

ョン)を含む核酸の等温増幅を行う別の方法である。修復連鎖反応(RCR)と呼ばれる同様の方法は、本発明において有用であり得る増幅の別の方法であり、増幅のために標的化された領域にわたる数種のプローブをアニーリングすること、および、続いて、4つの塩基のうちの2つのみが存在する修復反応を包含する。他の2つの塩基は、簡単な検出のためのビオチン化誘導体として添加され得る。同様のアプローチがSDAにおいて使用される。

【0153】

配列はまた、環状プローブ反応(CPR)を使用して検出され得る。CPRにおいて、非標的DNAの3'および5'配列およびその標的タンパク質特異的RNAの内部または「中間」配列を有するプローブが、サンプル中に存在するDNAにハイブリダイズする。ハイブリダイゼーションの際、その反応は、RNaseHで処理され、そしてそのプローブの産物は、消化後に放出されるシグナルを生成することによって明確な産物として同定される。そのオリジナルのテンプレートは、別のサイクリングプローブにアニーリングし、そしてその反応が繰り返される。このようにして、CPRは、標的遺伝子特異的に発現された核酸に対するプローブのハイブリダイゼーションによって生成されたシグナルを増幅することを包含する。

10

【0154】

Great Britain Pat. Appl. No. 2202328およびPCT/US89/01025(これらの各々はその全体が参考として本明細書中にて援用される)に記載されるなお他の増幅法は、本発明に従って使用され得る。前者の出願では、「改変された」プライマーは、PCR様のテンプレートおよび酵素依存的合成において使用される。そのプライマーは、捕捉部分(例えば、ビオチン)および/または検出部分(例えば、酵素)で標識することによって改変され得る。後者の出願では、過剰の標識されたプローブがサンプルに添加される。標的配列の存在下で、そのプローブが結合し、そして触媒により切断される。切断後、その標的配列が過剰のプローブによって結合されるようにインタクトのまま放出される。その標識されたプローブの切断は、その標的配列の存在の指標である。

20

【0155】

他の核酸増幅手順には、転写ベースの増幅システム(TAS)(Kwohら、1989; PCT国際特許出願公開WO88/10315、その全体が本明細書中で参考として援用される)が含まれ、これには、核酸配列ベースの増幅(NASBA)および3SRが含まれる。NASBAでは、核酸が、サンプルの標準的なフェノール/クロロホルム抽出、熱変性、溶解緩衝液での処理、およびDNAおよびRNAの単離のためのミニスパンカラム、またはRNAの塩化グアニジウム抽出による増幅のために調製され得る。これらの増幅技術には、その標的配列に特異的な配列を有するプライマーのアニーリングが含まれる。ポリマー化後、DNA/RNAハイブリッドは、RNaseHで消化され、一方、二本鎖DNA分子が、再び熱変性される。いずれかの場合において、その単一鎖DNAは、第二の標的特異的プライマーの付加およびその後のポリマー化によって完全に二本鎖にされる。次いで、その二本鎖DNA分子は、T7またはSP6のようなポリメラーゼによって多重に転写される。等温サイクリック反応において、そのRNAは、DNAに逆転され、そしてT7またはSP6のようなポリメラーゼによって再び転写される。得られた産物は、短縮型であろうと完全型であろうと、標的特異的配列を示す。

30

40

【0156】

その全体が本明細書によって参考として援用される、欧州特許出願番号329,822は、一本鎖RNA(「ssRNA」)、ssDNA、および二本鎖DNA(「dsDNA」)(これらは、本発明に従って使用され得る)を周期的に合成する工程を包含する、核酸増幅プロセスを開示する。ssRNAは、第一のプライマーオリゴヌクレオチドのための第一のテンプレートであり、これは、逆転写酵素(RNA依存性DNAポリメラーゼ)によって伸長される。次いで、このRNAは、生じたDNA:RNA二本鎖から、リボヌクレアーゼH(RNaseH、DNAまたはRNAのいずれかとの二本鎖におけるRNAに

50

特異的であるRNAse)の作用によって、除去される。生じたssDNAは、第二のプライマーのための第二のテンプレートであり、これはまた、そのテンプレートに対する相同性に対して、RNAポリメラーゼプロモーター(T7 RNAプロモーターによって例示される)5'の配列を含む。次いで、このプライマーは、DNAポリメラーゼ(E. coliのDNAポリメラーゼIの大きい「クレノウ」フラグメントによって例示される)によって伸長され、二本鎖DNA(「dsDNA」)分子として生じ、プライマー間の元のRNAの配列と同一の配列を有し、そしてさらに、一端でプロモーター配列を有する。このプロモーター配列は、適切なRNAポリメラーゼによって使用されて、DNAの多くのRNAコピーを作製し得る。次いで、これらのコピーは、再びサイクルに入って、非常に迅速な増幅を生じ得る。酵素の適切な選択によって、この増幅は、各サイクルでの酵素の添加なしに等温的になされ得る。このプロセスの周期的な性質に起因して、開始配列は、DNAまたはRNAのいずれかの形態であるように選択され得る。

10

【0157】

本明細書中でその全体が参考として援用されるPCT公開特許出願番号WO89/06700は、標的一本鎖DNA(「ssDNA」)に対するプロモーター/プライマー配列のハイブリダイゼーションに基づいた核酸配列増幅スキーム、それに続くこの配列の多くのRNAコピーの転写を開示する。このスキームは、周期的ではない。すなわち、新しい温度が、生じたRNA転写物から生成されない。他の増幅方法は、当業者に周知の「RACE」(Frohman、1990)および「一面(one-side)PCR」(Ohara、1989)を含む。

20

【0158】

(癌の検出のための組成物およびキット)

本発明はさらに、上記の診断方法のいずれかにおける使用のためのキットを提供する。このようなキットは、代表的には、診断アッセイの実行のために必要な2つ以上の成分を含む。この成分は、化合物、試薬、容器および/または装置であり得る。例えば、キット内の1つの容器は、乳癌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、上記のような支持体材料に付着されて提供され得る。1つ以上のさらなる容器は、アッセイにおいて使用される構成要素(例えば、試薬または緩衝液)を封入する。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のために適切なレポーター基を含む、上記のような検出試薬を含む。

30

【0159】

本発明はまた、本発明の検出方法の実行のために適切なキットを提供する。例示的なキットは、その各々が別個の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対を含む。特定の実施形態において、本発明に従ったキットはまた、核酸ポリメラーゼおよび適切な緩衝液を含み得る。本発明のキットのために適切な、例示的なオリゴヌクレオチドプライマーは、配列番号33~71によって本明細書中に開示される。本発明のキットに適切な例示的なオリゴヌクレオチドプライマーは、配列番号73、配列番号74、配列番号75、配列番号1、配列番号3、配列番号5、配列番号6、配列番号7、配列番号11、配列番号13、配列番号15、配列番号17、配列番号19、配列番号20、配列番号21、配列番号22、配列番号23、配列番号24、配列番号30、配列番号32、およびリポフィリンBに開示される。

40

【0160】

あるいは、キットは、生物学的なサンプル中の乳癌タンパク質をコードするmRNAのレベルを検出するために設計され得る。このようなキットは一般に、乳癌タンパク質をコードするポリペプ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プローブまたはプライマーを含む。このようなオリゴヌクレオチドが、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使用され得る。このようなキット内に存在し得るさらなる成分は、第二のオリゴヌクレオチドおよび/もしくは診断試薬、または乳癌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ための容器を含む

50

。

【0161】

他の関連する局面において、本発明はさらに、本明細書中に開示された方法において有用な組成物を提供する。例示的な組成物としては、その各々が別個のポリヌクレオチドに特異的にハイブリダイズする、2つ以上のオリゴヌクレオチドプライマー対を含む。本発明の組成物のために適切な例示的なオリゴヌクレオチドプライマーは、配列番号33～71によって本明細書中に開示される。本発明の組成物に適切な例示的なポリヌクレオチドは、配列番号73、配列番号74、配列番号75、配列番号1、配列番号3、配列番号5、配列番号6、配列番号7、配列番号11、配列番号13、配列番号15、配列番号17、配列番号19、配列番号20、配列番号21、配列番号22、配列番号23、配列番号24、配列番号30、配列番号32、およびリポフィリンBに開示される。

10

【0162】

以下の実施例は、例示のために提供され、限定のためではない。

【0163】

(実施例)

(実施例1)

(ディファレンシャルディスプレイ)

本実施例は、乳癌において過剰発現するポリヌクレオチドを富化するためのディファレンシャルディスプレイの使用を開示する。

【0164】

ディファレンシャルディスプレイを、文献(例えば、その全体が本明細書によって参考として援用される、Liang, Pら、Science 257:9687-971(1993))を参照のこと)中に記載のように、以下の改変と共に実施した:(a)PCR増幅産物を、銀染色ゲル上で可視化した(b)組織の遺伝的に適合した対を使用して、多型改変体を除外した(c)cDNAの2つの異なる希釈を、テンプレートとして使用して、任意の希釈の影響を除外した(その全体が本明細書によって参考として援用される、Mou, Eら、Biochem Biophys Res Commun. 199:564-569(1994))を参照のこと)。

20

【0165】

(実施例2)

(cDNAサブトラクションライブラリーの調製)

本実施例は、乳癌特異的なポリヌクレオチドが富化された、乳癌cDNAサブトラクションライブラリーの調製を開示する。

30

【0166】

cDNAライブラリーサブトラクションを、いくつかの改変と共に、記載されたように実行した。その全体が本明細書によって参考として援用される、Hara, T.ら、Blood 84:189-199(1994))を参照のこと。3つの乳癌のプールから作製した、乳癌ライブラリー(トレーサー)を、乳癌特異的な遺伝子を同定するために、正常な乳房ライブラリー(ドライバー)で差し引いた。より最近の差し引きは、共通の遺伝子をより効率的に差し引くために、1つの「免疫学的」組織(例えば、脾臓、リンパ節、またはPBMC)に沿った重要な組織上の強調と共に、ドライバーとして6～10の正常組織を使用して、腫瘍におけるリンパ球浸潤由来のcDNAの除去において評価した。この乳癌特異的な差し引かれたcDNAライブラリーを、以下のように作製した:ドライバーcDNAライブラリーをEcoRI、NotI、およびSfuI(SfuIは、ベクターを切断する)で消化し、DNAポリメラーゼクレノウフラグメントで満たした。フェノール-クロロホルム抽出およびエタノール沈殿の後、DNAをPhotoprobeビオチンで標識してH₂Oに溶解させた。トレーサーcDNAライブラリーを、BamHIおよびXhoIで消化し、フェノール-クロロホルムで抽出し、Chromaspin-400カラムを通し、エタノールで沈殿させ、そしてハイブリダイゼーションのために68で20時間(長時間ハイブリダイゼーション(LH))ドライバーDNAと混合した。次

40

50

いで、この反応混合物を、合計4回のストレプトアビジン処理、続くフェノール/クロロホルム抽出に供した。差し引いたDNAを沈殿し、そして再びドライバーDNAと682時間のハイブリダイゼーション(短時間ハイブリダイゼーション(SH))に供した。ピオチン化二本鎖DNAの除去後に、差し引いたcDNAを、クロラムフェニコール耐性pBCSK⁺のBamHI/XhoI部位に連結し、そしてElectroMax E.coli DH10B細胞にエレクトロポレーションによって形質導入して、差し引いたcDNAライブラリーを生成した。多くはない乳癌特異的遺伝子をクローニングするために、ドライバーcDNAライブラリーとこれら多くのcDNAとを有するトレーサーcDNAライブラリーを初期サブトラクションから差し引くことによって、cDNAライブラリーサブトラクションを反復した。これは、これら多くの配列の欠失および多くはない配列を含むサブトラクションライブラリーの生成を生じた。

10

【0167】

差し引かれたcDNAライブラリーを分析するために、プラスミドDNAを、100~200の独立したクローンから調製した。これらのクローンは、差し引きされたライブラリーから無作為に選択され、そしてDNA配列決定によって、特徴付けられた。決定されたcDNAおよび単離されたcDNAについての予測アミノ酸配列を、最も新しいGenbankデータベースおよびヒトESTデータベースを使用して公知の配列と比較した。

【0168】

(実施例3)

(PCRサブトラクション)

20

本実施例は、胸部腫瘍特異的ポリヌクレオチドを富化するためのPCRサブトラクションを開示する。

【0169】

PCRサブトラクションを、実質的に文献に記載されるように実施した。Diatchenko, L.ら、Proc Natl Acad Sci USA, 93:6025-6030(1996)およびYang, G. P.ら、Nucleic acid Res. 27:1517-23(1999)(これらの全体が参考として本明細書中に援用される)を参照のこと。簡単には、この型のサブトラクションは、2つの異なるアダプターを制限酵素消化したテスター(胸部腫瘍)cDNAサンプルの異なるアリコートに連結し、それに続いてこれらのテスターを過剰なドライバー(アダプターなし)と別々に混合することによって作用する。この第一のハイブリダイゼーションは、一本鎖テスター特異的cDNAの正規化をもたらし、これは、ハイブリダイゼーションの二番目の反応速度論に起因する。次いで、これらの別個のハイブリダイゼーション反応を変性なしに混合し、そして第二のハイブリダイゼーションを行ない、標的分子;異なるアダプターの両方を含む二本鎖cDNAフラグメントを産生する。2回のPCRを行い、これは標的集団分子(正規化されたテスター特異的cDNA)の対数的増幅をもたらすが、他のフラグメントは、増幅しないかまたは直線的様式で増幅したのみのいずれかであった。実施したサブトラクションは、正常乳房のプールを差し引かれた胸部腫瘍のプールおよびPBMC、脳、脾臓、肝臓、小腸、胃、心臓および腎臓を含む正常組織のプールを差し引かれた胸部腫瘍のプールを含んだ。

30

40

【0170】

cDNA合成の前に、RNAを、RNasin(Promega Biotech, Madison, WI)存在下で、DNase I(Ambion)で処理し、DNAの夾雑物を除去した。リアルタイムPCR組織パネルにおける利用のためのcDNAを、superscript II逆転写酵素(Gibco BRL, Bethesda, MD)を含む25μlのOligo dT(Boehringer-Mannheim)プライマーを使用して調製した。

【0171】

(実施例4)

(乳房特異的抗原を使用する胸部腫瘍の検出)

50

乳房特異的抗原 B 5 1 1 S および B 5 3 3 S の単離および特徴づけは、米国特許出願 0 9 / 3 4 6 , 3 2 7 (1 9 9 9 年 7 月 2 日出願) に開示される。この開示は、その全体が参考として本明細書中に援用される。B 5 1 1 S について決定した c D N A 配列を、配列番号 3 0 に示し、対応するアミノ酸配列を、配列番号 3 1 に示す。B 5 3 3 S について決定された c D N A 配列を、配列番号 3 2 に示す。乳房特異的抗原 B 7 2 6 P の単離および特徴づけは、米国特許出願 0 9 / 2 8 5 , 4 8 0 (1 9 9 9 年 4 月 2 日出願) および同 0 9 / 4 3 3 , 8 2 6 (1 9 9 9 年 1 1 月 3 日出願) に開示され、この開示は、その全体が参考として本明細書中に援用される。

【 0 1 7 2 】

B 7 2 6 P のスプライシング改変体について決定した c D N A 配列を、配列番号 1 3 、 1 5 、 1 7 および 1 9 ~ 2 4 に示し、対応するアミノ酸配列を、配列番号 1 4 、 1 6 、 1 8 および 2 5 ~ 2 9 に示す。 10

【 0 1 7 3 】

乳房特異的抗原 B 3 0 5 D (A および C を形成する) の単離および特徴づけは、米国特許出願 0 9 / 4 2 9 , 7 5 5 (1 9 9 9 年 1 0 月 2 8 日出願) に記載されており、この開示はその全体が参考として本明細書中に援用される。B 3 0 5 D アイソフォーム A および C について決定した c D N A 配列を、配列番号 1 、 3 および 5 ~ 7 に示し、対応するアミノ酸配列を、配列番号 2 、 4 および 8 ~ 1 0 に示す。

【 0 1 7 4 】

乳房特異的抗原 B 3 1 1 D の単離および特徴づけは、米国特許出願 0 9 / 2 8 9 , 1 9 8 (1 9 9 9 年 4 月 9 日出願) に記載されており、この開示はその全体が参考として本明細書中に援用される。B 3 1 1 D について決定した c D N A 配列を、配列番号 1 1 に示し、対応するアミノ酸配列を配列番号 1 2 に示す。 20

【 0 1 7 5 】

マンマグロビンについての c D N A 配列を図 4 および 5 に示し、G A B A についての c D N A 配列を、図 6 に示し、そしてそれぞれ配列番号 7 3 ~ 7 5 に開示する。

【 0 1 7 6 】

乳房特異的抗原リポフィリン B の単離および特徴づけは、米国特許出願 0 9 / 7 8 0 , 8 4 2 (2 0 0 1 年 2 月 8 日出願) に記載されており、この開示は、その全体が参考として本明細書中に援用される。リポフィリン B について決定した c D N A 配列を、配列番号 7 6 に示し、対応するアミノ酸配列を、配列番号 7 7 に示す。リポフィリン B のいくつかの配列改変体のヌクレオチド配列がまた、米国特許出願 0 9 / 7 8 0 , 8 4 2 に記載される。 30

【 0 1 7 7 】

(実施例 5)

(マイクロアレイ分析)

本実施例は、正常乳房組織サンプルと比較した場合に、胸部腫瘍組織サンプルにおいて、少なくとも 2 倍は過剰発現されるポリヌクレオチドを同定するためのマイクロアレイ分析の使用を開示する。

【 0 1 7 8 】

目的のポリヌクレオチドの m R N A の発現を以下のように実施した。異なる遺伝子についての c D N A を上記のように調製し、そしてスライドガラス (I n c y t e , P a l o A l t o , C A) 上に整列させた。次いで、整列した c D N A を、C y 3 または C y 5 で蛍光標識した第一鎖 c D N A (胸部腫瘍由来の p o l y A + R N A から得た) 、正常乳房および正常組織と他の腫瘍 (その全体が本明細書中で参考として援用される S h a l o n , D r a , G e n o m e R e s . 6 : 6 3 9 ~ 4 5 (1 9 9 6) に記載されている) の 1 : 1 混合物を用いてハイブリダイズした。代表的には、C y 3 (プローブ 1) を胸部腫瘍由来の c D N A に結合させ、そして C y 5 (プローブ 2) を正常乳房組織または他の正常組織に結合させた。両方のプローブを、チップ上で不死化遺伝子特異的 c D N A と競合させ、洗浄し、次いでハイブリダイゼーションの程度を決定するために個々の C y 3 およ 40 50

びCy5の蛍光の蛍光強度をスキャンした。データをGEMTOOLSソフトウェア(Incyte, Palo Alto, CA)(Cy3/Cy5の比によって正常組織と比較されるべき胸部腫瘍の過剰発現パターンを可能にする)を使用して解析した。蛍光強度はまた、個々の遺伝子の発現レベルに関連した。

【0179】

DNAマイクロアレイ分析を、主にスクリーニング道具として使用して、ディファレンシャルディスプレイ、cDNAライブラリーおよびPCRサブトラクションから回収されたcDNAの組織/腫瘍特異性を決定し、その後、定量的RT-PCR、ノーザンブロットイング、および免疫組織化学によるより厳密な解析を行った。マイクロアレイ分析を、2つのマイクロチップ上で実施した。合計で3603の差し引かれたcDNAテンプレートおよび197のディファレンシャルディスプレイテンプレートを評価し、定量的PCRによるさらなる解析のため40の候補を同定した。これらの候補のうちいくつかを、好ましい組織特異性プロファイル(B305D、B311D、B726P、B511SおよびB533Sを含む)に基づいて選択し、胸部腫瘍および/または正常乳房 対 他の正常組織におけるそれらの過剰発現プロファイルを示した。これらの遺伝子の発現が、胸部腫瘍および/または乳房組織に対する高度の特異性を示すことが明らかであった。さらに、これらの遺伝子は、多くの場合、相補的発現プロファイルを有する。

10

【0180】

また、2つの公知の乳房特異的遺伝子、マンマグロビンおよび - アミノ酪酸A型レセプターのサブユニット(GABA)を、マイクロアレイ分析に供した。マンマグロビンのmRNA発現は、胸部腫瘍を含む乳房組織の増殖においてアップレギュレーションされることが以前に記載されている。Watsonら、Cancer Res., 56:860~5(1996); Watsonら、Cancer Res., 59:3028~3031(1999); Watsonら、Oncogene, 16:817~24(1998)(これらは、その全体が参考として本明細書中に援用される)を参照のこと。GABA mRNAレベルを、胸部腫瘍において過剰発現した。以前の研究は、膀胱におけるその過剰発現、ならびに前立腺および肺におけるある程度の過剰発現を実証した(Hedblomら、J Biol. Chem., 272:15346~15350(1997))が、以前の研究では、胸部腫瘍での増大したレベルは示されていなかった。

20

【0181】

(実施例6)

(定量的リアルタイムPCR分析)

本実施例は、正常な胸部組織サンプルと比較して、胸部腫瘍組織サンプルにおいて少なくとも2倍過剰発現するポリヌクレオチドのマイクロアレイ同定を確認するための定量的リアルタイムPCRの使用を開示する。

30

【0182】

本明細書において実施例5で開示されたマイクロアレイ分析により同定されたポリヌクレオチドの腫瘍特異性および/または組織特異性を、定量的PCR分析によって確認した。胸部転移、胸部腫瘍、良性胸部障害および正常な胸部組織を、他の正常組織および腫瘍と共に、定量的(リアルタイム)PCRにおいて試験した。これを、ABI 7700 PrismまたはGeneAmp(登録商標)5700 sequence detection system(PE Biosystems, Foster City, CA)のいずれかにおいて実施した。7700システムは、順方向プライマーおよび逆方向プライマーを、この順方向プライマーとこの逆方向プライマーとの間の配列にアニーリングするように設計された特異的プローブと組み合わせて使用する。このプローブを、その5'末端で蛍光レポーター色素と結合体化し、そして他方の3'末端で消光剤色素(TaqmanTM)と結合体化した。PCRの間、Taq DNAポリメラーゼ(その5'-3'ヌクレアーゼ活性を有する)はプローブを切断し、これにより蛍光を発し始め、蛍光の増加によって反応をモニターすることを可能にする(リアルタイム)。Hollandら、Proc Natl Acad Sci U S A, 88:7276-7280(1991)

40

50

91)。5700システムは、二本鎖DNAに対してのみ結合する蛍光色素であるSYBR（登録商標）グリーン（Schneebergerら，PCR Methods Appl. 4:234-8（1995））、ならびに7700機器と同じ順方向プライマーおよび逆方向プライマーを使用した。プローブは必要とされなかった。一致するプライマーおよび蛍光プローブを、Primer Express program（PE Biosystems，Foster City，CA）に従って、各遺伝子について設計した。

【0183】

【表1】

表1.
目的の遺伝子についてのプライマーおよびプローブ配列

	順方向プライマー	逆方向プライマー	プローブ
マンマグロビン	TGCCATAGATG AATTGAAGGAA TG (SEQ ID NO:48)	TGTCATATATTAA TTGCATAAACACC TCA (SEQ ID NO:49)	TCTTAACCAAACG GATGAAACTCTGA GCAATG (SEQ ID NO:50)
B305D-C形態	AAAGCAGATGG TGGTTGAGGTT (SEQ ID NO:39)	CCTGAGACCAAA TGGCTTCTTC (SEQ ID NO:40)	ATTCCATGCCGGC TGCTTCTTCTG (SEQ ID NO:41)
B311D	CCGCTTCTGACA ACACTAGAGAT C (SEQ ID NO:63)	CCTATAAAGATGT TATGTACCAAAA ATGAAGT (SEQ ID NO:64)	CCCCTCCCTCAGG GTATGGCCC (SEQ ID NO:65)
B726P	TCTGGTTTTCTC ATTCTTTATTCA TTTATT (SEQ ID NO:42)	TGCCAAGGAGCG GATTATCT (SEQ ID NO:43)	CAACCACGTGACA AACACTGGAATTA CAGG (SEQ ID NO:44)
B533S	CCCTTCTCACC CACACACTGT (SEQ ID NO:66)	TGCATTCTCTCAT ATGTGGAAGCT (SEQ ID NO:67)	CCGGGCCTCAGGC ATATACTATTCTA CTGTCTG (SEQ ID NO:68)
GABA π	AAGCCTCAGAG TCCTTCCAGTAT G (SEQ ID NO:36)	AAATATAAGTGA AGAAAAAATTA GTAGAT (SEQ ID NO:72)	AATCCATTGTATC TTAGAACCAGGG ATTTGTTTAGA (SEQ ID NO:38)
B511S	GACATTCCAGTT TTACCCAAATG G (SEQ ID NO:69)	TGCAGAAGACTC AAGCTGATTCC (SEQ ID NO:70)	TCTCAGGGACACA CTCTACCATTCGG GA (SEQ ID NO:71)

定量的PCRにおいて、マンマグロビン（mammaglobin）、GABA、B305D-C形態、B311D、B511S、B533SおよびB726Pについての順方向プライマーについて使用された濃度は、それぞれ、900、900、300、900、900、300および300nMであった。逆方向プライマーについて、これらはそれぞれ、300、900、900、900、300、900および900nMであった。このようにして生成されたプライマーおよびプローブを、リアルタイムPCRにおいて、汎用熱サイクルプログラム（universal thermal cycling program）で使用した。これらを滴定し、チェッカー盤アプローチを使用して最適濃度を決定した。標的腫瘍由来のcDNAのプールを、この最適化プロセスにおいて使用した。反応を、25 μ l容量で実施した。すべての場合において、最終プローブ濃度は160nMであった。dATP、dCTPおよびdGTPは、0.2mMであり、そしてdUTPは、0.4mMであった。Amplitaq goldおよびAmperase UNG（PE Biosystems，Foster City，CA）を、1回の反応あたり0.625ユニットおよび0.25ユニットで使用した。MgCl₂は、5mMの最終濃度であった。微量のグリセロール、ゼラチン、およびTween 20（Sigma C

10

20

30

40

50

hem Co, St Louis, MO) を添加して、反応を安定化させた。各反応は、
2 μ l の希釈テンプレートを含んだ。上記のように調製された RT 反応からの cDNA を
、目的の遺伝子について 1 : 10 で希釈し、そして - アクチンについて 1 : 100 で希
釈した。 - アクチン (PE Biosystems, Foster City, CA)
についてのプライマーおよびプローブを、同様の様式で使用して、サンプル中の - アク
チンの存在を定量した。SYBR (登録商標) グリーンアッセイの場合、反応混合物 (2
5 μ l) は、2.5 μ l の SYBR グリーン緩衝液、2 μ l の cDNA テンプレート、な
らびに目的の遺伝子についての順方向プライマーおよび逆方向プライマー (各々 2.5 μ
l) を含んだ。この混合物はまた、3 mM MgCl₂、0.25 ユニットの AmpErase
UNG、0.625 ユニットの AmpliTaq gold、0.08 % グリセ
ロール、0.05 % ゼラチン、0.0001 % Tween 20 および 1 mM dNTP
混合物を含んだ。両方の形式において、40 サイクルの増幅を実施した。

10

20

30

40

【0184】

サンプル中の特定の cDNA (従って、最初の mRNA) の量を定量するために、目的の
遺伝子を含むプラスミドを使用した各実行について、検量線を作成した。検量線を、アッ
セイに使用された最初の cDNA 濃度と関連する、リアルタイム PCR で決定された Ct
値を使用して作成した。20 ~ 2 x 10⁶ コピーの範囲の目的遺伝子の基準希釈 (sta
ndard dilution) を、この目的のために使用した。さらに、一定量の -
アクチンに対する正規化を可能にするために、200 fg ~ 2000 pg の範囲のハウス
キーピング遺伝子 - アクチンについて検量線を作成した。これにより、各遺伝子で観察
される過剰発現レベルの評価が可能となった。

【0185】

遺伝子 B311D、B533S および B726P を、以下からなる 2 つの異なるパネルに
おいて、上記で表 1 において示されるプライマーおよびプローブを使用し、上記のような
定量的 PCR において評価した：(a) 胸部腫瘍、胸部正常および正常組織；ならびに (b)
胸部腫瘍転移 (主に、リンパ節)。パネル (a) についてのデータを、3 つすべての
遺伝子について図 1 に示す。この 3 つの遺伝子は、同一の胸部組織発現プロフィールを示
した。しかし、遺伝子発現の相対的レベルは、各場合において非常に異なった。一般的に
B311D は、B533S よりも低いレベルで発現され、そしてこの両方とも、B726
P よりも少なかったが、3 つすべてが胸部組織に制限されていた。従って、定量的 PCR
によって、3 つすべての遺伝子について、正常胸部組織と胸部腫瘍との間で示差的な発現
が存在すること、および約 50 % の胸部腫瘍が、これらの遺伝子を過剰発現することが確
認された。乳癌に由来する遠位転移のパネルにおいて試験される場合、3 つすべての遺伝
子は 14 / 21 の転移と反応し、そして同様のプロフィールを提示した。3 つすべての遺
伝子はまた、図 2 において B533S について示されたように、腫瘍の病理学的段階の関
数として発現レベルの増加を示した。

【0186】

マンマグロピンは、ウサギのウテログロビンおよびラットのステロイド結合タンパク質サ
ブユニット C3 のホモログであり、そして高度にグリコシル化された低分子量タンパク質
である。そのホモログとは対照的に、マンマグロピンは、胸部特異的であることが報告さ
れており、そして転移性乳癌患者由来の 91 % のリンパ節におけるマンマグロピン mRNA
発現の検出の報告と共に、胸部腫瘍生検 (23 %) ならびに原発性および転移性の胸部
腫瘍 (約 75 %) における過剰発現が記載されている。しかし、上記のようなマイクロア
レイおよび定量的 PCR によるマンマグロピン遺伝子発現のより綿密な分析 (パネル (a)
および (b)、ならびに他の腫瘍および正常組織ならびにさらなる胸部腫瘍のパネル)
により、皮膚および唾液腺における有意なレベルでの発現と共に、食道および気管におけ
るはるかに低レベルでの発現が示された (以下の表 2 に示される)。

【0187】

【表 2】

表 2

種々の組織におけるマンマグロビンおよびB511SのmRNAの正規化分布

組織	平均コピー マンマグロビン /ng β -アクトニン \pm SD	PCR 陽性	平均コピー-B511S /ng β -アクトニン \pm SD	PCR 陽性	PCR陽性 (マンマグロビン /B511S)
胸部腫瘍	1233.88 \pm 3612.74	31/42	1800.40 \pm 3893.24	33/42	38/42
胸部腫瘍 転移	1912.54 \pm 4625.85	14/24	3329.50 \pm 10820.71	14/24	17/24
良性 乳房障害	121.87 \pm 78.63	3/3	524.66 \pm 609.43	2/3	3/3
正常胸部	114.19 \pm 94.40	11/11	517.64 \pm 376.83	8/9	11/11
胸部縮小 (breast reduction)	231.50 \pm 276.68	2/3	482.54 \pm 680.28	1/2	2/3
他の腫瘍	0.13 \pm 0.65	1/39	24.17 \pm 36.00	5/23	
唾液腺	435.65 \pm 705.11	2/3	45766.61 \pm 44342.43	3/3	
皮膚	415.74 \pm 376.14	7/9	7039.05 \pm 7774.24	9/9	
食道	4.45 \pm 3.86	2/3	1.02 \pm 0.14	0/3	
気管支	0.16	0/1	84.44 \pm 53.31	2/2	
他の 正常組織	0.33 \pm 1.07	0/85	5.49 \pm 10.65	3/75	

10

20

30

40

胸部特異的遺伝子B511Sは、胸部腫瘍と正常胸部組織上ではマンマグロビン(mam maglobin)に対して異なる反応性のプロフィールを有するが、マンマグロビンと同じ正常組織のサブセットと反応した。P S O R T分析によって、B511Sは、90アミノ酸のORFを有することおよびマンマグロビンの場合のように分泌タンパク質であることが示された。B511Sは、膜貫通ドメインの証拠を有さなかったが、切断可能なシグナル配列を含み得る。マンマグロビンは、遠位転移性胸部腫瘍24個のうち14個、胸部腫瘍42個のうち31個を検出し、そして正常な胸部組織と比較して、腫瘍および転移において10倍の過剰発現を示した。マンマグロビン発現に関しては本質的に陰性であった試験した他の陰性の正常組織および腫瘍に対して、正常な胸部組織において少なくとも300倍の過剰発現が存在した。B511Sは、胸部腫瘍42個のうち33個および遠位転移24個のうち14個を検出したが、B511Sとマンマグロビンとの組合せは、胸部腫瘍42個のうち38個および転移性病変24個のうち17個を検出することが予想される(表2、上記)。B511Sおよびマンマグロビンの定量的な発現レベルはまた、これら2つの遺伝子に関して観察されたマイクロアレイプロフィールに従って、類似した範囲であった。マンマグロビンと付加的であった他の遺伝子を、表3に示す。

【0188】

【表3】

表 3
マンマグロビンと他の遺伝子の mRNA 補完性

	マンマグロビン 陽 性	マンマグロビン陰性				
		B305D	GABA π	B726P	B305D + GABA π	B305D + GABA π + B726P
胸部 転 移	13/21	2/8	5/8	3/8	7/8	8/8
胸部腫瘍	18/25	3/7	4/7	5/7	7/7	7/7

10

B 3 0 5 D は、正常組織（正常胸部組織を含む）と比較して、胸部腫瘍、前立腺腫瘍、正常前立腺組織および精巣において高度に過剰発現されたことが示された。B 3 0 5 D の異なるスプライス改変体が、最も豊富な形態 A および形態 C で同定されたが、試験した全ては、定量的 P C R において類似の組織プロファイルを有する。A 形態および C 形態は、それぞれ、3 2 0 アミノ酸および 3 8 5 アミノ酸の O R F を含む。B 3 0 5 D は、P S O R T によって、一連のアンキリン反復を含む I I 型膜タンパク質であることが予想された。胸部腫瘍において B 3 0 5 D と補完性であることが示される既知遺伝子は、G A B A であつた。この遺伝子は、G A B A_A レセプターファミリーのメンバーであり、ファミリーの他のメンバーと 3 0 ~ 4 0 % のアミノ酸相同性を有するタンパク質をコードし、そして

ノーザンブロット分析によって肺、胸腺および前立腺において低めのレベルで過剰発現され、子宮において高度に過剰発現されることが示された。胸部組織におけるその発現は、以前に記載されていない。ニューロン組織において認識可能な発現を有することは、他の G A B A_A レセプターと対照的である。この遺伝子の組織発現プロファイリングは、この遺伝子が、B 3 0 5 D 遺伝子と逆の関係で胸部腫瘍において過剰発現されることを示した（表 3）。G A B A は、腫瘍 2 5 個のうちの 1 5 個および転移 2 1 個のうちの 6 個（マンマグロビンによっては見落とされた 4 個の腫瘍および 5 個の転移を含む）を検出した。対照的に、B 3 0 5 D は、胸部腫瘍 2 5 個のうちの 1 3 個および転移 2 1 個のうちの 8 個（これもまた、マンマグロビンによっては見落とされた 3 個の腫瘍および 2 個の転移を含む）を検出した。ちょうど B 3 0 5 D および G A B A の組合せは、胸部腫瘍 2 5 個のうちの 2 2 個および転移 2 1 個のうちの 1 4 個を同定すると予想される。胸部転移を検出する場合の B 3 0 5 D および G A B A とマンマグロビンとの組合せを、上記の表 3 ならびに図 3 A および 3 B に示す。この組合せは、胸部転移 2 1 個のうちの 2 0 個ならびに胸部腫瘍 2 5 個のうちの 2 5 個を検出した（これは、3 つ全ての遺伝子に対して同じパネル上で評価された）。これらの 3 つの遺伝子に関して陰性であつた 1 つの胸部転移は、B 7 2 6 P に関して強く陽性であつた（図 3 A および 3 B）。

20

30

【 0 1 8 9 】

循環している腫瘍細胞の存在を評価するために、免疫捕獲（細胞捕獲）方法を、R T - P C R 分析の前に上皮細胞について最初に富化するために使用した。上皮細胞表面上の 2 つの糖タンパク質に対する特異的なモノクローナル抗体を用いてコーティングされた免疫磁気ポリスチレンビーズを、この目的のために使用した。このような富化手順は、表 4 に示されるように、ポリ A⁺ R N A の直接的な単離と比較した場合、検出感度を増加させた（約 1 0 0 倍）。

40

【 0 1 9 0 】

【 表 4 】

表4

全血中にスパイクされたマンマグロビン陽性系細胞(MB415)
の検出およびリアルタイムPCRによる検出

MB415系細胞/ml血液	上皮系細胞検出 (ポリA ⁺ RNA)	直接検出 (ポリA ⁺ RNA)
	マンマグロビン/ngβ-アクトン	
100000	54303.2	58527.1
10000	45761.9	925.9
1000	15421.2	61.6
100	368.0	5.1
10	282.0	1.1
1	110.2	0
0	0	0

10

マンマグロビン陽性細胞(MB415)を、種々の濃度中に全血にスパイクし、次いで、上皮細胞富化または血液からの直接的な単離のいずれかを用いて検出した。富化手順を用いて、マンマグロビンmRNAは、直接的な単離が使用された場合よりもはるかに低いレベルで検出されることが見られた。転移性乳癌を有する患者由来の全血サンプルを、引き続き免疫磁気ビーズを用いて処理した。次いで、ポリA⁺RNAを単離し、cDNAを調製し、そして2つの遺伝子特異的プライマー(表1)および蛍光プローブ(TaqmanTM)を用いて定量的PCR中で実行させた。乳癌組織において観察されるように、補完性がまた、乳癌由来の循環している腫瘍細胞の検出において見出された。また、マンマグロビンPCRは、正常な血液サンプルと比較した場合、転移性乳癌患者由来の血液サンプルの高い割合において(20/32)(低いレベルではあるが)循環している腫瘍細胞を検出したが(表5)、現在までに試験されたいくつかの他の遺伝子はさらに、この検出率を増加した。これには、B726P、B305D、B311D、B533SおよびGABAが挙げられる。転移性乳癌患者由来の血液サンプルにおけるマンマグロビンの検出レベルは、おそらく、本研究における上皮マーカー特異的富化の使用に起因して、より少ない血液容積を試験しているのにもかかわらず、以前に記載されたものよりも高い(62対49%)。試験した遺伝子全ての組み合わせは、サンプル32個のうち27個が、これらの遺伝子の1つ以上によって陽性であったことを示す。

20

30

【0191】

【表5】

表5
正常な個体および転移性乳癌患者の血液から単離された
上皮細胞における遺伝子発現

サンプルID	マンマグロビン	B305D	B311D	B533S	B726P	GABAR	Combo
2	+	-	-	+	-	-	+
3	+	-	-	+	-	-	+
5	+	+	+	-	+	-	+
6	+	-	+	+	+	-	+
8	-	-	+	-	-	-	+
9	+	+	+	-	+	-	+
10	+	-	+	-	+	-	+
11	-	-	-	-	-	-	-
12	+	+	+	-	-	-	+
13	-	-	-	+	-	-	+
15	-	-	-	-	-	-	-
18	+	-	-	-	-	+	+
19	+	-	-	-	-	-	+
21	+	-	-	-	-	-	-
22	-	-	-	-	-	-	+
23	+	-	-	-	-	-	+
24	+	-	-	-	-	-	+
25	-	+	-	-	-	-	+
26	-	-	-	-	-	-	+
29	+	-	+	+	+	-	+
31	+	-	-	+	-	-	+
32	-	-	-	-	-	±	±
33	-	-	-	-	+	-	+
34	+	-	-	+	-	+	+
35	+	-	-	-	+	-	+
36	-	-	-	-	-	+	+
37	+	-	-	+	-	-	+
38	-	-	-	-	-	-	-
40	+	-	-	-	-	-	+
41	+	-	-	+	-	-	+
42	+	-	-	-	-	-	+
43	-	-	-	-	-	+	+
ドナー104	-	-	-	-	-	Nd	-
ドナー348	-	-	-	-	-	Nd	-
ドナー392	-	-	-	-	-	Nd	-
ドナー408	-	-	-	-	-	-	-
ドナー244	-	-	-	-	-	-	-
ドナー355	-	-	-	-	-	-	-
ドナー264	-	-	-	-	-	Nd	-
ドナー232	-	-	-	-	-	-	-
ドナー12	-	-	-	-	-	Nd	-
ドナー415	-	-	-	-	-	-	-
ドナー35	-	-	-	-	-	-	-
総数	20/32	4/32	7/32	9/32	7/32	4/32	27/32

10

20

30

40

さらなる研究において、マンマグロビン、GABA、B305D(C形態)およびB726Pに特異的なプライマーならびに特異的なTaqmanプローブを、異なる組み合わせで使用して、リアルタイムPCRを用いて胸部転移(B.met)サンプルおよび胸部腫瘍(B.tumor)サンプルにおけるその含まれたmRNA発現プロファイルを分析した。マンマグロビン、B305D(C形態)およびB726Pに対して使用した順方向および逆方向プライマーならびにプローブを、表1に示す。GABAに対して使用した順方向プライマーおよびプローブを、表1に示し、逆方向プライマーは、以下である：TTCAAATATAAGTGAAGAAAAAATTAG-TAGATCAA(配列番号51)。以下の表6に示されるように、マンマグロビン、GABA、B305C(C形態)およびB726Pの組み合わせは、胸部腫瘍サンプル22個のうち22個を検出することが見出され、5サンプルにおいて発現の増加が見られた(++によって示す)。

【0192】

【表6】

表6

異なるプライマーの組み合わせを用いた腫瘍サンプルのリアルタイムPCR検出

腫瘍サンプル	マンマグロビン	マンマグロビン+ GABA	マンマグロビン+ GABA + B305D	マンマグロビン+GABA+ B305D + B726P
B. Met 316A		+	+	+
B. Met 317A	+	+	+	++
B. Met 318A		+	+	+
B. Met 595A	+	+	+	+
B. Met 611A	+	+	+	+
B. Met 612A	+	+	+	+
B. Met 614A		+	+	+
B. Met 616A		+	+	+
B. Met 618A	+	+	+	+
B. Met 620A	+	+	+	+
B. Met 621A	+	+	+	+
B. Met 624A	+	+	+	+
B. Met 625A			+	+
B. Met 627A	+		+	+
B. Met 629A		+	+	+
B. Met 631A	+	+	+	++
B. Tumor 154A	+	+	+	++
B. Tumor 155A	+	+	+	++
B. Tumor 81D			+	+
B. Tumor 209A		+	+	++
B. Tumor 208A		+	+	+
B. Tumor 10A	+	+	+	+

10

20

特異的プライマーの添加によるメッセージシグナルの増加がさらに、4つの腫瘍サンプル(B. met 316A、B. met 317A、B. tumor 81DおよびB. tumor 209A)を用いる1プレート実験において実証された。

【0193】

胸部腫瘍サンプルおよび正常組織サンプルのパネルにおけるマンマグロビン、GABA、B305D(C形態)およびB726Pの組み合わせの発現がまた、Taqmanプロブアプローチの代わりにSYBR緑色検出系を用いるリアルタイムPCRを用いて検出された。この系を用いて得た結果を、図7に示す。

【0194】

(実施例7)

(乳癌患者の末梢血における定量的PCR)

この研究において評価した既知遺伝子は、マンマグロビンおよびアミノ酪酸A型レセプターサブユニット(GABA)であった。乳癌において過剰発現される新規遺伝子を同定するために、本発明者らは、改良されたバージョンのディファレンシャルディスプレイRT-PCR(DDPCR)技術(Liangら、Science 257:967-971(1993); Mouら、Biochem Biophys Res Commun. 199:564-569(1994)); cDNAライブラリー抽出方法(Haraら、Blood 84:189-199(1994))およびPCRサブトラクション(Diatchenkoら、Proc Natl Acad Sci U S A.、93:6025-6030(1996); Yangら、Nucleic Acids Res. 27:1517-23(1999))を使用した。

【0195】

ディファレンシャルディスプレイは、B305DおよびB311Dと称する2つのcDNAフラグメントの回収をもたらした(Houghtonら、Cancer Res. 40:Abstract#217,32-33(1999))。B511SおよびB533Sは、cDNAライブラリーサブトラクションアプローチ(原稿準備中)を使用して単離した2つのcDNAフラグメントであるが、B726P cDNAフラグメントは、PCRサブトラクションに由来した(Jainら、Proceedings of the Amer Assoc Cancer Res. 40:Abstract#216,32

30

40

50

(1999); Xuら、Proceedings of the Amer Assoc Cancer Res. 40: Abstract #2115, 319 (1999); および Moleshera, Proceedings of Amer Assoc Cancer Res. 41: Abstract #4330, 681 (2000)。

【0196】

3つの新規な遺伝子、B311D、B533SおよびB726Pは、定量的PCR分析により同一の乳房組織発現プロファイルを示した。これらの遺伝子を、以下からなる2つの別個のパネルに対する定量的PCRにおいて評価した：(a)胸部腫瘍、乳房正常組織および正常な組織、ならびに(b)胸部腫瘍転移の群(主にリンパ節)。使用するプライマーおよびプローブを表1に示す。パネル(a)についてのデータを、3つすべての遺伝子について図2に示す。概して、発現プロファイルは同等であり、そして同じランクにあるが、発現のレベルは、かなり異なる。B311Dは一般に、B533Sよりも低いレベルで発現され、そしてその両方が、B726Pより低い。3つすべてが、乳房組織に限定された。3つすべての配列を使用して、Genbankデータベースに対して検索した。B311D配列およびB533S配列の両方が、異なる反復配列を含み、そしていずれについてもORFは同定されていない。B726Pは新規な遺伝子であり、mRNAスプライシングは、いくつかの異なる推定ORFを生じる。

10

【0197】

この定量的PCRにより、正常な乳房組織と胸部腫瘍との間に異なるmRNAが存在し、胸部腫瘍の約50%がこれらの遺伝子を過剰発現することが、確認された。乳癌に由来する一群の遠隔転移に対して試験した場合、3つすべての遺伝子が、転移21個のうちの14個と反応し、そして類似するプロファイルを示した(データ示さず)。興味深いことに、前立腺癌群に対して試験した場合、3つすべての遺伝子が、前立腺腫瘍24個のうちの同じ3つを同定したが、乳房よりも大いに低い発現レベルであった。この群の遺伝子は、B533Sについて示されるのと同様にその腫瘍の病理学的段階の関数として、漸増レベルの発現を示した。

20

【0198】

マイクロアレイおよび定量的PCRによるマンマグロビン遺伝子発現のより厳密な分析は、皮膚および唾液腺において有意なレベルの発現を示し、そして食道および気管において大いに低いレベルの発現を示した。B511Sは、マンマグロビンと比較した場合に胸部腫瘍および正常乳房組織に対してわずかに異なる反応性プロファイルを有したが、マンマグロビンと類似する正常組織の部分集合と反応した。マンマグロビンは、遠隔転移性胸部腫瘍24個のうちの14個を検出し、胸部腫瘍42個のうちの31個を検出し、そして正常な乳房組織と比較した場合に、腫瘍および転移において10倍過剰発現を示した。正常乳房組織対他のネガティブな正常組織および試験した腫瘍において、少なくとも300倍のマンマグロビン過剰発現が存在した。B511Sは、胸部腫瘍42個のうちの33個および遠隔転移24個のうちの14個を検出した。B511Sとマンマグロビンとの組み合わせは、胸部腫瘍42個のうちの38個および転移性病変24個のうちの17個を検出すると推測される。B511Sおよびマンマグロビンの発現の定量レベルもまた、これらの2つの遺伝子について観察されたマイクロアレイプロファイルと一致して、類似した範囲にあった。

30

40

【0199】

特定の遺伝子が、マンマグロビンの発現プロファイルを補完した(すなわち、マンマグロビンが発現しなかった腫瘍において発現することが示された)。B305Dは、正常な組織(正常な乳房組織を含む)と比較した場合に、胸部腫瘍、前立腺腫瘍、正常な前立腺組織および精巣において、高度に過剰発現された。B305Dの異なるスプライス改変体を、最も豊富である形態Aおよび形態Cを用いて同定した。試験したすべての形態は、定量的PCRにおいて類似の組織プロファイルを有した。このA形態およびC形態は、それぞれ、320アミノ酸のORFおよび385アミノ酸のORFを含む。胸部腫瘍において、B305Dと相補的であることが示された公知の遺伝子は、GABAであった。この組

50

組織発現プロファイルは、ニューロン組織においてかなりの発現を代表的に有する、他の G A B A_A レセプターと対照的である。さらなる知見は、この遺伝子の組織発現プロファイルが、この遺伝子が、B 3 0 5 D 遺伝子に逆比例して胸部腫瘍において過剰発現されることを示したことであった(表 3)。G A B A は、腫瘍 2 5 個のうちの 1 5 個および転移 2 1 個のうちの 6 個を検出し、これらは、マンマグロビンにより見逃された 4 つの腫瘍および 5 つの転移を含んだ。対照的に、B 3 0 5 D は、胸部腫瘍 2 5 個のうちの 1 3 個および転移 2 1 個のうちの 8 個を検出し、これらはまた、マンマグロビンにより見逃された 3 つの腫瘍および 2 つの転移を含んだ。まさに B 3 0 5 D と G A B A との組み合わせは、胸部腫瘍 2 5 個のうちの 2 2 個および転移 2 1 個のうちの 1 4 個を同定すると推定される。この組み合わせは、3 つすべての遺伝子について同じ群に対して評価された、乳房転移 2 1 個のうちの 2 0 個、ならびに乳癌 2 5 個のうちの 2 5 個を検出した。これら 3 つの遺伝子についてネガティブであった 1 つの乳房転移は、B 7 2 6 P について強くポジティブであった。

10

【0200】

マイクロアレイ分析およびその後の定量的 P C R の使用により、乳房組織(腫瘍および正常)ならびに正常組織の両方における乳癌遺伝子の発現を性格に決定するため、そしてそれらの診断および治療能力を評価するための、方法論が提供される。5 つの新規な遺伝子および 2 つの既知の遺伝子を、これらの技術を使用して評価した。これらの遺伝子のうちの 3 つ、B 3 1 1 D、B 5 3 3 S および B 7 2 6 P は、一致した m R N A 発現を示した。そして集合的にこのデータは、転写制御のレベルでのこれら 3 つの遺伝子座の一致した発現と一致する。3 つすべての遺伝子が、胸部腫瘍対正常乳房組織における差次的発現を示した。そして過剰発現のレベルは、その腫瘍の病理学的段階に関連するようであった。マンマグロビンの場合、発現が、乳房組織から離れた他の組織において見出された。発現が、皮膚、唾液腺において観察され、そして気管においてかなり劣った程度に観察された。

20

【0201】

胸部腫瘍における G A B A の発現もまた、新規な知見であった。マンマグロビンとともに観察されたいくつかの遺伝子の発現が相補した場合、特に 2 つの遺伝子、B 3 0 5 D および G A B A が、マンマグロビン検出の診断感度を増加した。これらの 3 つの遺伝子の組み合わせは、評価した胸部腫瘍および転移 4 6 個のうち 4 5 個(97.8%)を検出した。B 7 2 6 P を含めると、2 5 個すべての胸部腫瘍および 2 1 個の遠隔転移の検出が可能であった。

30

【0202】

(実施例 8)

(免疫捕捉による循環する乳癌細胞の富化)

本実施例は、循環する乳癌細胞の富化のための免疫捕捉細胞捕捉方法論の使用により達成される、感度の増強を開示する。

【0203】

循環する腫瘍細胞の存在を評価するために、免疫捕捉方を採用して、R T - P C R 分析の前にまず、上皮細胞について富化した。上皮細胞を、免疫磁気ビーズ分離法(Dynal A . S , O s l o , N o r w a y) を用いて血液サンプルから富化した。この方法は、ヒト上皮細胞の表面上の糖ポリペプチド抗原に特異的なモノクローナル抗体でコートした磁気ビーズを使用する。適切な例示的細胞表面抗原は、例えば、M o m b u r g , F . ら、C a n c e r R e s . 4 1 : 2 8 8 3 ~ 9 1 (1 9 9 7) ; N a u m e , B . ら、J o u r n a l o f H e m o t h e r a p y . 6 : 1 0 3 ~ 1 1 3 (1 9 9 7) ; N a u m e , B . ら、I n t J C a n c e r . 7 8 : 5 5 6 ~ 6 0 (1 9 9 8) ; M a r t i n , V . M . ら、E x p H e m a t o l . , 2 6 : 2 5 2 ~ 6 4 (1 9 9 8) ; H i l d e b r a n d t , M . ら、E x p H e m a t o l . 2 5 : 5 7 ~ 6 5 (1 9 9 7) ; E a t o n , M . C . ら、B i o t e c h n i q u e s 2 2 : 1 0 0 ~ 5 (1 9 9 7) ; B r a n d t , B . ら、C l i n E x p M e t a s t a s e s 1 4 : 3 9 9 ~ 4 0 8 (1 9 9 6) に記載され、この各々が、本明細書中にその全体が参考として援用

40

50

される。この方法で単離された細胞を溶解し、そして磁気ビーズを除去した。その後、この溶解物を、 $Oligo(dT)_{25}$ でコートした磁気ビーズ (Dynabeads) を使用して、ポリA⁺ mRNA 単離について処理した。キット緩衝液中でこのビーズを洗浄した後、ビーズ/ポリA⁺ RNA サンプルを、最終的に 10 mM Tris HCl (pH 8) に懸濁し、そして逆転写に供した。その後、遺伝子特異的なプライマーおよびプローブを使用して、このRNAをリアルタイムPCRに供し、反応条件は、本明細書中で上記に概説したとおりであった。 - アクチン含量もまた決定し、そして正規化のために使用した。正常なサンプルの平均 + 3 標準偏差よりも大きい (目的の遺伝子 コピー) / (- アクチン ng) であるサンプルを、ポジティブであると見なした。血液サンプルに対するリアルタイムPCRを、もっぱら TaqmanTM 手順であるが 50 サイクルまで延長して使用して、実施した。 10

【0204】

富化手順を使用して、マンマグロビン mRNA が、直接単離を使用した場合よりも大いに低いレベルで検出可能であることが、見出された。転移性乳癌を伴う患者由来の全血サンプルを、その後、免疫磁気ビーズで処理し、その後、ポリA⁺ RNA を単離し、cDNA を作製し、そしてマンマグロビンに対する2つの遺伝子特異的プライマーおよび蛍光プローブ (TaqmanTM) を使用して、定量的PCRを実施した。乳癌組織において観察されたように、相補はまた、乳癌に由来する循環する腫瘍細胞の検出においても観察された。また、マンマグロビンPCRは、正常な血液サンプルと比較した場合に、低レベルであるにも関わらず、転移性乳癌からの高いパーセンテージの血液 (20 個 / 32 個) において循環する腫瘍細胞を検出した。今日まで試験した他の遺伝子のうちのいくつかは、この検出率をさらに増加し得；これは、B726P、B305D、B311D、B533S および GABA を包含する。試験した遺伝子すべての組み合わせは、32 個中 27 個のサンプルが、これらの遺伝子のうちの1つ以上によりポジティブであったことを示す。 20

【0205】

(実施例9)

(乳癌の多重検出)

さらなる多重リアルタイムPCRアッセイを、4つの乳癌特異的遺伝子：LipophilinB、Gaba (B899P)、B305D - C および B726P の発現を同時に検出するために確立した。単一乳癌特異的遺伝子の発現分析に頼る検出アプローチとは対照的に、このマルチプレックスアッセイは、試験された全ての胸部腫瘍サンプルを検出し得た。 30

【0206】

このマルチプレックスアッセイを、Mammaglobin の代わりに LipophilinB 発現を検出するように設計した。それらの類似する発現プロファイルに起因して、LipophilinB は、乳癌検出のためのこの多重PCRアッセイにおいて Mammaglobin を置換し得る。このアッセイを以下のようにして行なった：LipophilinB 特異的プライマー、B899P (Gaba) 特異的プライマー、B305D 特異的プライマー、および B726P 特異的プライマー、ならびに特異的 Taqman プローブを使用して、胸部腫瘍におけるそれらの組み合わせられた mRNA 発現プロファイルを分析した。これらのプライマーおよびプローブを以下に示す： 40

LipophilinB：正方向プライマー (配列番号33)：5' TGCCCTCCGGAAGCT。逆方向プライマー (配列番号34)：5' CGTTTCTGAAGGACATCTGATC。プローブ (配列番号35) (FAM - 5' - 3' - TAMRA)：TTGCAGCCAAAGTTAGGAGTGAAAGAGATGCA。

【0207】

GABA (B899P)：正方向プライマー (配列番号36)：5' AAGCCTCAGAGGTCCTTCCAGTATG。逆方向プライマー (配列番号37)：5' TTCAAATAATAAGTGAAAGAAAAAATTAGTAGATCA。プローブ (配列番号38) (FAM - 5' - 3' - TAMRA)：AATCCATTGTATCTTAGAA 50

C C G A G G G A T T T G T T T A G A。

【0208】

B 3 0 5 D (C 型) : 正方向プライマー (配列番号 3 9) : 5 ' A A A G C A G A T G G T G G T T G A G G G T T。逆方向プライマー (配列番号 4 0) : 5 ' C C T G A G A C C A A A T G G C T T C T T C。プローブ (配列番号 4 1) (F A M - 5 ' - 3 ' - T A M R A) A T T C C A T G C C G G C T G C T T C T T C T T C T G。

【0209】

B 7 2 6 P : 正方向プライマー (配列番号 4 2) : 5 ' T C T G G T T T T C T C A T T C T T T A T T C A T T T A T T。逆方向プライマー (配列番号 4 3) : 5 ' T G C C A A G G A G C G G A T T A T C T。プローブ (配列番号 4 4) (F A M - 5 ' - 3 ' - T A M R A) : C A A C C A C G T G A C A A A C A C T G G A A T T A C A G G。 10

【0210】

アクチン : 正方向プライマー (配列番号 4 5) : 5 ' A C T G G A A C G G T G A A G G T G A C A。逆方向プライマー (配列番号 4 6) : 5 ' C G G C C A C A T T G T G A A C T T T G。プローブ (配列番号 4 7) : (F A M - 5 ' - 3 ' - T A M R A) : C A G T C G G T T G G A G C G A G C A T C C C。

【0211】

アッセイ条件は以下のものである :

T a q m a n プロトコール (7 7 0 0 P e r k i n E l m e r) :
2 5 μ l 最終容量中 : 1 × 緩衝液 A、5 m M M g C l、0 . 2 m M d C T P、0 . 2 m M d A T P、0 . 4 m M d U T P、0 . 2 m M d G T P、0 . 0 1 U / μ l A m p E r a s e U N G、0 . 0 2 5 μ l T a q G o l d、8 % (v / v) グリセロール、0 . 0 5 % (v / v) ゼラチン、0 . 0 1 % (v / v) T w e e n 2 0、4 p m o l の各遺伝子特異的 T a q m a n プローブ (L i p o p h i l i n B + G a b a + B 3 0 5 D + B 7 2 6 P)、1 0 0 n M の B 7 2 6 P - F + B 7 2 6 P - R、3 0 0 n M の G a b a - R、および 5 0 n M の L i p o p h i l i n B - F + L o p o p h i l i n B - R + B 3 0 5 D - R + G a b a - R、鋳型 c D N A (0 . 0 2 μ g ポリ A + R N A に由来する)。 20

【0212】

L i p o p h i l i n B 発現を、2 7 個の胸部腫瘍サンプルのうち 1 4 サンプルにおいて 30 検出した。

【0213】

しかし、L i p o p h i l i n B、B 8 9 9 P、B 3 0 5 D - C および B 7 2 6 P についてのマルチプレックスアッセイは、2 7 腫瘍のうち 2 7 腫瘍において発現シグナルを検出し、検出レベルは、大多数のサンプルで 1 0 m R N A コピー / 1 0 0 0 p g アクチンを超え、そして試験された 2 7 サンプルのうち 5 サンプルにおいて 1 0 0 m R N A コピー / 1 0 0 0 p g アクチンを超えた (図 8)。

【0214】

(実施例 1 0)

(多重検出最適化)

上記の多重リアルタイム P C R アッセイを使用して、マンマグロビン (またはリポフィリン B)、G a b a (B 8 9 9 P)、B 3 0 5 D - C および B 7 2 6 P の発現を同時に検出した。この実施例に従って、アッセイ条件およびプライマー配列を、異なる長さの 4 つの P C R 産物の対応する増幅を得るように最適化した。このアッセイのポジティブサンプルを、ゲル電気泳動によってさらに特徴付け得、そして目的の発現遺伝子を、検出されたアンプリコンのサイズに従って決定し得る。 40

【0215】

マンマグロビン (またはリポフィリン B)、G a b a (B 8 9 9 P)、B 3 0 5 D および B 7 2 6 P の特異的プライマーおよび特異的 T a q m a n プローブを使用して、これらの発現を同時に検出する。この実施例において使用されるプライマーおよびプローブを以下 50

に示す。

【0216】

マンマグロピン：正方向プライマー（配列番号48）：5' T G C C A T A G A T G A A T T G A A G G A A T G。逆方向プライマー（配列番号49）：5' T G T C A T A T A T T A A T T G C A T A A A C A C C T C A。プローブ（配列番号50）（FAM - 5' - 3' - T A M R A）：T C T T A A C C A A A C G G A T G A A A C T C T G A G C A A T G。

【0217】

G A B A (B 8 9 9 P) : 正方向プライマー（配列番号36）：5' A A G C C T C A G A G T C C T T C C A G T A T G。逆方向プライマー（配列番号51）：5' A T C A T T G A A A A T T C A A A T A T A A G T G A A G。プローブ（配列番号38）（FAM - 5' - 3' - T A M R A）：A A T C C A T T G T A T C T T A G A A C C G A G G G A T T T G T T T A G A。

10

【0218】

B 3 0 5 D (C 型) : 正方向プライマー（配列番号39）：5' A A A G C A G A T G G T G G T T G A G G G T T。逆方向プライマー（配列番号40）：5' C C T G A G A C C A A A T G G C T T C T T C。プローブ（配列番号41）（FAM - 5' - 3' - T A M R A）A T T C C A T G C C G G C T G C T T C T T C T T C T G。

【0219】

B 7 2 6 P : 正方向プライマー（配列番号52）：5' G T A G T T G T G C A T T G A A A T A A T T A T C A T T A T。逆方向プライマー（配列番号43）：5' T G C C A A G G A G C G G A T T A T C T。プローブ（配列番号44）（FAM - 5' - 3' - T A M R A）：C A A C C A C G T G A C A A A C A C T G G A A T T A C A G G。

20

【0220】

プライマー位置およびアッセイ条件を、異なる長さの4つのPCR産物の対応する増幅を得るために最適化した。アッセイ条件は以下のようである：

Taqmanプロトコール（7700 Perkin Elmer）：

25 μ l 最終容量中：1 \times 緩衝液A、5 mM MgCl、0.2 mM dCTP、0.2 mM dATP、0.4 mM dUTP、0.2 mM dGTP、0.01 U / μ l AmpErase UNG、0.0375 μ l TaqGold、8% (v/v) グリセロール、0.05% (v/v) ゼラチン、0.01% (v/v) Tween20、4 pmol の各遺伝子特異的Taqmanプローブ（Mammaglobin + Gaba + B305D + B726P）、300 nMのGaba - R + Gaba - F、100 nMのMammaglobin - F + R；B726P - F + R、および50 nMのB305D - F + R 鋳型cDNA（0.02 μ g ポリA + RNAに由来する）。

30

【0221】

（PCRプロトコール）

50 2分間を1回、95 10分間を1回、および95 15秒間 / 60 ° 1分間 / 68 1分間を50回。

【0222】

マルチプレックスアッセイにおいて設定された各プライマーは、独特の長さのバンドを生じるので、目的の4つの遺伝子の発現シグナルは、アガロースゲル分析によって個々に測定し得るか（図9を参照のこと）、または4つ全ての遺伝子の組み合わせ発現シグナルを、ABI 7700 Prism配列検出システム（PE Biosystems, Foster City, CA）でリアルタイムで測定し得る。リポフィリンBの発現もまた、マンマグロピンの代わりに検出し得る。特定のプライマーが本明細書中に記載されているが、異なるプライマー配列、異なるプライマーまたはプローブの標識、および異なる検出系を使用して、このマルチプレックスアッセイを行ない得る。例えば、第2の蛍光発生的レポーター色素を、リアルタイムPCRによって参照遺伝子の対応する検出のために組み込み得る。または、例えば、Taqmanプローブアプローチの代わりに、SYBR

40

50

Green 検出系を使用し得る。

【0223】

(実施例11)

(乳癌マルチプレックスアッセイのためのゲノムDNAを排除しイントロン-エキソン境界に広がるプライマー対の設計および使用)

本明細書中に記載される多重リアルタイムPCRアッセイは、マンマグロビン、Gaba (B899P)、B305D-CおよびB726Pの発現を同時に検出し得る。これらの遺伝子の組み合わせられた発現レベルを、ABI 7700 Prism 配列検出システム (PE Biosystems, Foster City, CA) でリアルタイムで測定する。個別に発現された遺伝子をまた、ゲル電気泳動を介して異なるアンプリコンサイズに起因して、同定し得る。DNase 処理していないRNAに由来するサンプル (例えば、リンパ節cDNA) を用いてこのアッセイを使用するするため、そして低分子のRNAサンプル (例えば、血液標本、腫瘍およびリンパ節吸引物由来) のためのDNase 処理を避けるために、イントロンに広がるプライマー対を設計して、ゲノムDNAの増幅を排除し、従って非特異的シグナルおよび偽陽性シグナルを排除する。偽陽性シグナルは、cDNA 標本におけるゲノムDNAの混入によって引き起こされる。本明細書中に記載される最適マルチプレックスアッセイは、ゲノムDNAの増幅を排除し、そしてRNAサンプルのDNase 前処理の必要性を伴わずに標的遺伝子発現の特異的検出を可能にする。さらに、ゲノムの適合およびイントロン-エキソン境界の位置は、これらのプライマー組で検証され得る。

【0224】

マンマグロビン特異的プライマー、Gaba (B899P) 特異的プライマー、B305D 特異的プライマーおよびB726P 特異的プライマーおよび特異的Taqman プローブを使用して、これらの発現を同時に検出した (表7)。(イントロン-エキソン境界に広がる) プライマー位置を最適化して、cDNA のみを検出し、そしてゲノムDNAを増幅から排除する。発現された遺伝子の身元を、ゲル電気泳動によって決定した。

【0225】

【表7】

表 7
乳癌マルチプレックスアッセイに関するイントロン-エキソン境界のプライマーおよびプロブ配列

遺伝子	3'側方向プライマー	逆方向プライマー	Taqman 700-F" (FAM-5'-3'TAMRA)
マンマグロビン	tgccatagatgaatgaagga alg (配列番号 48)	tgtcatatattaatgcataaacct ca (配列番号 49)	tcitaccaaaccggatgaactcigagca atg (配列番号 50)
B899P	aagccctcagagtcctccagta tg (配列番号 36)	ttcaaatataagtgagaaaaaatta gtatgcaa (配列番号 37)	aatccattgtatcttagaaccgaggttt gttt (配列番号 62)
B305D	aaagcagatgggtgaggt t (配列番号 39)	cctgagaccatggctcttc (配列番号 40)	attccatgccgctgcttcttg (配列 番号 41)
B726P	tctggcttctcattcttattcatt tatt (配列番号 42)	tgccaaggagcggattatct (配列番号 43)	caaccacgtgacaacactggaattaca gg (配列番号 44)
Actin	actggaacgggaaggtgac a (配列番号 45)	cggccacattgtgaactttg (配列番号 46)	cagtcgggtggagcgagcatccc (配列番号 47)
B899P-INT	caatttgggtggagaacccg (配列番号 53)	gctgtcggaggtataggtg (配列番号 54)	cattcagagagtaacaaggactacaca (配列番号 55)
B305D-INT	tctgataaaggccgtacaatg (配列番号 56)	tcacgactgtgtgtttgtctc (配列番号 57)	atcaaaaaaacagcatggcctcacacca ct (配列番号 58)
B726P-INT	gcaagtgcacatgatcagagg (配列番号 59)	atatagactcaggatcacact (配列番号 60)	tcacatcagatccaacaagaggaaga ig (配列番号 61)

プライマーの位置およびアッセイ条件を、4つのPCR産物の対応する増幅を達成するために最適化した。このアッセイ条件は、以下の通りであった：

(Taqman プロトコル (7700 Perkin Elmer))

25 μ l の最終容積中：1 \times 緩衝液 A、5 mM MgCl₂ 0.2 mM dCTP、0.2 mM dATP、0.4 mM dUTP、0.2 mM dGTP、0.01 U / Amp

E r a s e U N G、8 % (v / v) グリセロール、0 . 0 5 % (v / v) ゼラチン、0 . 0 1 % (v / v) T w e e n 2 0、4 p m o l のそれぞれの遺伝子特異的 T a q m a n プローブ (マンマグロビン + B 8 9 9 P - I N T + B 3 0 5 D - I N T + B 7 2 6 P - I N T)、3 0 0 n M の B 3 0 5 D - I N T - F ; B 8 9 9 P - I N T - F , 1 0 0 n M のマンマグロビン - F + R ; B 7 2 6 P - I N T - F + R、5 0 n M の B 8 9 9 P - I N T - R ; B 3 0 5 - I N T - R、テンプレート c D N A (0 . 0 2 μ g ポリ A + R N A を起源とする)。

【 0 2 2 6 】

(P C R サイクリング条件)

5 0 、2 分間を 1 サイクル、9 5 、1 0 分間を 1 サイクル、9 5 1 分間および 6 8 10
、1 分間を 5 0 サイクル。

【 0 2 2 7 】

図 1 0 は、リンパ節組織のパネルにおける乳癌細胞を検出するために、イントロン - エキソン境界スパニングプライマーを使用するマルチプレックスアッセイ (下のパネル) および非最適化プライマーを使用するマルチプレックスアッセイ (上のパネル) の比較を示す。この実験は、サンプル中のゲノム D N A 夾雑物から生じるバックグラウンドの減少を、本発明のイントロン - エキソンスパニングプライマーを使用して達成することを示す。

【 0 2 2 8 】

(実施例 1 2)

(歩哨リンパ節生検サンプルにおける転位した乳癌細胞の多重検出)

20

リンパ節病期分類は、適切なアジュバントホルモンおよび化学療法を決定するために重要である。従来の腋窩の解剖と対照的に、最小の後遺疾患の病期分類に関するより侵襲性でないアプローチは、歩哨リンパ節生検である。歩哨リンパ節生検 (S L N B) は、転位の検出を改善し、そして予後値を提供する可能性があり、完全なリンパ節解剖に関連する最小の罹患率を有する治療を導く。S L N B は、最初に腫瘍を排出する 1 つまたは 2 つのリンパ節のマッピングを実行し、従って、たいてい、転位性疾患 (歩哨節) を有するようである。リンパ節の日常的な病理学的な分析は、高い偽陰性比 (f a l s e - n e g a t i v e r a t e) を生じる：病理学的陰性リンパ節を有する 3 分の 1 の女性が、再発性疾患に発達する [B l a n d : T h e B r e a s t : S a u n d e r s 1 9 9 1]。腫瘍細胞に関するより感受性の検出技術は、P T - P C R であるが、この適用は、単一の特定 30
のマーカの欠失によって制限される。上記のマルチマーカアッセイは、正常のリンパ節から生じる偽陰性を生じることなく、集団を介して癌検出の可能性を増大する。

【 0 2 2 9 】

上記のように、原発性腫瘍由来のリンパ性求心性神経を、領域性のリンパ性骨盤へのドレナージが起きる前に、単一の節 (歩哨リンパ節) に排出する。歩哨リンパ節を、一次病変部位に注入された色素および / または放射標識コロイドを用いて位置付け、そして歩哨リンパ節生検は、微小転位性沈着物に関する病理学的試験を可能し、従って、腋窩の病期分類は、不必要な腋窩の解剖を避け得る。結節性の微小転位を、サイトケラチンタンパク質に関して、染色 (ヘマトキシリンまたはエオシン) または免疫組織化学的分析を使用して位置付け得る。免疫組織化学的染色技術は、節の不適切な切片化に起因して、小転位性病 40
巣を見落とすことによって頻繁に偽陰性結果を生じ得る。免疫組織化学は、サイトケラチン (細網細胞) の非摘出の発現に起因して、偽陰性結果を生じ得るか、または対応する抗原が、全ての腫瘍細胞に対して発現しない抗体 B e r - E p 4 を使用する場合、偽陰性結果を生じ得る。

【 0 2 3 0 】

本明細書中に記載されるマルチプレックスアッセイは、陽性歩哨リンパ節に関するより感受性の検出を提供し得た。さらに、骨髓サンプル、末梢血および小ニードル吸引サンプルにおける乳癌細胞の検出は、乳癌患者における診断および予後に所望される。

【 0 2 3 1 】

2 2 個の転位性リンパ節サンプル (加えて、1 5 個のサンプルをまた、以前に分析し、そ 50

して図3Aに示す)を、本明細書中に記載されるイントロン-エキソン境界スパニング多重PCRアッセイを使用して分析した。この分析からの結果を、表8に要約する。27個の原発性腫瘍をまた、分析し、そして結果を表9に示す。このアッセイを使用して試験した20個の正常なリンパ節サンプルは、すべて陰性であった。

【0232】

【表8】

表8
571個の転位性リンパ節の多重PCR分析

転位性リンパ節 サンプル	マクロヒン	B305D	B899P	B726P	多重化
B.Met 317A	++	+		+	+++
B.Met 318A			++		+++
B.Met 595A	+			+	+++
B.Met 611A	+	+	+++		++
B.Met 612A	++	++		+	++
B.Met 614A		++		++	+++
B.Met 616A			+		++
B.Met 618A	+++	+			+++
B.Met 620A	++	++		++	+++
B.Met 621A	+	+++		+	+++
B.Met 624A			++		+++
B.Met 625A		++		++	+
B.Met 627A		+		+	+
B.Met 629A	++				+++
B.Met 631A	+		++		+
1255	+++	++		++	++
1257	+++	+	+	+	++
769	+++			+	++
1258	++	+	+		+
1259		++	++		+++
1250	+++	+		+	+++
1726	+++	+		+	+++
786	+++	+	+		+++
281-L1-r	+++				+++
289-L2	++	+			++
366-S	+				+
374-S+	+++	++			+++
376-S	++			+	++
381-S	+	+			+
383-Sx	+++	++			+++
496-M	+++	++			+++
591-SI-A	+	+			+
652-I		+	++		+++
772	+				+
777	+	+		++	++
778	+++				+++
779	+		++		++

【0233】

【表9】

10

20

30

40

表 9
271個の原発性乳癌のリアルタイムPCR分析

原発性乳癌形態	マンマグロビン	B305D	B899P	B726P	増悪化
T443	+	++		+++	+++
T457		+	+		++
T395			++		++
T10A	+	+++		+++	+++
T446		+		++	++
T11C	+		+++		+++
T23B	+	++			+++
T207A		++			+
T437	+	+		++	+++
T391	+	++		+++	+++
T392	+	+			++
TS76	+	++			+++
T483	++	+			+++
T81G	+	+	++	++	+++
T430	+		++		++
T465	+	+		+	++
TS80			+		+
T469	+			+	+++
T467	+			++	+++
T439		+			+
T387	++		+	+	++
T318			+		++
T154A				+	+
T387A	+++		+	+	+++
T155A	+		++	+	+
T209A		++			++
T208A		+		+	++

10

20

前述の結果から、本発明の特定の実施形態が、例示の目的のために本明細書中に記載されるが、種々の改変が、本発明の精神および範囲を逸脱することなく行なわ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以外によって限定されない。

【図面の簡単な説明】

【図 1】

図 1 は、定量 PCR (TaqmanTM) を用いて決定される、B 3 1 1 D、B 5 3 3 S、および B 7 2 6 P についての mRNA 発現プロフィールを示す。略語：B . T . : 胸部腫瘍；B . M . : 骨髄；B . R . : 乳房の整復。

30

【図 2】

図 2 は、腫瘍の病理的段階に対する B 5 3 3 S 発現の関係を示す。正常な乳房 (8) 由来の組織、良性の乳房障害 (3)、および乳房の腫瘍の第 I 段階 (5)、第 II 段階 (6)、第 III 段階 (7)、第 IV 段階 (3)、ならびに転移 (1 つのリンパ節および 3 つの胸水) を、リアルタイム PCR において試験した。このデータを、試験された各グループについて平均コピー / ng - アクチンとして表わし、線は、算出された傾向線である。

【図 3】

図 3 A および 3 B は、それぞれ、B 3 0 5 D C 形態、B 7 2 6 P、G A B A、およびの転移におけるマンマグロビン (mammaglobin) ならびに原発性腫瘍の遺伝子の補完性を示す。それぞれの遺伝子についてのカットオフは、陰性の正常組織の平均および 3 つの標準的偏差に基づく 6 . 5 7、1 . 6 5、4 . 5 8、および 3 . 5 6 コロニー / ng アクチンであった。

40

【図 4】

図 4 は、マンマグロビンについての全長 cDNA 配列を示す。

【図 5】

図 5 は、E . c o l i i において発現されるマンマグロビン組換えポリペプチドをコードするオープンリーディングフレームの決定された cDNA 配列を示す。

50

【 図 6 】

図 6 は、G A B A についての全長 c D N A 配列を示す。

【 図 7 】

図 7 は、リアルタイム P C R および S Y B R 検出系を用いて決定された、胸部腫瘍および正常サンプル中のマンマグロビン、G A B A 、B 3 0 5 D (C 形態)、および B 7 2 6 P についての m R N A 発現レベルを示す。略称：B T：胸部腫瘍；B R：乳房の整復；A . P B M C：活性化された末梢血単核細胞；R . P B M C：静止状態の P B M C；T . G l a n d：甲状腺；S . C o r d：脊髄；A . G l a n d：副腎；B . M a r r o w：骨髓；S . M u s c l e：骨格筋。

【 図 8 】

図 8 は、胸部腫瘍サンプルのパネル上で試験されたりポフィリン B 単独とリポフィリン B - B 8 9 9 P - B 3 0 5 D - C - B 7 2 6 のマルチプレックスアッセイとの間の比較を示す棒グラフである。略称：B T：胸部腫瘍；B R：乳房の整復；S C I D：重症複合免疫不全。

【 圖 9 】

図 9 は、マルチプレックスリアルタイム PCR アッセイにて試験された目的の腫瘍遺伝子の 4 つの増幅産物（マンマグロビン、B305D、B899P、B726P）の独特のバンドの長さを示すゲルである。

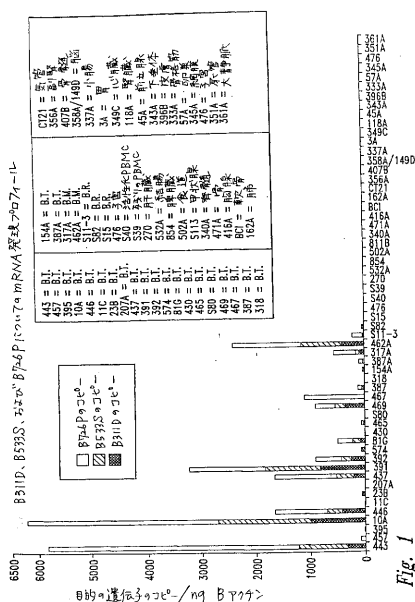
【 図 1 0 】

図 10 は、リンパ節組織のパネルにおける乳癌細胞を検出するための、イントロン - エキソンに広範にわたるプライマー（パネル下）を用いたマルチプレックスアッセイと非最適化プライマー（パネル上）を用いたマルチプレックスアッセイとの比較を示す。

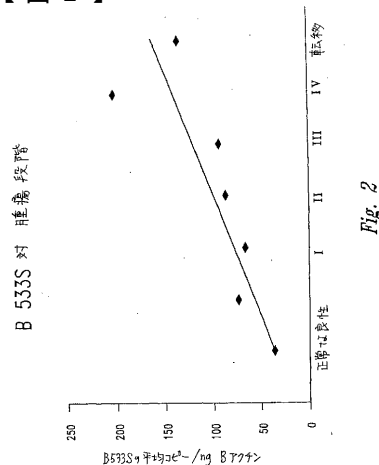
10

20

【 圖 1 】



【 図 2 】



【 図 3 】

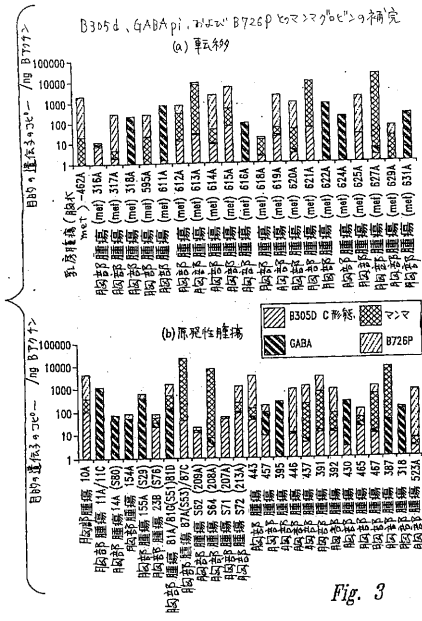


Fig. 3

【 図 7 】

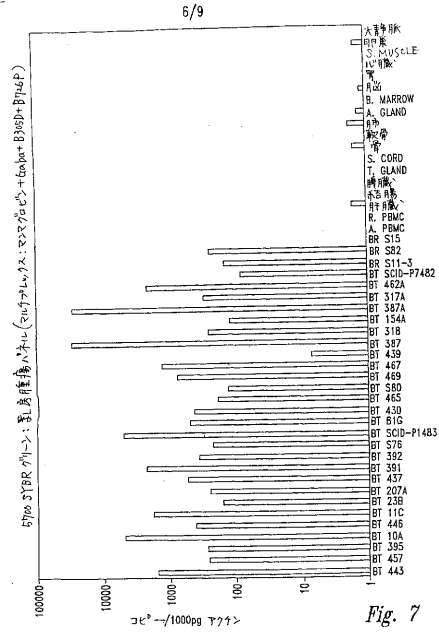


Fig. 7

【 図 8 】

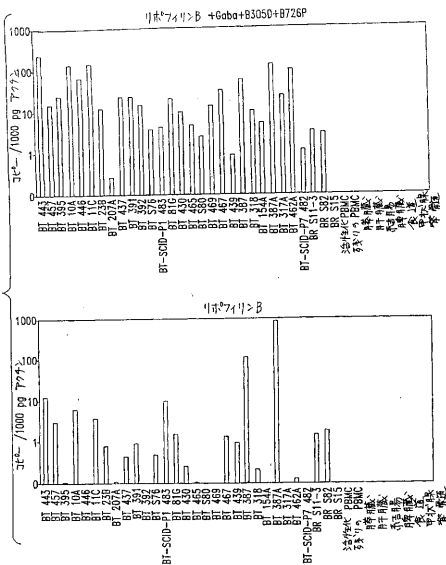


Fig. 8

【 図 9 】

マルチプレックスPCRで:
アンプリコンの大きさによる遺伝子決定

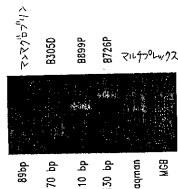


Fig. 9

【図 10】

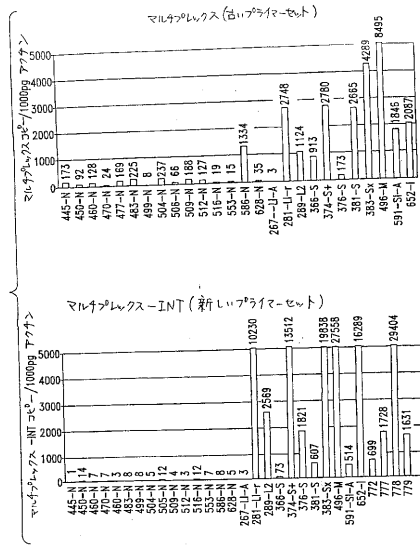


Fig. 10

【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
11 October 2001 (11.10.2001)

PCT

(10) International Publication Number
WO 01/75171 A2(51) International Patent Classification: **C12Q 1/68**,
G01N 33/543, 33/574 // C12N 5/08

(21) International Application Number: PCT/US01/10631

(22) International Filing Date: 2 April 2001 (02.04.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/194,241 3 April 2000 (03.04.2000) US
60/219,862 20 July 2000 (20.07.2000) US
60/221,300 27 July 2000 (27.07.2000) US
60/256,592 18 December 2000 (18.12.2000) 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

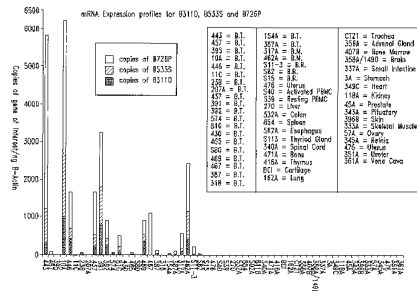
(75) Inventors/Applicants (for US only): **HOUGHTON, Raymond, L.** [US/US]; 2636 242nd Place S.E., Bothell, WA 98021 (US). **DILLON, Davin, C.** [US/US]; 18112 N.W. Montreux Drive, Issaquah, WA 98027 (US). **MOLESH,****David, Alan** [US/US]; 12385 N.E. Klabo Road, Kingston, WA 98346 (US). **XU, Jiangchun** [US/US]; 15805 S.E. 43rd Place, Bellevue, WA 98006 (US). **ZEHEHNTNER, Barbara** [DE/US]; 4420 Bluff Lane, Bainbridge Island, WA 98110 (US). **PERSING, David, H.** [US/US]; 22401 N.E. 25th Way, Redmond, WA 98053 (US).(74) Agents: **POTTER, Jane, E., R.** et al.; Seed Intellectual Property Law Group PLLC, Suite 6300, 701 Fifth Avenue, Seattle, WA 98104-7092 (US).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

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

[Continued on next page]

(54) Title: METHODS, COMPOSITIONS AND KITS FOR THE DETECTION AND MONITORING OF BREAST CANCER



WO 01/75171 A2**Published:**

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

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

WO 01/75171

PCT/US01/10631

METHODS, COMPOSITIONS AND KITS
FOR THE DETECTION AND MONITORING OF BREAST CANCER

TECHNICAL FIELD OF THE INVENTION

The present invention relates generally to the field of cancer diagnostics.

5 More specifically, the present invention relates to methods, compositions and kits for the detection of cancer that employ oligonucleotide hybridization and/or amplification to simultaneously detect two or more tissue-specific polynucleotides in a biological sample suspected of containing cancer cells.

BACKGROUND OF THE INVENTION

10 Cancer remains a significant health problem throughout the world. The failure of conventional cancer treatment regimens can commonly be attributed, in part, to delayed disease diagnosis. Although significant advances have been made in the area of cancer diagnosis, there still remains a need for improved detection methodologies that permit early, reliable and sensitive determination of the presence of cancer cells.

15 Breast cancer is second only to lung cancer in mortality among women in the U.S., affecting more than 180,000 women each year and resulting in approximately 40,000-50,000 deaths annually. For women in North America, the life-time odds of getting breast cancer are one in eight.

Management of the disease currently relies on a combination of early
20 diagnosis (through routine breast screening procedures) 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 analysis of specific tumor markers. See, e.g., Porter-Jordan et al., *Breast Cancer* 8:73-100
25 (1994). The use of established markers often leads, however, to a result that is difficult to interpret, and the high mortality observed in breast cancer patients indicates that improvements are needed in the diagnosis of the disease.

The recent introduction of immunotherapeutic approaches to breast cancer treatment which are targeted to Her2/neu have provided significant motivation to

WO 01/75171

PCT/US01/10631

2

identify additional breast cancer specific genes as targets for therapeutic antibodies and T-cell vaccines as well as for diagnosis of the disease. To this end, mammaglobin, has been identified as one of the most breast-specific genes discovered to date, being expressed in approximately 70-80% of breast cancers. Because of its highly tissue-specific distribution, detection of mammaglobin gene expression has been used to identify micrometastatic lesions in lymph node tissues and, more recently, to detect circulating breast cancer cells in peripheral blood of breast cancer patients with known primary and metastatic lesions.

Mammaglobin is a homologue of a rabbit uteroglobin and the rat steroid binding protein subunit C3 and is a low molecular weight protein that is highly glycosylated. Watson et al., *Cancer Res.* 56:860-5 (1996); Watson et al., *Cancer Res.* 59:3028-3031 (1999); Watson et al., *Oncogene* 16:817-24 (1998). In contrast to its homologs, mammaglobin has been reported to be breast specific and overexpression has been described in breast tumor biopsies (23%), primary and metastatic breast tumors (~75%) with reports of the detection of mammaglobin mRNA expression in 91% of lymph nodes from metastatic breast cancer patients. Leygue et al., *J. Pathol.* 189:28-33 (1999) and Min et al., *Cancer Res.* 58:4581-4584 (1998).

Since mammaglobin gene expression is not a universal feature of breast cancer, the detection of this gene alone may be insufficient to permit the reliable detection of all breast cancers. Accordingly, what is needed in the art is a methodology that employs the detection of two or more breast cancer specific genes in order to improve the sensitivity and reliability of detection of micrometastases, for example, in lymph nodes and bone marrow and/or for recognition of anchorage-independent cells in the peripheral circulation.

The present invention achieves these and other related objectives by providing methods that are useful for the identification of tissue-specific polynucleotides, in particular tumor-specific polynucleotides, as well as methods, compositions and kits for the detection and monitoring of cancer cells in a patient afflicted with the disease.

WO 01/75171

PCT/US01/10631

3

SUMMARY OF THE INVENTION

By certain embodiments, the present invention provides methods for identifying one or more tissue-specific polynucleotides which methods comprise the steps of: (a) performing a genetic subtraction to identify a pool of polynucleotides from a tissue of interest; (b) performing a DNA microarray analysis to identify a first subset of said pool of polynucleotides of interest wherein each member polynucleotide of said first subset is at least two-fold over-expressed in said tissue of interest as compared to a control tissue; and (c) performing a quantitative polymerase chain reaction analysis on polynucleotides within said first subset to identify a second subset of polynucleotides that are at least two-fold over-expressed as compared to the control tissue. Preferred genetic subtractions are selected from the group consisting of differential display and cDNA subtraction and are described in further detail herein below.

Alternate embodiments of the present invention provide methods of identifying a subset of polynucleotides showing concordant and/or complementary tissue-specific expression profiles in a tissue of interest. Such methods comprise the steps of, (a) performing an expression analysis selected from the group consisting of DNA microarray and quantitative PCR to identify a first polynucleotides that is at least two-fold over-expressed in a tissue of interest as compared to a control tissue; and (b) performing an expression analysis selected from the group consisting of DNA microarray and quantitative PCR to identify a first polynucleotides that is at least two-fold over-expressed in a tissue of interest as compared to a control tissue.

Further embodiments of the present invention provide methods for detecting the presence of a cancer cell in a patient. Such methods comprise the steps of: (a) obtaining a biological sample from the patient; (b) contacting the biological sample with a first oligonucleotide pair wherein the members of the first oligonucleotide pair hybridize, under moderately stringent conditions, to a first polynucleotide and the complement thereof, respectively; (c) contacting the biological sample with a second oligonucleotide pair wherein the members of the second oligonucleotide pair hybridize, under moderately stringent conditions, to a second polynucleotide and the complement thereof, respectively and wherein the first polynucleotide is unrelated in nucleotide

WO 01/75171

PCT/US01/10631

4

sequence to the second polynucleotide; (d) amplifying the first polynucleotide and the second polynucleotide; and (e) detecting the amplified first polynucleotide and the amplified second polynucleotide; wherein the presence of the amplified first polynucleotide or amplified second polynucleotide indicates the presence of a cancer cell in the patient.

By some embodiments, detection of the amplified first and/or second polynucleotides may be preceded by a fractionation step such as, for example, gel electrophoresis. Alternatively or additionally, detection of the amplified first and/or second polynucleotides may be achieved by hybridization of a labeled oligonucleotide probe that hybridizes specifically, under moderately stringent conditions, to the first or second polynucleotide. Oligonucleotide labeling may be achieved by incorporating a radiolabeled nucleotide or by incorporating a fluorescent label.

In certain preferred embodiments, cells of a specific tissue type may be enriched from the biological sample prior to the steps of detection. Enrichment may be achieved by a methodology selected from the group consisting of cell capture and cell depletion. Exemplary cell capture methods include immunocapture and comprise the steps of: (a) adsorbing an antibody to a tissue-specific cell surface to cells said biological sample; (b) separating the antibody adsorbed tissue-specific cells from the remainder of the biological sample. Exemplary cell depletion may be achieved by cross-linking red cells and white cells followed by a subsequent fractionation step to remove the cross-linked cells.

Alternative embodiments of the present invention provide methods for determining the presence or absence of a cancer in a patient, comprising the steps of: (a) contacting a biological sample obtained from the patient with an oligonucleotide that hybridizes to a polynucleotide that encodes a breast tumor protein; (b) detecting in the sample a level of a polynucleotide (such as, for example, 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

WO 01/75171

PCT/US01/10631

5

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.

5 In related aspects, methods are provided for monitoring the progression 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 breast tumor protein; (b) detecting in the sample an amount of a polynucleotide that hybridizes to the oligonucleotide; (c) repeating steps (a) and (b)
10 using a biological sample obtained from the patient at a subsequent point in time; and (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.

Certain embodiments of the present invention provide that the step of
15 amplifying said first polynucleotide and said second polynucleotide is achieved by the polymerase chain reaction (PCR).

Within certain embodiments, the cancer cell to be detected may be selected from the group consisting of prostate cancer, breast cancer, colon cancer, ovarian cancer, lung cancer head & neck cancer, lymphoma, leukemia, melanoma, liver
20 cancer, gastric cancer, kidney cancer, bladder cancer, pancreatic cancer and endometrial cancer. Still further embodiments of the present invention provide that the biological sample is selected from the group consisting of blood, a lymph node and bone marrow. The lymph node may be a sentinel lymph node.

Within specific embodiments of present invention it is provided that the
25 first polynucleotide is selected from the group consisting of mammaglobin, lipophilin B, GABA π (B899P), B726P, B511S, B533S, B305D and B311D. Other embodiments provide that the second polynucleotide is selected from the group consisting of mammaglobin, lipophilin B, GABA π (B899P), B726P, B511S, B533S, B305D and B311D.

30 Alternate embodiments of the present invention provide methods for detecting the presence or absence of a cancer in a patient, comprising the steps of: (a)

WO 01/75171

PCT/US01/10631

6

contacting a biological sample obtained from a patient with a first oligonucleotide that hybridizes to a polynucleotide selected from the group consisting of mammaglobin and lipophilin B; (b) contacting the biological sample with a second oligonucleotide that hybridizes to a polynucleotide sequence selected from the group consisting of GABA π (B899P), B726P, B511S, B533S, B305D and B311D; (c) detecting in the sample an amount of a polynucleotide that hybridizes to at least one of the oligonucleotides; and (d) comparing the amount of polynucleotide that hybridizes to the oligonucleotide to a predetermined cut-off value, and therefrom determining the presence or absence of a cancer in the patient.

10 According to certain embodiments, oligonucleotides may be selected from those disclosed herein such as those presented in SEQ ID Nos:33-72. By other embodiments, the amount of polynucleotide that hybridizes to the oligonucleotide is determined using a polymerase chain reaction. Alternatively, the amount of polynucleotide that hybridizes to the oligonucleotide may be determined using a
15 hybridization assay.

Still other embodiments of the present invention provide methods for determining the presence or absence of a cancer cell in a patient, comprising the steps of: (a) contacting a biological sample obtained from a patient with a first oligonucleotide that hybridizes to a polynucleotide selected from the group consisting
20 of a polynucleotide depicted in SEQ ID NO:73 and SEQ ID NO:74 or complement thereof; (b) contacting the biological sample with a second oligonucleotide that hybridizes to a polynucleotide depicted in SEQ ID NO:75 or complement thereof; (c) contacting the biological sample with a third oligonucleotide that hybridizes to a polynucleotide selected from the group consisting of a polynucleotide depicted in SEQ
25 ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:6 and SEQ ID NO:7 or complement thereof; (d) contacting the biological sample with a fourth oligonucleotide that hybridizes to a polynucleotide selected from the group consisting of a polynucleotide depicted in SEQ ID NO:11 or complement thereof; (e) contacting the biological sample with a fifth oligonucleotide that hybridizes to a polynucleotide
30 selected from the group consisting of a polynucleotide depicted in SEQ ID NO:13, 15 and 17 or complement thereof; (f) contacting the biological sample with a sixth

WO 01/75171

PCT/US01/10631

7

oligonucleotide that hybridizes to a polynucleotide selected from the group consisting of a polynucleotide depicted in SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23 and SEQ ID NO:24 or complement thereof; (g) contacting the biological sample with a seventh oligonucleotide that hybridizes to a polynucleotide depicted in SEQ ID NO:30 or complement thereof; (h) contacting the biological sample with an eighth oligonucleotide that hybridizes to a polynucleotide depicted in SEQ ID NO:32 or complement thereof; (i) contacting the biological sample with a ninth oligonucleotide that hybridizes to a polynucleotide depicted in SEQ ID NO:76 or complement thereof; (j) detecting in the sample a hybridized oligonucleotide of any one of steps (a) through (i); and (j) comparing the amount of polynucleotide that hybridizes to the oligonucleotide to a predetermined cut-off value, wherein the presence of a hybridized oligonucleotide in any one of steps (a) through (i) in excess of the predetermined cut-off value indicates the presence of a cancer cell in the biological sample of said patient.

Other related embodiments of the present invention provide methods for determining the presence or absence of a cancer cell in a patient, comprising the steps of: (a) contacting a biological sample obtained from a patient with a first oligonucleotide and a second oligonucleotide wherein said first and second oligonucleotides hybridize under moderately stringent conditions to a first and a second polynucleotide selected from the group consisting of SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:30, SEQ ID NO:32, and SEQ ID NO:76 and wherein said first polynucleotide is unrelated structurally to said second polynucleotide; (b) detecting in the sample said first and said second hybridized oligonucleotides; and (c) comparing the amount of polynucleotide that hybridizes to the oligonucleotide to a predetermined cut-off value, wherein the presence of a hybridized first oligonucleotide or a hybridized second oligonucleotide in excess of the pre-determined cut-off value indicates the presence of a cancer cell in the biological sample of said patient.

WO 01/75171

PCT/US01/10631

8

Other related embodiments of the present invention provide methods for determining the presence or absence of a cancer cell in a patient, comprising the steps of: (a) contacting a biological sample obtained from a patient with a first oligonucleotide and a second oligonucleotide wherein said first and second
5 oligonucleotides hybridize under moderately stringent conditions to a first and a second polynucleotide are both tissue-specific polynucleotides of the cancer to be detected and wherein said first polynucleotide is unrelated structurally to said second polynucleotide; (b) detecting in the sample said first and said second hybridized oligonucleotides; and (c) comparing the amount of polynucleotide that hybridizes to the oligonucleotide to a
10 predetermined cut-off value, wherein the presence of a hybridized first oligonucleotide or a hybridized second oligonucleotide in excess of the pre-determined cut-off value indicates the presence of a cancer cell in the biological sample of said patient.

In other related aspects, the present invention further provides compositions useful in the methods disclosed herein. Exemplary compositions
15 comprise two or more oligonucleotide primer pairs each one of which specifically hybridizes to a distinct polynucleotide. Exemplary oligonucleotide primers suitable for compositions of the present invention are disclosed herein by SEQ ID NOs: 33-71. Exemplary polynucleotides suitable for compositions of the present invention are disclosed in SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, SEQ ID NO:1, SEQ ID
20 NO:3, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:30, SEQ ID NO:32, and SEQ ID NO:76.

The present invention also provides kits that are suitable for performing
25 the detection methods of the present invention. Exemplary kits comprise oligonucleotide primer pairs each one of which specifically hybridizes to a distinct polynucleotide. Within certain embodiments, kits according to the present invention may also comprise a nucleic acid polymerase and suitable buffer. Exemplary oligonucleotide primers suitable for kits of the present invention are disclosed herein by
30 SEQ ID NOs: 33-71. Exemplary polynucleotides suitable for kits of the present invention are disclosed in SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, SEQ ID

WO 01/75171

PCT/US01/10631

9

NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:30, SEQ ID NO:32, and SEQ ID NO:76.

5 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 each was incorporated individually.

BRIEF DESCRIPTION OF THE DRAWINGS AND SEQUENCE IDENTIFIERS

10 Figure 1 shows the mRNA expression profiles for B311D, B533S and B726P as determined using quantitative PCR (Taqman™). Abbreviations: B.T.: Breast tumor; B.M.: Bone marrow; B.R.: Breast reduction.

 Figure 2 shows the relationship of B533S expression to pathological stage of tumor. Tissues from normal breast (8), benign breast disorders (3), and breast tumors stage I (5), stage II (6), stage III (7), stage IV (3) and metastases (1 lymph node and 3 pleural effusions) were tested in real-time PCR. The data is expressed as the mean copies/ng β -actin for each group tested and the line is the calculated trend line.

 Figures 3A and 3B show the gene complementation of B305D C-form, B726P, GABA π and mammaglobin in metastases and primary tumors, respectively.

20 The cut-off for each of the genes was 6.57, 1.65, 4.58 and 3.56 copies/ng β -Actin based on the mean of the negative normal tissues plus 3 standard deviations.

 Figure 4 shows the full-length cDNA sequence for mammaglobin.

 Figure 5 shows the determined cDNA sequence of the open reading frame encoding a mammaglobin recombinant polypeptide expressed in *E. coli*.

25 Figure 6 shows the full-length cDNA sequence for GABA π .

 Figure 7 shows the mRNA expression levels for mammaglobin, GABA π , B305D (C form) and B726P in breast tumor and normal samples determined using real-time PCR and the SYBR detection system. Abbreviations: BT: Breast tumor; BR: Breast reduction; A. PBMC: Activated peripheral blood mononuclear cells;

30 R. PBMC: resting PBMC; T. Gland: Thyroid gland; S. Cord: Spinal Cord; A. Gland: Adrenal gland; B. Marrow: Bone marrow; S. Muscle: Skeletal muscle.

WO 01/75171

PCT/US01/10631

10

Figure 8 is a bar graph showing a comparison between the LipophilinB alone and the LipophilinB-B899P-B305D-C-B726 multiplex assays tested on a panel of breast tumor samples. Abbreviations: BT: Breast tumor; BR: Breast reduction; SCID: severe combined immunodeficiency.

5 Figure 9 is a gel showing the unique band length of four amplification products of tumor genes of interest (mammaglobin, B305D, B899P, B726P) tested in a multiplex Real-time PCR assay.

Figure 10 shows a comparison of a multiplex assay using intron-exon border spanning primers (bottom panel) and those using non-optimized primers (top panel), to detect breast cancer cells in a panel of lymph node tissues.

SEQ ID NO: 1 is the determined cDNA sequence for a first splice variant of B305D isoform A.

SEQ ID NO: 2 is the amino acid sequence encoded by the sequence of SEQ ID NO: 1.

15 SEQ ID NO: 3 is the determined cDNA sequence for a second splice variant of B305D isoform A.

SEQ ID NO: 4 is the amino acid sequence encoded by the sequence of SEQ ID NO: 3.

SEQ ID NO: 5-7 are the determined cDNA sequences for three splice variants of B305D isoform C.

SEQ ID NO: 8-10 are the amino acid sequences encoded by the sequence of SEQ ID NO: 5-7, respectively.

SEQ ID NO: 11 is the determined cDNA sequence for B311D.

25 SEQ ID NO: 12 is the amino acid sequence encoded by the sequence of SEQ ID NO: 11.

SEQ ID NO: 13 is the determined cDNA sequence of a first splice variant of B726P.

SEQ ID NO: 14 is the amino acid sequence encoded by the sequence of SEQ ID NO: 13.

30 SEQ ID NO: 15 is the determined cDNA sequence of a second splice variant of B726P.

SEQ ID NO: 16 is the amino acid sequence encoded by the sequence of SEQ ID NO: 15.

WO 01/75171

PCT/US01/10631

11

SEQ ID NO: 17 is the determined cDNA sequence of a third splice variant of B726P.

SEQ ID NO: 18 is the amino acid sequence encoded by the sequence of SEQ ID NO: 17.

5 SEQ ID NO: 19-24 are the determined cDNA sequences of further splice variants of B726P.

SEQ ID NO: 25-29 are the amino acid sequences encoded by SEQ ID NO: 19-24, respectively.

SEQ ID NO: 30 is the determined cDNA sequence for B511S.

10 SEQ ID NO: 31 is the amino acid sequence encoded by SEQ ID NO: 30.

SEQ ID NO: 32 is the determined cDNA sequence for B533S.

SEQ ID NO:33 is the DNA sequence of Lipophilin B forward primer.

SEQ ID NO:34 is the DNA sequence of Lipophilin B reverse primer.

SEQ ID NO:35 is the DNA sequence of Lipophilin B probe.

15 SEQ ID NO:36 is the DNA sequence of GABA (B899P) forward primer.

SEQ ID NO:37 is the DNA sequence of GABA (B899P) reverse primer.

SEQ ID NO:38 is the DNA sequence of GABA (B899P) probe.

SEQ ID NO:39 is the DNA sequence of B305D (C form) forward

primer.

20 SEQ ID NO:40 is the DNA sequence of B305D (C form) reverse primer.

SEQ ID NO:41 is the DNA sequence of B305D (C form) probe.

SEQ ID NO:42 is the DNA sequence of B726P forward primer.

SEQ ID NO:43 is the DNA sequence of B726P reverse primer.

SEQ ID NO:44 is the DNA sequence of B726P probe.

25 SEQ ID NO:45 is the DNA sequence of Actin forward primer.

SEQ ID NO:46 is the DNA sequence of Actin reverse primer.

SEQ ID NO:47 is the DNA sequence of Actin probe.

SEQ ID NO:48 is the DNA sequence of Mammaglobin forward primer.

SEQ ID NO:49 is the DNA sequence of Mammaglobin reverse primer.

30 SEQ ID NO:50 is the DNA sequence of Mammaglobin probe.

SEQ ID NO:51 is the DNA sequence of a second GABA (B899P)

reverse primer.

SEQ ID NO:52 is the DNA sequence of a second B726P forward primer.

WO 01/75171

PCT/US01/10631

12

- SEQ ID NO:53 is the DNA sequence of a GABA B899P-INT forward primer.
- SEQ ID NO:54 is the DNA sequence of a GABA B899P-INT reverse primer.
- 5 SEQ ID NO:55 is the DNA sequence of a GABA B899P-INT Taqman probe.
- SEQ ID NO:56 is the DNA sequence of a B305D-INT forward primer.
- SEQ ID NO:57 is the DNA sequence of a B305D-INT reverse primer.
- SEQ ID NO:58 is the DNA sequence of a B305D-INT Taqman probe.
- 10 SEQ ID NO:59 is the DNA sequence of a B726-INT forward primer.
- SEQ ID NO:60 is the DNA sequence of a B726-INT reverse primer.
- SEQ ID NO:61 is the DNA sequence of a B726-INT Taqman probe.
- SEQ ID NO:62 is the DNA sequence of a GABA B899P Taqman probe.
- SEQ ID NO:63 is the DNA sequence of a B311D forward primer.
- 15 SEQ ID NO:64 is the DNA sequence of a B311D reverse primer.
- SEQ ID NO:65 is the DNA sequence of a B311D Taqman probe.
- SEQ ID NO:66 is the DNA sequence of a B533S forward primer.
- SEQ ID NO:67 is the DNA sequence of a B533S reverse primer.
- SEQ ID NO:68 is the DNA sequence of a B533S Taqman probe.
- 20 SEQ ID NO:69 is the DNA sequence of a B511S forward primer.
- SEQ ID NO:70 is the DNA sequence of a B511S reverse primer.
- SEQ ID NO:71 is the DNA sequence of a B511S Taqman probe.
- SEQ ID NO:72 is the DNA sequence of a GABA π reverse primer.
- SEQ ID NO:73 is the full-length cDNA sequence for mammaglobin.
- 25 SEQ ID NO:74 is the determined cDNA sequence of the open reading frame encoding a mammaglobin recombinant polypeptide expressed in *E. coli*.
- SEQ ID NO:75 is the full-length cDNA sequence for GABA π .
- SEQ ID NO:76 is the full-length cDNA sequence for lipophilin B.
- SEQ ID NO:77 is the amino acid sequence encoded by the sequence of
- 30 SEQ ID NO:76.

WO 01/75171

PCT/US01/10631

13

DETAILED DESCRIPTION OF THE INVENTION

As noted above, the present invention is directed generally to methods that are suitable for the identification of tissue-specific polynucleotides as well as to methods, compositions and kits that are suitable for the diagnosis and monitoring of cancer. While certain exemplary methods, compositions and kits disclosed herein are directed to the identification, detection and monitoring of breast cancer, in particular breast cancer-specific polynucleotides, it will be understood by those of skill in the art that the present invention is generally applicable to the identification, detection and monitoring of a wide variety of cancers, and the associated over-expressed polynucleotides, including, for example, prostate cancer, breast cancer, colon cancer, ovarian cancer, lung cancer, head & neck cancer, lymphoma, leukemia, melanoma, liver cancer, gastric cancer, kidney cancer, bladder cancer, pancreatic cancer and endometrial cancer. Thus, it will be apparent that the present invention is not limited solely to the identification of breast cancer-specific polynucleotides or to the detection and monitoring of breast cancer.

Identification of Tissue-specific Polynucleotides

Certain embodiments of the present invention provide methods, compositions and kits for the detection of a cancer cell within a biological sample. These methods comprise the step of detecting one or more tissue-specific polynucleotide(s) from a patient's biological sample the over-expression of which polynucleotides indicates the presence of a cancer cell within the patient's biological sample. Accordingly, the present invention also provides methods that are suitable for the identification of tissue-specific polynucleotides. As used herein, the phrases "tissue-specific polynucleotides" or "tumor-specific polynucleotides" are meant to include all polynucleotides that are at least two-fold over-expressed as compared to one or more control tissues. As discussed in further detail herein below, over-expression of a given polynucleotide may be assessed, for example, by microarray and/or quantitative real-time polymerase chain reaction (Real-time PCRTM) methodologies.

Exemplary methods for detecting tissue-specific polynucleotides may comprise the steps of: (a) performing a genetic subtraction to identify a pool of

WO 01/75171

PCT/US01/10631

14

polynucleotides from a tissue of interest; (b) performing a DNA microarray analysis to identify a first subset of said pool of polynucleotides of interest wherein each member polynucleotide of said first subset is at least two-fold over-expressed in said tissue of interest as compared to a control tissue; and (c) performing a quantitative polymerase
5 chain reaction analysis on polynucleotides within said first subset to identify a second subset of polynucleotides that are at least two-fold over-expressed as compared to said control tissue.

Polynucleotides Generally

As used herein, the term "polynucleotide" refers generally to either DNA
10 or RNA molecules. Polynucleotides may be naturally occurring as normally found in a biological sample such as blood, serum, lymph node, bone marrow, sputum, urine and tumor biopsy samples. Alternatively, polynucleotides may be derived synthetically by, for example, a nucleic acid polymerization reaction. As will be recognized by the skilled artisan, polynucleotides may be single-stranded (coding or antisense) or double-
15 stranded, and may be DNA (genomic, cDNA or synthetic) or RNA molecules. RNA molecules 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 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
20 linked to other molecules and/or support materials.

Polynucleotides may comprise a native sequence (*i.e.* an endogenous sequence that encodes a tumor protein, such as a breast tumor protein, or a portion thereof) or may comprise a variant, or a biological or antigenic functional equivalent of such a sequence. Polynucleotide variants may contain one or more substitutions,
25 additions, deletions and/or insertions, as further described below. The term "variants" also encompasses homologous genes of xenogenic origin.

When comparing polynucleotide or polypeptide sequences, two sequences are said to be "identical" if the sequence of nucleotides or amino acids in the two sequences is the same when aligned for maximum correspondence, as described
30 below. Comparisons between two sequences are typically performed by comparing the sequences over a comparison window to identify and compare local regions of sequence

WO 01/75171

PCT/US01/10631

15

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.

- 5 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 alignment schemes described in the following references: Dayhoff, M.O. (1978) A model of evolutionary change in proteins – Matrices for detecting distant relationships.
- 10 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) 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,
- 15 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 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
- 20 conducted by the local identity algorithm of Smith and Waterman (1981) *Add. APL. Math* 2:482, by the identity alignment algorithm of Needleman and Wunsch (1970) *J. 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
- 25 Software Package, Genetics Computer Group (GCG), 575 Science Dr., Madison, WI), 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
- 30 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

WO 01/75171

PCT/US01/10631

16

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. In one illustrative example, cumulative scores can be calculated using, for nucleotide sequences, the parameters M (reward score for a pair of
5 matching residues; always >0) and N (penalty score for mismatching residues; always <0). 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 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
10 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,
15 (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 or polypeptide sequence in the comparison window may comprise additions or deletions (*i.e.*, gaps) of 20 percent or
20 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 or amino acid residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the
25 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.

Therefore, the present invention encompasses polynucleotide and polypeptide sequences having substantial identity to the sequences disclosed herein, for example those comprising at least 50% sequence identity, preferably at least 55%, 60%,
30 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% or higher, sequence identity compared to a polynucleotide or polypeptide sequence of this invention using

WO 01/75171

PCT/US01/10631

17

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 corresponding identity of proteins encoded by two nucleotide sequences by taking into account codon degeneracy, amino acid similarity, 5 reading frame positioning and the like.

In additional embodiments, the present invention provides isolated polynucleotides and polypeptides 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 10 least about 15, 20, 30, 40, 50, 75, 100, 150, 200, 300, 400, 500 or 1000 or more 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, 15 *etc.*; 150, 151, 152, 153, *etc.*; including all integers through 200-500; 500-1,000, and the like.

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 20 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 DNA segments with total lengths of about 10,000, about 5000, 25 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.

In other embodiments, the present invention is directed to polynucleotides that are capable of hybridizing under moderately stringent conditions to 30 a polynucleotide sequence provided herein, or a fragment thereof, or a complementary sequence thereof. Hybridization techniques are well known in the art of molecular

WO 01/75171

PCT/US01/10631

18

biology. For purposes of illustration, suitable moderately stringent conditions for 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-65°C, 5 X SSC, overnight; followed by washing
5 twice at 65°C for 20 minutes with each of 2X, 0.5X and 0.2X SSC containing 0.1% SDS.

Moreover, 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
10 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
15 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).

Microarray Analyses

20 Polynucleotides that are suitable for detection according to the methods of

the present invention may be identified, as described in more detail below, by screening a microarray of cDNAs for tissue and/or tumor-associated expression (*e.g.*, expression that is at least two-fold greater in a tumor than in normal tissue, as determined using a
25 representative assay provided herein). Such screens may be performed, for example, using a Synteni microarray (Palo Alto, CA) according to the manufacturer's instructions (and essentially as described by Schena *et al.*, *Proc. Natl. Acad. Sci. USA* 93:10614-10619 (1996) and Heller *et al.*, *Proc. Natl. Acad. Sci. USA* 94:2150-2155 (1997)).

30 Microarray is an effective method for evaluating large numbers of genes but due to its limited sensitivity it may not accurately determine the absolute tissue

WO 01/75171

PCT/US01/10631

19

distribution of low abundance genes or may underestimate the degree of overexpression of more abundant genes due to signal saturation. For those genes showing overexpression by microarray expression profiling, further analysis was performed using quantitative RT-PCR based on Taqman™ probe detection, which comprises a greater dynamic range of sensitivity. Several different panels of normal and tumor tissues, distant metastases and cell lines were used for this purpose.

Quantitative Real-time Polymerase Chain Reaction

Suitable polynucleotides according to the present invention may be further characterized or, alternatively, originally identified by employing a quantitative PCR methodology such as, for example, the Real-time PCR methodology. By this methodology, tissue and/or tumor samples, such as, e.g., metastatic tumor samples, may be tested along side the corresponding normal tissue sample and/or a panel of unrelated normal tissue samples.

Real-time PCR (see Gibson et al., *Genome Research* 6:995-1001, 1996; Heid et al., *Genome Research* 6:986-994, 1996) is a technique that evaluates the level of PCR product accumulation during amplification. This technique permits quantitative evaluation of mRNA levels in multiple samples. Briefly, mRNA is extracted from tumor and normal tissue and cDNA is prepared using standard techniques.

Real-time PCR may, for example, be performed either on the ABI 7700 Prism or on a GeneAmp® 5700 sequence detection system (PE Biosystems, Foster City, CA). The 7700 system uses a forward and a reverse primer in combination with a specific probe with a 5' fluorescent reporter dye at one end and a 3' quencher dye at the other end (Taqman™). When the Real-time PCR is performed using Taq DNA polymerase with 5'-3' nuclease activity, the probe is cleaved and begins to fluoresce allowing the reaction to be monitored by the increase in fluorescence (Real-time). The 5700 system uses SYBR® green, a fluorescent dye, that only binds to double stranded DNA, and the same forward and reverse primers as the 7700 instrument. Matching primers and fluorescent probes may be designed according to the primer express program (PE Biosystems, Foster City, CA). Optimal concentrations of primers and probes are initially determined by those of ordinary skill in the art. Control (e.g., β-

WO 01/75171

PCT/US01/10631

20

actin) primers and probes may be obtained commercially from, for example, Perkin Elmer/Applied Biosystems (Foster City, CA).

To quantitate the amount of specific RNA in a sample, a standard curve is

5 generated using a plasmid containing the gene of interest. Standard curves are generated using the Ct values determined in the real-time PCR, which are related to the initial cDNA concentration used in the assay. Standard dilutions ranging from 10^{-10} to 10^{-6} copies of the gene of interest are generally sufficient. In addition, a standard curve is generated for the control sequence. This permits standardization of initial RNA content of a tissue sample to the amount of control for comparison purposes.

In accordance with the above, and as described further below, the present

invention provides the illustrative breast tissue- and/or tumor-specific polynucleotides mammaglobin, lipophilin B, GABA π (B899P), B726P, B511S, B533S, B305D and
15 B311D having sequences set forth in SEQ ID NO: 1, 3, 5-7, 11, 13, 15, 17, 19-24, 30, 32, and 73-76 illustrative polypeptides encoded thereby having amino acid sequences set forth in SEQ ID NO: 2, 4, 8-10, 12, 14, 16, 18, 25-29 and 31 and 77 that may be suitably employed in the detection of cancer, more specifically, breast cancer.

The methods disclosed herein will also permit the identification of
20 additional and/or alternative polynucleotides that are suitable for the detection of a wide range of cancers including, but not limited to, prostate cancer, breast cancer, colon cancer, ovarian cancer, lung cancer head & neck cancer, lymphoma, leukemia, melanoma, liver cancer, gastric cancer, kidney cancer, bladder cancer, pancreatic cancer and endometrial cancer.

25

Methodologies for the Detection of Cancer

In general, a cancer cell may be detected in a patient based on the presence of one or more polynucleotides within cells of a biological sample (for example, blood, lymph nodes, bone marrow, sera, sputum, urine and/or tumor biopsies)
30 obtained from the patient. In other words, such polynucleotides may be used as markers to indicate the presence or absence of a cancer such as, *e.g.*, breast cancer.

WO 01/75171

PCT/US01/10631

21

As discussed in further detail herein, the present invention achieves these and other related objectives by providing a methodology for the simultaneous detection of more than one polynucleotide, the presence of which is diagnostic of the presence of cancer cells in a biological sample. Each of the various cancer detection methodologies disclosed herein have in common a step of hybridizing one or more oligonucleotide primers and/or probes, the hybridization of which is demonstrative of the presence of a tumor- and/or tissue-specific polynucleotide. Depending on the precise application contemplated, it may be preferred to employ one or more intron-spanning oligonucleotides that are inoperative against polynucleotide of genomic DNA and, thus, these oligonucleotides are effective in substantially reducing and/or eliminating the detection of genomic DNA in the biological sample.

Further disclosed herein are methods for enhancing the sensitivity of these detection methodologies by subjecting the biological samples to be tested to one or more cell capture and/or cell depletion methodologies.

By certain embodiments of the present invention, the presence of a cancer cell in a patient may be determined by employing the following steps: (a) obtaining a biological sample from said patient; (b) contacting said biological sample with a first oligonucleotide that hybridizes to a first polynucleotide said first polynucleotide selected from the group consisting of polynucleotides depicted in SEQ ID NO:73 and SEQ ID NO:74; (c) contacting said biological sample with a second oligonucleotide that hybridizes to a second polynucleotide selected from the group consisting of SEQ ID NO: 1, 3, 5-7, 11, 13, 15, 17, 19-24, 30, 32, and 75; (d) detecting in said sample an amount of a polynucleotide that hybridizes to at least one of the oligonucleotides; and (e) comparing the amount of the polynucleotide that hybridizes to said oligonucleotide to a predetermined cut-off value, and therefrom determining the presence or absence of a cancer in the patient.

Alternative embodiments of the present invention provide methods wherein the presence of a cancer cell in a patient is determined by employing the steps of: (a) obtaining a biological sample from said patient; (b) contacting said biological sample with a first oligonucleotide that hybridizes to a first polynucleotide said first polynucleotide depicted in SEQ ID NO:76; (c) contacting said biological sample with a

WO 01/75171

PCT/US01/10631

22

second oligonucleotide that hybridizes to a second polynucleotide selected from the group consisting of SEQ ID NO: 1, 3, 5-7, 11, 13, 15, 17, 19-24, 30, 32, and 75; (d) detecting in said sample an amount of a polynucleotide that hybridizes to at least one of the oligonucleotides; and (e) comparing the amount of the polynucleotide that
5 hybridizes to said oligonucleotide to a predetermined cut-off value, and therefrom determining the presence or absence of a cancer in the patient.

Other embodiments of the present invention provide methods for determining the presence or absence of a cancer in a patient. Such methods comprise the steps of: (a) obtaining a biological sample from said patient; (b) contacting said
10 biological sample obtained from a patient with a first oligonucleotide that hybridizes to a polynucleotide sequence selected from the group consisting of polynucleotides depicted in SEQ ID NO:73, SEQ ID NO:74 and SEQ ID NO:76; (c) contacting said biological sample with a second oligonucleotide that hybridizes to a polynucleotide as depicted in SEQ ID NO:75; (d) contacting said biological sample with a third
15 oligonucleotide that hybridizes to a polynucleotide selected from the group consisting of polynucleotides depicted in SEQ ID NO:5, SEQ ID NO:6 and SEQ ID NO:7; (e) contacting said biological sample with a fourth oligonucleotide that hybridizes to a polynucleotide selected from the group consisting of polynucleotides depicted in SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20; SEQ
20 ID NO:21, SEQ ID NO:22, SEQ ID NO:23 and SEQ ID NO:24; (f) detecting in said biological sample an amount of a polynucleotide that hybridizes to at least one of said oligonucleotides; and (g) comparing the amount of polynucleotide that hybridizes to the oligonucleotide to a predetermined cut-off value, and therefrom determining the presence or absence of a cancer in the patient.

25 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 breast tumor protein that is at least 10 nucleotides, and preferably at least 20 nucleotides, in length. Preferably,
30 oligonucleotide primers hybridize to a polynucleotide encoding a polypeptide described herein under moderately stringent conditions, as defined above. Oligonucleotide

WO 01/75171

PCT/US01/10631

23

primers 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 recited in SEQ ID NO: 1, 3, 5-7, 11, 13, 15, 17, 19-24, 30, 32 and 73-76. 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*).

The present invention also provides amplification-based methods for detecting the presence of a cancer cell in a patient. Exemplary methods comprise the steps of (a) obtaining a biological sample from a patient; (b) contacting the biological sample with a first oligonucleotide pair the first pair comprising a first oligonucleotide and a second oligonucleotide wherein the first oligonucleotide and the second oligonucleotide hybridize to a first polynucleotide and the complement thereof, respectively; (c) contacting the biological sample with a second oligonucleotide pair the second pair comprising a third oligonucleotide and a fourth oligonucleotide wherein the third and the fourth oligonucleotide hybridize to a second polynucleotide and the complement thereof, respectively, and wherein the first polynucleotide is unrelated in nucleotide sequence to the second polynucleotide; (d) amplifying the first polynucleotide and the second polynucleotide; and (e) detecting the amplified first polynucleotide and the amplified second polynucleotide; wherein the presence of the amplified first polynucleotide or the amplified second polynucleotide indicates the presence of a cancer cell in the patient.

Methods according to the present invention are suitable for identifying polynucleotides obtained from a wide variety of biological sample such as, for example, blood, serum, lymph node, bone marrow, sputum, urine and tumor biopsy sample. In certain preferred embodiments, the biological sample is either blood, a lymph node or bone marrow. In other embodiments of the present invention, the lymph node may be a sentinel lymph node.

It will be apparent that the present methods may be employed in the detection of a wide variety of cancers. Exemplary cancers include, but are not limited

WO 01/75171

PCT/US01/10631

24

to, prostate cancer, breast cancer, colon cancer, ovarian cancer, lung cancer head & neck cancer, lymphoma, leukemia, melanoma, liver cancer, gastric cancer, kidney cancer, bladder cancer, pancreatic cancer and endometrial cancer.

Certain exemplary embodiments of the present invention provide
 5 methods wherein the polynucleotides to be detected are selected from the group consisting of mammaglobin, lipophilin B, GABA π (B899P), B726P, B511S, B533S, B305D and B311D. Alternatively and/or additionally, polynucleotides to be detected may be selected from the group consisting of those depicted in SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:6,
 10 SEQ ID NO:7, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:30, SEQ ID NO:32, and SEQ ID NO:76.

Suitable exemplary oligonucleotide probes and/or primers that may be used according to the methods of the present invention are disclosed herein by SEQ ID
 15 NOs:33-35 and 63-72. In certain preferred embodiments that eliminate the background detection of genomic DNA, the oligonucleotides may be intron spanning oligonucleotides. Exemplary intron spanning oligonucleotides suitable for the detection of various polynucleotides disclosed herein are depicted in SEQ ID NOs:36-62.

20 Depending on the precise application contemplated, the artisan may prefer to detect the tissue- and/or tumor-specific polynucleotides by detecting a radiolabel and detecting a fluorophore. More specifically, the oligonucleotide probe and/or primer may comprises a detectable moiety such as, for example, a radiolabel and/or a fluorophore.

25 Alternatively or additionally, methods of the present invention may also comprise a step of fractionation prior to detection of the tissue- and/or tumor-specific polynucleotides such as, for example, by gel electrophoresis.

In other embodiments, methods described herein may be used as to monitor the progression of cancer. By these embodiments, assays as provided for the
 30 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

WO 01/75171

PCT/US01/10631

25

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 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
5 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 binding agent may then be detected directly or indirectly via a reporter group. Such binding agents may also be used in histological applications. Alternatively,
10 polynucleotide probes may be used within such applications.

As noted above, to improve sensitivity, multiple breast tumor protein markers may be assayed within a given sample. It will be apparent that binding agents specific 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
15 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.

WO 01/75171

PCT/US01/10631

26

Cell Enrichment

In other aspects of the present invention, cell capture technologies may be used prior to polynucleotide detection to improve the sensitivity of the various detection methodologies disclosed herein.

Exemplary cell enrichment methodologies employ immunomagnetic beads that are coated with specific monoclonal antibodies to surface cell markers, or tetrameric antibody complexes, may be used to first enrich or positively select cancer cells in a sample. Various commercially available kits may be used, including Dynabeads® Epithelial Enrich (DynaL Biotech, Oslo, Norway), StemSep™ (StemCell Technologies, Inc., Vancouver, BC), and RosetteSep (StemCell Technologies). The skilled artisan will recognize that other readily available methodologies and kits may also be suitably employed to enrich or positively select desired cell populations.

Dynabeads® Epithelial Enrich contains magnetic beads coated with mAbs specific for two glycoprotein membrane antigens expressed on normal and neoplastic epithelial tissues. The coated beads may be added to a sample and the sample then applied to a magnet, thereby capturing the cells bound to the beads. The unwanted cells are washed away and the magnetically isolated cells eluted from the beads and used in further analyses.

RosetteSep can be used to enrich cells directly from a blood sample and consists of a cocktail of tetrameric antibodies that target a variety of unwanted cells and crosslinks them to glycophorin A on red blood cells (RBC) present in the sample, forming rosettes. When centrifuged over Ficoll, targeted cells pellet along with the free RBC.

The combination of antibodies in the depletion cocktail determines which cells will be removed and consequently which cells will be recovered. Antibodies that are available include, but are not limited to: CD2, CD3, CD4, CD5, CD8, CD10, CD11b, CD14, CD15, CD16, CD19, CD20, CD24, CD25, CD29, CD33, CD34, CD36, CD38, CD41, CD45, CD45RA, CD45RO, CD56, CD66B, CD66c, HLA-DR, IgE, and TCRαβ. Additionally, it is contemplated in the present invention that mAbs specific for breast tumor antigens, can be developed and used in a similar

WO 01/75171

PCT/US01/10631

27

manner. For example, mAbs that bind to tumor-specific cell surface antigens may be conjugated to magnetic beads, or formulated in a tetrameric antibody complex, and used to enrich or positively select metastatic breast tumor cells from a sample.

Once a sample is enriched or positively selected, cells may be further
5 analysed. For example, the cells may be lysed and RNA isolated. RNA may then be subjected to RT-PCR analysis using breast tumor-specific multiplex primers in a Real-time PCR assay as described herein.

In another aspect of the present invention, cell capture technologies may be used in conjunction with Real-time PCR to provide a more sensitive tool for
10 detection of metastatic cells expressing breast tumor antigens. Detection of breast cancer cells in bone marrow samples, peripheral blood, and small needle aspiration samples is desirable for diagnosis and prognosis in breast cancer patients.

Probes and Primers

As noted above and as described in further detail herein, certain
15 methods, compositions and kits according to the present invention utilize two or more oligonucleotide primer pairs for the detection of cancer. 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 biological sample.

Alternatively, in other embodiments, the probes and/or primers of the
20 present invention may be employed for detection via 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 of a polynucleotide to be detected will find particular utility. Longer contiguous identical or complementary
25 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.

Oligonucleotide primers having sequence regions consisting of contiguous nucleotide stretches of 10-14, 15-20, 30, 50, or even of 100-200 nucleotides
30 or so (including intermediate lengths as well), identical or complementary to a polynucleotide to be detected, are particularly contemplated as primers for use in

WO 01/75171

PCT/US01/10631

28

amplification reactions such as, *e.g.*, the polymerase chain reaction (PCRTM). This would allow a polynucleotide to be analyzed, both in diverse biological samples such as, for example, blood, lymph nodes and bone marrow. .

The use of a primer of about 15-25 nucleotides in length allows the
5 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 primers having gene-complementary stretches of 15 to
10 25 contiguous nucleotides, or even longer where desired.

Primers may be selected from any portion of the polynucleotide to be detected. All that is required is to review the sequence, such as those exemplary polynucleotides set forth in SEQ ID NO: 1, 3, 5-7, 11, 13, 15, 17, 19-24, 30, 32, 73-75 (Figures 3-6, respectively) and SEQ ID NO:76 (lipophilin B) or to any continuous
15 portion of the sequence, from about 15-25 nucleotides in length up to and including the full length sequence, that one wishes to utilize as a primer. The choice of primer sequences may be governed by various factors. For example, one may wish to employ primers from towards the termini of the total sequence. The exemplary primers disclosed herein may optionally be used for their ability to selectively form duplex
20 molecules with complementary stretches of the entire polynucleotide of interest such as those set forth in SEQ ID NO: 1, 3, 5-7, 11, 13, 15, 17, 19-24, 30, 32, 73-75 (Figures 3-6, respectively), and SEQ ID NO:76 (lipophilin B).

The present invention further provides the nucleotide sequence of various exemplary oligonucleotide primers and probes, set forth in SEQ ID NOs: 33-71,
25 that may be used, as described in further detail herein, according to the methods of the present invention for the detection of cancer.

Oligonucleotide primers according to the present invention may be readily prepared routinely by methods commonly available to the skilled artisan including, for example, directly synthesizing the primers by chemical means, as is
30 commonly practiced using an automated oligonucleotide synthesizer. Depending on the application envisioned, one will typically desire to employ varying conditions of

WO 01/75171

PCT/US01/10631

29

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.

10 Polynucleotide Amplification Techniques

Each of the specific embodiments outlined herein for the detection of cancer has in common the detection of a tissue- and/or tumor-specific polynucleotide via the hybridization of one or more oligonucleotide primers and/or probes. Depending on such factors as the relative number of cancer cells present in the biological sample and/or the level of polynucleotide expression within each cancer cell, it may be preferred to perform an amplification step prior to performing the steps of detection. For example, at least two oligonucleotide primers may be employed in a polymerase chain reaction (PCR) based assay to amplify a portion of a breast 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 breast tumor protein. The amplified cDNA may optionally be subjected to a fractionation step such as, for example, gel electrophoresis.

A number of template dependent processes are available to amplify the target sequences of interest present in a sample. One of the best known amplification methods is the polymerase chain reaction (PCRTM) 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 PCRTM, 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 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

WO 01/75171

PCT/US01/10631

30

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

One preferred methodology for polynucleotide amplification employs RT-PCR, in which PCR is applied in conjunction with reverse transcription. Typically, RNA is extracted from a biological sample, such as blood, serum, lymph node, bone marrow, sputum, urine and tumor biopsy samples, 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 patient and from an individual who is not afflicted with a cancer. The amplification reaction may be 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.

Any of a variety of commercially available kits may be used to perform the amplification step. 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

WO 01/75171

PCT/US01/10631

31

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 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.

Another method for amplification is the ligase chain reaction (referred to as LCR), disclosed in Eur. Pat. Appl. Publ. No. 320,308 (specifically incorporated herein by reference in its entirety). In LCR, two complementary probe pairs are prepared, and in the presence of the target sequence, each pair will bind to opposite complementary strands of the target such that they abut. In the presence of a ligase, the two probe pairs will link to form a single unit. By temperature cycling, as in PCRTM, bound ligated units dissociate from the target and then serve as "target sequences" for ligation of excess probe pairs. U.S. Patent No. 4,883,750, incorporated herein by reference in its entirety, describes an alternative method of amplification similar to LCR for binding probe pairs to a target sequence.

Qbeta Replicase, described in PCT Intl. Pat. Appl. Publ. No. PCT/US87/00880, incorporated herein by reference in its entirety, may also be used as still another amplification method in the present invention. In this method, a replicative sequence of RNA that has a region complementary to that of a target is added to a sample in the presence of an RNA polymerase. The polymerase will copy the replicative sequence that can then be detected.

An isothermal amplification method, in which restriction endonucleases and ligases are used to achieve the amplification of target molecules that contain nucleotide 5'-[α-thio]triphosphates in one strand of a restriction site (Walker *et al.*, 1992, incorporated herein by reference in its entirety), may also be useful in the amplification of nucleic acids in the present invention.

Strand Displacement Amplification (SDA) is another method of carrying out isothermal amplification of nucleic acids which involves multiple rounds of strand displacement and synthesis, *i.e.* nick translation. A similar method, called Repair Chain Reaction (RCR) is another method of amplification which may be useful in the present

WO 01/75171

PCT/US01/10631

32

invention and is involves annealing several probes throughout a region targeted for amplification, followed by a repair reaction in which only two of the four bases are present. The other two bases can be added as biotinylated derivatives for easy detection. A similar approach is used in SDA.

5 Sequences can also be detected using a cyclic probe reaction (CPR). In CPR, a probe having a 3' and 5' sequences of non-target DNA and an internal or "middle" sequence of the target protein specific RNA is hybridized to DNA which is present in a sample. Upon hybridization, the reaction is treated with RNaseH, and the products of the probe are identified as distinctive products by generating a signal that is released after digestion. The original template is annealed to another cycling probe and the reaction is repeated. Thus, CPR involves amplifying a signal generated by
10 hybridization of a probe to a target gene specific expressed nucleic acid.

Still other amplification methods described in Great Britain Pat. Appl. No. 2 202 328, and in PCT Intl. Pat. Appl. Publ. No. PCT/US89/01025, each of which
15 is incorporated herein by reference in its entirety, may be used in accordance with the present invention. In the former application, "modified" primers are used in a PCR-like, template and enzyme dependent synthesis. The primers may be modified by labeling with a capture moiety (*e.g.*, biotin) and/or a detector moiety (*e.g.*, enzyme). In the latter application, an excess of labeled probes is added to a sample. In the presence
20 of the target sequence, the probe binds and is cleaved catalytically. After cleavage, the target sequence is released intact to be bound by excess probe. Cleavage of the labeled probe signals the presence of the target sequence.

Other nucleic acid amplification procedures include transcription-based amplification systems (TAS) (Kwoh *et al.*, 1989; PCT Intl. Pat. Appl. Publ. No. WO
25 88/10315, incorporated herein by reference in its entirety), including nucleic acid sequence based amplification (NASBA) and 3SR. In NASBA, the nucleic acids can be prepared for amplification by standard phenol/chloroform extraction, heat denaturation of a sample, treatment with lysis buffer and minispin columns for isolation of DNA and RNA or guanidinium chloride extraction of RNA. These amplification techniques
30 involve annealing a primer that has sequences specific to the target sequence. Following polymerization, DNA/RNA hybrids are digested with RNase H while double

WO 01/75171

PCT/US01/10631

33

stranded DNA molecules are heat-denatured again. In either case the single stranded DNA is made fully double stranded by addition of second target-specific primer, followed by polymerization. The double stranded DNA molecules are then multiply transcribed by a polymerase such as T7 or SP6. In an isothermal cyclic reaction, the
5 RNAs are reverse transcribed into DNA, and transcribed once again with a polymerase such as T7 or SP6. The resulting products, whether truncated or complete, indicate target-specific sequences.

Eur. Pat. Appl. Publ. No. 329,822, incorporated herein by reference in its entirety, disclose a nucleic acid amplification process involving cyclically synthesizing
10 single-stranded RNA ("ssRNA"), ssDNA, and double-stranded DNA (dsDNA), which may be used in accordance with the present invention. The ssRNA is a first template for a first primer oligonucleotide, which is elongated by reverse transcriptase (RNA-dependent DNA polymerase). The RNA is then removed from resulting DNA:RNA duplex by the action of ribonuclease H (RNase H, an RNase specific for
15 RNA in a duplex with either DNA or RNA). The resultant ssDNA is a second template for a second primer, which also includes the sequences of an RNA polymerase promoter (exemplified by T7 RNA polymerase) 5' to its homology to its template. This primer is then extended by DNA polymerase (exemplified by the large "Klenow" fragment of *E. coli* DNA polymerase I), resulting as a double-stranded DNA
20 ("dsDNA") molecule, having a sequence identical to that of the original RNA between the primers and having additionally, at one end, a promoter sequence. This promoter sequence can be used by the appropriate RNA polymerase to make many RNA copies of the DNA. These copies can then re-enter the cycle leading to very swift amplification. With proper choice of enzymes, this amplification can be done
25 isothermally without addition of enzymes at each cycle. Because of the cyclical nature of this process, the starting sequence can be chosen to be in the form of either DNA or RNA.

PCT Intl. Pat. Appl. Publ. No. WO 89/06700, incorporated herein by reference in its entirety, disclose a nucleic acid sequence amplification scheme based on
30 the hybridization of a promoter/primer sequence to a target single-stranded DNA ("ssDNA") followed by transcription of many RNA copies of the sequence. This

WO 01/75171

PCT/US01/10631

34

scheme is not cyclic; *i.e.* new templates are not produced from the resultant RNA transcripts. Other amplification methods include "RACE" (Frohman, 1990), and "one-sided PCR" (Ohara, 1989) which are well-known to those of skill in the art.

Compositions and Kits for the Detection of Cancer

5 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 breast
10 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.

15 The present invention also provides kits that are suitable for performing the detection methods of the present invention. Exemplary kits comprise oligonucleotide primer pairs each one of which specifically hybridizes to a distinct polynucleotide. Within certain embodiments, kits according to the present invention may also comprise a nucleic acid polymerase and suitable buffer. Exemplary
20 oligonucleotide primers suitable for kits of the present invention are disclosed herein by SEQ ID NOs: 33-71. Exemplary polynucleotides suitable for kits of the present invention are disclosed in SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, SEQ
25 ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:30, SEQ ID NO:32, and lipophilin B.

 Alternatively, a kit may be designed to detect the level of mRNA encoding a breast tumor protein in a biological sample. Such kits generally comprise at least one oligonucleotide probe or primer, as described above, that hybridizes to a
30 polynucleotide encoding a breast tumor protein. Such an oligonucleotide may be used,

WO 01/75171

PCT/US01/10631

35

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 breast tumor protein.

In other related aspects, the present invention further provides
5 compositions useful in the methods disclosed herein. Exemplary compositions comprise two or more oligonucleotide primer pairs each one of which specifically hybridizes to a distinct polynucleotide. Exemplary oligonucleotide primers suitable for compositions of the present invention are disclosed herein by SEQ ID NOs: 33-71. Exemplary polynucleotides suitable for compositions of the present invention are
10 disclosed in SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:30, SEQ ID NO:32, and lipophilin B.

15 The following Examples are offered by way of illustration and not by way of limitation.

WO 01/75171

PCT/US01/10631

36

EXAMPLES

EXAMPLE 1

DIFFERENTIAL DISPLAY

5 This example discloses the use of differential display to enrich for polynucleotides that are over-expressed in breast tumor tissues.

Differential display was performed as described in the literature (*see, e.g., Liang, P. et al., Science 257:967-971 (1993)*, incorporated herein by reference in its entirety) with the following modifications: (a) PCR amplification products were
10 visualized on silver stained gels (b) genetically matched pairs of tissues were used to eliminate polymorphic variation (c) two different dilutions of cDNA were used as template to eliminate any dilutional effects (*see, Mou, E. et al., Biochem Biophys Res Commun. 199:564-569 (1994)*, incorporated herein by reference in its entirety).

15

EXAMPLE 2

PREPARATION OF cDNA SUBTRACTION LIBRARY

This example discloses the preparation of a breast tumor cDNA subtraction library enriched in breast tumor specific polynucleotides.

cDNA library subtraction was performed as described with some
20 modification. *See, Hara, T. et al., Blood 84: 189-199 (1994)*, incorporated herein by reference in its entirety. The breast tumor library (tracer) that was made from a pool of three breast tumors was subtracted with normal breast library (driver) to identify breast tumor specific genes. More recent subtractions utilized 6-10 normal tissues as driver to subtract out common genes more efficiently, with an emphasis on essential tissues
25 along with one "immunological" tissue (*e.g., spleen, lymph node, or PBMC*), to assist in the removal of cDNAs derived from lymphocyte infiltration in tumors. The breast tumor specific subtracted cDNA library was generated as follows: driver cDNA library was digested with EcoRI, NotI, and SfuI (SfuI cleaves the vector), filled in with DNA polymerase klenow fragment. After phenol-chloroform extraction and ethanol
30 precipitation, the DNA was labeled with Photoprobe biotin and dissolved in H₂O. Tracer cDNA library was digested with BamHI and XhoI, phenol chloroform extracted,

WO 01/75171

PCT/US01/10631

37

passed through Chroma spin-400 columns, ethanol precipitated, and mixed with driver DNA for hybridization at 68°C for 20 hours [long hybridization (LH)]. The reaction mixture was then subjected to the streptavidin treatment followed by phenol/chloroform extraction for a total of four times. Subtracted DNA was precipitated and subjected to a
5 hybridization at 68°C for 2 hours with driver DNA again [short hybridization (SH)]. After removal of biotinylated double-stranded DNA, subtracted cDNA was ligated into BamHI/XhoI site of Chloramphenicol resistant pBCSK⁺ and transformed into ElectroMax *E. coli* DH10B cells by electroporation to generate subtracted cDNA library. To clone less abundant breast tumor specific genes, cDNA library subtraction
10 was repeated by subtracting the tracer cDNA library with the driver cDNA library plus abundant cDNAs from primary subtractions. This resulted in the depletion of these abundant sequences and the generation of subtraction libraries that contain less abundant sequences.

To analyze the subtracted cDNA library, plasmid DNA was prepared
15 from 100-200 independent clones, which were randomly picked from the subtracted library, and characterized by DNA sequencing. The determined cDNA and expected amino acid sequences for the isolated cDNAs were compared to known sequences using the most recent Genbank and human EST databases.

20

EXAMPLE 3

PCR-SUBTRACTION

This example discloses PCR subtraction to enrich for breast tumor
specific
25 polynucleotides.

PCR-subtraction was performed essentially as described in the literature.
See, Diatchenko, L. et al., *Proc Natl Acad Sci U S A.* 93:6025-6030 (1996) and Yang, G.P. et al., *Nucleic acids Res.* 27:1517-23 (1999), incorporated herein by reference in their entirety. Briefly, this type of subtraction works by ligating two different adapters
30 to different aliquots of a restriction enzyme digested tester (breast tumor) cDNA sample, followed by mixing of the testers separately with excess driver (without

WO 01/75171

PCT/US01/10631

38

adapters). This first hybridization results in normalization of single stranded tester specific cDNA due to the second order kinetics of hybridization. These separate hybridization reactions are then mixed without denaturation, and a second hybridization performed which produces the target molecules; double stranded cDNA fragments containing both of the different adapters. Two rounds of PCR were performed, which results in the exponential amplification of the target population molecules (normalized tester specific cDNAs), while other fragments were either unamplified or only amplified in a linear manner. The subtractions performed included a pool of breast tumors subtracted with a pool of normal breast and a pool of breast tumors subtracted with a pool of normal tissues including PBMC, brain, pancreas, liver, small intestine, stomach, heart and kidney.

Prior to cDNA synthesis RNA was treated with DNase I (Ambion) in the presence of RNasin (Promega Biotech, Madison, WI) to remove DNA contamination. The cDNA for use in real-time PCR tissue panels was prepared using 25µl Oligo dT (Boehringer-Mannheim) primer with superscript II reverse transcriptase (Gibco BRL, Bethesda, MD).

EXAMPLE 4

DETECTION OF BREAST CANCER USING BREAST-SPECIFIC ANTIGENS

The isolation and characterization of the breast-specific antigens B511S and B533S is described in U.S. Patent Application 09/346,327, filed July 2, 1999, the disclosure of which is hereby incorporated by reference in its entirety. The determined cDNA sequence for B511S is provided in SEQ ID NO: 30, with the corresponding amino acid sequence being provided in SEQ ID NO: 31. The determined cDNA sequence for B533S is provided in SEQ ID NO: 32. The isolation and characterization of the breast-specific antigen B726P is described in U.S. Patent Applications 09/285,480, filed April 2, 1999, and 09/433,826, filed November 3, 1999, the disclosures of which are hereby incorporated by reference in their entirety.

The determined cDNA sequences for splice variants of B726P are provided in SEQ ID NO: 13, 15, 17 and 19-24, with the corresponding amino acid sequences being provided in SEQ ID NO: 14, 16, 18 and 25-29.

WO 01/75171

PCT/US01/10631

39

The isolation and characterization of the breast-specific antigen B305D forms A and C has been described in U.S. Patent Application 09/429,755, filed October 28, 1999, the disclosure of which is hereby incorporated by reference in its entirety. Determined cDNA sequences for B305D isoforms A and C are provided in SEQ ID NO: 1, 3 and 5-7, with the corresponding amino acid sequences being provided in SEQ ID NO: 2, 4 and 8-10.

The isolation and characterization of the breast-specific antigen B311D has been described in U.S. Patent Application 09/289,198, filed April 9, 1999, the disclosure of which is hereby incorporated by reference in its entirety. The determined cDNA sequence for B311D is provided in SEQ ID NO:11, with the corresponding amino acid sequence being provided in SEQ ID NO:12.

cDNA sequences for mammaglobin are provided in Figs. 4 and 5, with the cDNA sequence for GABA π being provided in Fig 6 and are disclosed in SEQ ID NOs: 73-75, respectively.

The isolation and characterization of the breast-specific antigen lipophilin B has been described in U.S. Patent Application 09/780,842, filed February 8, 2001, the disclosure of which is hereby incorporated by reference in its entirety. The determined cDNA sequence for lipophilin B is provided in SEQ ID NO:76, with the corresponding amino acid sequence being provided in SEQ ID NO:77. The nucleotide sequences of several sequence variants of lipophilin B are also described in the 09/780,842 application.

EXAMPLE 5

MICROARRAY ANALYSIS

This example discloses the use of microarray analyses to identify polynucleotides that are at least two-fold overexpressed in breast tumor tissue samples as compared to normal breast tissue samples.

mRNA expression of the polynucleotides of interest was performed as follows. cDNA for the different genes was prepared as described above and arrayed on a glass slide (Incyte, Palo Alto, CA). The arrayed cDNA was then hybridized with a 1:1 mixture of Cy3 or Cy5 fluorescent labeled first strand cDNAs obtained from

WO 01/75171

PCT/US01/10631

40

polyA+ RNA from breast tumors, normal breast and normal tissues and other tumors as described in Shalon, D. et al., *Genome Res.* 6:639-45 (1996), incorporated herein by reference in its entirety. Typically Cy3 (Probe 1) was attached to cDNAs from breast tumors and Cy5 (Probe 2) to normal breast tissue or other normal tissues. Both probes
5 were allowed to compete with the immobilized gene specific cDNAs on the chip, washed then scanned for fluorescence intensity of the individual Cy3 and Cy5 fluorescence to determine extent of hybridization. Data were analyzed using GEMTOOLS software (Incyte, Palo Alto, CA) which enabled the overexpression patterns of breast tumors to be compared with normal tissues by the ratios of Cy3/Cy5.
10 The fluorescence intensity was also related to the expression level of the individual genes.

DNA microarray analyses was used primarily as a screening tool to determine tissue/tumor specificity of cDNA's recovered from the differential display, cDNA library and PCR subtractions, prior to more rigorous analysis by quantitative
15 RT-PCR, northern blotting, and immunohistochemistry. Microarray analysis was performed on two microchips. A total of 3603 subtracted cDNA's and 197 differential display templates were evaluated to identify 40 candidates for further analysis by quantitative PCR. From these candidates, several were chosen on the basis of favorable tissue specificity profiles, including B305D, B311D, B726P, B511S and B533S,
20 indicating their overexpression profiles in breast tumors and/or normal breast versus other normal tissues. It was evident that the expression of these genes showed a high degree of specificity for breast tumors and/or breast tissue. In addition, these genes have in many cases complementary expression profiles.

The two known breast-specific genes, mammaglobin and γ -aminobutyrate type A receptor π subunit (GABA π) were also subjected to microarray
25 analysis. mRNA expression of mammaglobin has been previously described to be upregulated in proliferating breast tissue, including breast tumors. See, (Watson et al., *Cancer Res.*, 56: 860-5 (1996); Watson et al., *Cancer Res.*, 59: 3028-3031 (1999); Watson et al., *Oncogene*, 16:817-24 (1998), incorporated herein by reference in their
30 entirety). The GABA π mRNA levels were over-expressed in breast tumors. Previous studies had demonstrated its overexpression in uterus and to some degree in prostate

WO 01/75171

PCT/US01/10631

41

and lung (Hedblom et al., J Biol. Chem. 272:15346-15350 (1997)) but no previous study had indicated elevated levels in breast tumors.

EXAMPLE 6

5 QUANTITATIVE REAL-TIME PCR ANALYSIS

This example discloses the use of quantitative Real-time PCR to confirm the

microarray identification polynucleotide that are at least two-fold overexpressed in breast tumor tissue samples as compared to normal breast tissue samples.

10 The tumor- and/or tissue-specificity of the polynucleotides identified by the microarray analyses disclosed herein in Example 5, were confirmed by quantitative PCR analyses. Breast metastases, breast tumors, benign breast disorders and normal breast tissue along with other normal tissues and tumors were tested in quantitative (Real time) PCR. This was performed either on the ABI 7700 Prism or on a
15 GeneAmp® 5700 sequence detection system (PE Biosystems, Foster City, CA). The 7700 system uses a forward and a reverse primer in combination with a specific probe designed to anneal to sequence between the forward and reverse primer. This probe was conjugated at the 5' end with a fluorescent reporter dye and a quencher dye at the other 3' end (Taqman™). During PCR the Taq DNA polymerase with its 5'-3'
20 nuclease activity cleaved the probe which began to fluoresce, allowing the reaction to be monitored by the increase in fluorescence (Real-time). Holland et al., *Proc Natl Acad Sci U S A.* 88:7276-7280 (1991). The 5700 system used SYBR® green, a fluorescent dye, that only binds to double stranded DNA (Schneeberger et al., *PCR Methods Appl.* 4:234-8 (1995)), and the same forward and reverse primers as the 7700
25 instrument. No probe was needed. Matching primers and fluorescent probes were designed for each of the genes according to the Primer Express program (PE Biosystems, Foster City, CA).

Table 1.
Primer and Probe Sequences for the Genes of Interest

	Forward Primer	Reverse primer	Probe
Mammaglobin	TGCCATAGATG	TGTCATATATTAA	TCTTAACCAAACG
	AATTGAAGGAA	TTGCATAAACACC	GATGAAACTCTGA
	TG (SEQ ID	TCA (SEQ ID	GCAATG (SEQ ID

WO 01/75171

PCT/US01/10631

42

	Forward Primer NO:48)	Reverse primer NO:49)	Probe NO:50)
B305D-C form	AAAGCAGATGG TGGTTGAGGTT (SEQ ID NO:39)	CCTGAGACCAAA TGGCTTCTTC (SEQ ID NO:40)	ATTCCATGCCGGC TGCTTCTTCTG (SEQ ID NO:41)
B311D	CCGCITCTGACA ACACTAGAGAT C (SEQ ID NO:63)	CCTATAAAGATGT TATGTACCAAAA ATGAAAGT (SEQ ID NO:64)	CCCCTCCCTCAGG GTATGGCCC (SEQ ID NO:65)
B726P	TCTGGTTTCTC ATTCTTTATCA TTTATT (SEQ ID NO:42)	TGCCAAGGAGCG GATTATCT (SEQ ID NO:43)	CAACCACGTGACA AACACTGGAATTA CAGG (SEQ ID NO:44)
B533S	CCCTTTCTCACC CACACACTGT (SEQ ID NO:66)	TGCATTCTCAT ATGTGGAAGCT (SEQ ID NO:67)	CCGGGCCTCAGGC ATATACTATTCTA CTGTCTG (SEQ ID NO:68)
GABA π	AAGCCTCAGAG TCCTTCCAGTAT G (SEQ ID NO:36)	AAATATAAGTGA AGAAAAAATTA GTAGAT (SEQ ID NO:72)	AATCCATTGTATC TTAGAACCAGGGG ATTTGTTTAGA (SEQ ID NO:38)
B511S	GACATTCCAGTT TTACCCAAATG G (SEQ ID NO:69)	TGCAGAAGACTC AAGCTGATTCC (SEQ ID NO:70)	TCTCAGGGACACA CTCTACCAATCGG GA (SEQ ID NO:71)

The concentrations used in the quantitative PCR for the forward primers for mammaglobin, GABA π , B305D C form, B311D, B511S, B533S and B726P were 900, 900, 300, 900, 900, 300 and 300nM respectively. For the reverse primers they were 300, 900, 900, 900, 300, 900 and 900nM respectively. Primers and probes so produced were used in the universal thermal cycling program in real-time PCR. They were titrated to determine the optimal concentrations using a checkerboard approach. A pool of cDNA from target tumors was used in this optimization process. The reaction was performed in 25 μ l volumes. The final probe concentration in all cases was 160nM. dATP, dCTP and dGTP were at 0.2mM and dUTP at 0.4mM. Amplitaq gold and Amperase UNG (PE Biosystems, Foster City, CA) were used at 0.625 units and 0.25 units per reaction. MgCl₂ was at a final concentration of 5mM. Trace amounts of glycerol, gelatin and Tween 20 (Sigma Chem Co, St Louis, MO) were added to stabilize the reaction. Each reaction contained 2 μ l of diluted template. The cDNA from RT reactions prepared as above was diluted 1:10 for the gene of interest and 1:100 for β -Actin. Primers and probes for β -Actin (PE Biosystems, Foster City, CA) were used in

WO 01/75171

PCT/US01/10631

43

a similar manner to quantitate the presence of β -actin in the samples. In the case of the SYBR® green assay, the reaction mix (25 μ l) included 2.5 μ l of SYBR green buffer, 2 μ l of cDNA template and 2.5 μ l each of the forward and reverse primers for the gene of interest. This mix also contained 3mM MgCl₂, 0.25units of AmpErase UNG, 0.625 units of Amplitaq gold, 0.08% glycerol, 0.05% gelatin, 0.0001% Tween 20 and 1mM dNTP mix. In both formats, 40 cycles of amplification were performed.

In order to quantitate the amount of specific cDNA (and hence initial mRNA) in the sample, a standard curve was generated for each run using the plasmid containing the gene of interest. Standard curves were generated using the Ct values determined in the real-time PCR which were related to the initial cDNA concentration used in the assay. Standard dilutions ranging from 20-2x10⁶ copies of the gene of interest were used for this purpose. In addition, a standard curve was generated for the housekeeping gene β -actin ranging from 200fg-2000pg to enable normalization to a constant amount of β -Actin. This allowed the evaluation of the over-expression levels seen with each of the genes.

The genes B311D, B533S and B726P were evaluated in quantitative PCR as described above on two different panels consisting of: (a) breast tumor, breast normal and normal tissues; and (b) breast tumor metastases (primarily lymph nodes), using the primers and probes shown above in Table 1. The data for panel (a) is shown in Figure 1 for all three genes. The three genes showed identical breast tissue expression profiles. However, the relative level of gene expression was very different in each case. B311D in general was expressed at lower levels than B533S and both less than B726P, but all three were restricted to breast tissue. The quantitative PCR thus confirmed there was a differential expression between normal breast tissue and breast tumors for all three genes, and that approximately 50% of breast tumors over-expressed these genes. When tested on a panel of distant metastases derived from breast cancers all three genes reacted with 14/21 metastases and presented similar profiles. All three genes also exhibited increasing levels of expression as a function of pathological stage of the tumor, as shown for B533S in Figure 2.

Mammaglobin is a homologue of a rabbit uteroglobin and the rat steroid binding protein subunit C3 and is a low molecular weight protein that is highly glycosylated. In contrast to its homologs, mammaglobin has been reported to be breast

WO 01/75171

PCT/US01/10631

44

specific and over-expression has been described in breast tumor biopsies (23%) and primary and metastatic breast tumors (~75%) with reports of the detection of mammaglobin mRNA expression in 91% of lymph nodes from metastatic breast cancer patients. However, more rigorous analysis of mammaglobin gene expression by
5 microarray and quantitative PCR as described above (panels (a) and (b) and a panel of other tumors and normal tissues and additional breast tumors), showed expression at significant levels in skin and salivary gland with much lower levels in esophagus and trachea, as shown in Table 2 below.

10

WO 01/75171

PCT/US01/10631

45

Table 2

Normalized Distribution of Mammaglobin and B511S mRNA in Various Tissues

Tissue	Mean Copies Mammaglobin /ng β -Actin \pm SD	PCR Positive	Mean Copies B511S /ng β -Actin \pm SD	PCR Positive	PCR Positive (Mammaglobin/ B511S)
Breast Tumors	1233.88 \pm 3612.74	31/42	1800.40 \pm 3893.24	33/42	38/42
Breast tumor Metastases	1912.54 \pm 4625.85	14/24	3329.50 \pm 10820.71	14/24	17/24
Benign Breast disorders	121.87 \pm 78.63	3/3	524.66 \pm 609.43	2/3	3/3
Normal breast	114.19 \pm 94.40	11/11	517.64 \pm 376.83	8/9	11/11
Breast reduction	231.50 \pm 276.68	2/3	482.54 \pm 680.28	1/2	2/3
Other tumors	0.13 \pm 0.65	1/39	24.17 \pm 36.00	5/23	
Salivary gland	435.65 \pm 705.11	2/3	45766.61 \pm 44342.43	3/3	
Skin	415.74 \pm 376.14	7/9	7039.05 \pm 7774.24	9/9	
Esophagus	4.45 \pm 3.86	2/3	1.02 \pm 0.14	0/3	
Bronchia	0.16	0/1	84.44 \pm 53.31	2/2	
Other normal tissues	0.33 \pm 1.07	0/85	5.49 \pm 10.65	3/75	

5 The breast-specific gene B511S, while having a different profile of reactivity on breast tumors and normal breast tissue to mammaglobin, reacted with the same subset of normal tissues as mammaglobin. B511S by PSORT analysis is indicated to have an ORF of 90aa and to be a secreted protein as is the case for mammaglobin. B511S has no evidence of a transmembrane domain but may harbor a

10 cleavable signal sequence. Mammaglobin detected 14/24 of distant metastatic breast tumors, 31/42 breast tumors and exhibited ten-fold over-expression in tumors and metastases as compared to normal breast tissue. There was at least 300-fold over-expression in normal breast tissue versus other negative normal tissues and tumors tested, which were essentially negative for mammaglobin expression. B511S detected

WO 01/75171

PCT/US01/10631

46

33/42 breast tumors and 14/24 distant metastases, while a combination of B511S with
mammaglobin would be predicted to detect 38/42 breast tumors and 17/24 metastatic
lesions (Table 2 above). The quantitative level of expression of B511S and
mammaglobin were also in similar ranges, in concordance with the microarray profiles
5 observed for these two genes. Other genes that were additive with mammaglobin are
shown in Table 3.

WO 01/75171

PCT/US01/10631

47

Table 3
mRNA Complementation of Mammaglobin with Other Genes

	Mammaglobin Positive	Mammaglobin Negative				
		B305D	GABA π	B726P	B305D + GABA π	B305D + GABA π + B726P
Breast Metastases	13/21	2/8	5/8	3/8	7/8	8/8
Breast tumors	18/25	3/7	4/7	5/7	7/7	7/7

B305D was shown to be highly over-expressed in breast tumors, prostate
5 tumors, normal prostate tissue and testis compared to normal tissues, including normal
breast tissue. Different splice variants of B305D have been identified with form A and
C being the most abundant but all tested have similar tissue profiles in quantitative
PCR. The A and C forms contain ORF's of 320 and 385 aa, respectively. B305D is
predicted by PSORT to be a Type II membrane protein that comprises a series of
10 ankyrin repeats. A known gene shown to be complementary with B305D, in breast
tumors, was GABA π . This gene is a member of the GABA π receptor family and
encodes a protein that has 30-40% amino acid homology with other family members,
and has been shown by Northern blot analysis to be over-expressed in lung, thymus and
prostate at low levels and highly over-expressed in uterus. Its expression in breast tissue
15 has not been previously described. This is in contrast to other GABA π receptors that
have appreciable expression in neuronal tissues. Tissue expression profiling of this
gene showed it to be over-expressed in breast tumors in an inverse relationship to the
B305D gene (Table 3). GABA π detected 15/25 tumors and 6/21 metastases including 4
tumors and 5 metastases missed by mammaglobin. In contrast, B305D detected 13/25
20 breast tumors and 8/21 metastases, again including 3 tumors and 2 metastases missed
by mammaglobin. A combination of just B305D and the GABA π would be predicted
to identify 22/25 breast tumors and 14/21 metastases. The combination of B305D and
GABA π with mammaglobin in detecting breast metastases is shown in Table 3 above
and Figures. 3A and 3B. This combination detected 20/21 of the breast metastases as
25 well as 25/25 breast tumors that were evaluated on the same panels for all three genes.

WO 01/75171

PCT/US01/10631

48

The one breast metastasis that was negative for these three genes was strongly positive for B726P (Figs. 3A and 3B).

To evaluate the presence of circulating tumor cells, an immunocapture (cell capture) method was employed to first enrich for epithelial cells prior to RT-PCR
5 analysis. Immunomagnetic polystyrene beads coated with specific monoclonal antibodies to two glycoproteins on the surface of epithelial cells were used for this purpose. Such an enrichment procedure increased the sensitivity of detection (~100 fold) as compared to direct isolation of poly A⁺ RNA, as shown in Table 4.

10

WO 01/75171

PCT/US01/10631

49

Table 4
Extraction of Mammaglobin Positive Cells (MB415) Spiked
into Whole Blood and Detection by Real-time PCR

MB415 cells/ml Blood	Epithelial cell extraction (Poly A ⁺ RNA)	Direct Extraction (Poly A ⁺ RNA)
Copies Mammaglobin/ng β -Actin		
100000	54303.2	58527.1
10000	45761.9	925.9
1000	15421.2	61.6
100	368.0	5.1
10	282.0	1.1
1	110.2	0
0	0	0

5 Mammaglobin-positive cells (MB415) were spiked into whole blood at various concentrations and then extracted using either epithelial cell enrichment or direct isolation from blood. Using enrichment procedures, mammaglobin mRNA was found to be detectable at much lower levels than when direct isolation was used. Whole blood samples from patients with metastatic breast cancer were subsequently treated

10 with the immunomagnetic beads. Poly A⁺ RNA was then isolated, cDNA prepared and run in quantitative PCR using two gene specific primers (Table 1) and a fluorescent probe (TaqmanTM). As observed in breast cancer tissues, complementation was also seen in the detection of circulating tumor cells derived from breast cancers. Again, mammaglobin PCR detected circulating tumor cells in a high percentage of blood

15 samples, albeit at low levels, from metastatic breast cancer patients (20/32) when compared to the normal blood samples (Table 5) but several of the other genes tested to date further increased this detection rate. This included B726P, B305D, B311D, B533S and GABA π . The detection level of mammaglobin in blood samples from metastatic breast cancer patients is higher than described previously (62 vs. 49%), despite testing

20 smaller blood volumes, probably because of the use of epithelial marker-specific enrichment in our study. A combination of all the genes tested indicate that 27/32 samples were positive by one or more of these genes.

WO 01/75171

PCT/US01/10631

50

Table 5
Gene Complementation in Epithelial Cells Isolated from Blood of Normal Individuals
and Metastatic Breast Cancer Patients

Sample ID	Mammaglobin	B305D	B311D	B533S	B726P	GABA π	Combo
2	+	-	-	+	-	-	+
3	+	-	-	+	-	-	+
5	+	+	+	-	+	-	+
6	+	-	+	+	+	-	+
8	-	-	+	-	-	-	+
9	+	+	+	-	+	-	+
10	+	-	+	-	+	-	+
11	-	-	-	-	-	-	-
12	+	+	+	-	-	-	+
13	-	-	-	+	-	-	+
15	-	-	-	-	-	-	-
18	+	-	-	-	-	-	+
19	+	-	-	-	-	+	+
21	+	-	-	-	-	-	+
22	-	-	-	-	-	-	-
23	+	-	-	-	-	-	+
24	+	-	-	-	-	-	+
25	-	+	-	-	-	-	+
26	-	-	-	-	-	-	-
29	+	-	+	+	+	-	+
31	+	-	-	+	-	-	+
32	-	-	-	-	-	±	±
33	-	-	-	-	+	-	+
34	+	-	-	+	-	+	+
35	+	-	-	-	+	-	+
36	-	-	-	-	-	+	+
37	+	-	-	+	-	-	+
38	-	-	-	-	-	-	-
40	+	-	-	-	-	-	+
41	+	-	-	+	-	-	+
42	+	-	-	-	-	-	+
43	-	-	-	-	-	+	+
Donor 104	-	-	-	-	-	+	+
Donor 348	-	-	-	-	-	Nd	-
Donor 392	-	-	-	-	-	Nd	-
Donor 408	-	-	-	-	-	Nd	-
Donor 244	-	-	-	-	-	-	-
Donor 355	-	-	-	-	-	-	-
Donor 264	-	-	-	-	-	-	-
Donor 232	-	-	-	-	-	Nd	-
Donor 12	-	-	-	-	-	-	-
Donor 415	-	-	-	-	-	Nd	-
Donor 35	-	-	-	-	-	-	-
Sensitivity	20/32	4/32	7/32	9/32	7/32	4/32	27/32

5 In further studies, mammaglobin, GABA π , B305D (C form) and B726P specific primers and specific Taqman probes were employed in different combinations

WO 01/75171

PCT/US01/10631

51

to analyze their combined mRNA expression profile in breast metastases (B. met) and breast tumor (B. tumors) samples using real-time PCR. The forward and reverse primers and probes employed for mammaglobin, B305D (C form) and B726P are shown in Table 1. The forward primer and probe employed for GABA π are shown in Table 1, with the reverse primer being as follows: TTCAAATATAAGTGAAGAAAAATTAG-TAGATCAA (SEQ ID NO:51). As shown below in Table 6, a combination of mammaglobin, GABA π , B305D (C form) and B726P was found to detect 22/22 breast tumor samples, with an increase in expression being seen in 5 samples (indicated by ++).

10

WO 01/75171

PCT/US01/10631

52

Table 6
Real-time PCR Detection of Tumor Samples using Different Primer Combinations

Tumor sample	Mammaglobin	Mammaglobin + GABA	Mammaglobin + GABA + B305D	Mammaglobin + GABA + B305D + B726P
B. Met 316A		+	+	+
B. Met 317A	+	+	+	+
B. Met 318A		+	+	++
B. Met 595A	+	+	+	+
B. Met 611A	+	+	+	+
B. Met 612A	+	+	+	+
B. Met 614A		+	+	+
B. Met 616A		+	+	+
B. Met 618A	+	+	+	+
B. Met 620A	+	+	+	+
B. Met 621A	+	+	+	+
B. Met 624A	+	+	+	+
B. Met 625A			+	+
B. Met 627A	+		+	+
B. Met 629A		+	+	+
B. Met 631A	+	+	+	+
B. Tumor 154A	+	+	+	++
B. Tumor 155A	+	+	+	++
B. Tumor 81D			+	++
B. Tumor 209A		+	+	+
B. Tumor 208A		+	+	++
B. Tumor 10A	+	+	+	+

The increase of message signals by the addition of specific primers was further demonstrated in a one plate experiment employing the four tumor samples B. met 316A, B. met 317A, B. tumor 81D and B. tumor 209A.

Expression of a combination of mammaglobin, GABA π , B305D (C form) and B726P in a panel of breast tumor and normal tissue samples was also detected using real-time PCR with a SYBR Green detection system instead of the Taqman probe approach. The results obtained using this system are shown in Figure 7.

EXAMPLE 7

QUANTITATIVE PCR IN PERIPHERAL BLOOD OF BREAST CANCER PATIENTS

The known genes evaluated in this study were mammaglobin and γ aminobutyrate type A receptor π subunit (GABA π). In order to identify novel genes which are over-expressed in breast cancer we have used an improved version of the differential display RT-PCR (DDPCR) technique (Liang et al., Science 257:967-971 (1993); Mou et al., Biochem Biophys Res Commun. 199:564-569 (1994)); cDNA library

WO 01/75171

PCT/US01/10631

53

extraction methods (Hara et al., *Blood* 84:189-199 (1994)) and PCR subtraction (Diatchenko et al., *Proc Natl Acad Sci U S A.*, 93:6025-6030 (1996); Yang et al., *Nucleic Acids Res.* 27:1517-23 (1999)).

Differential display resulted in the recovery of two cDNA fragments designated as B305D and B311D (Houghton et al., *Cancer Res.* 40:Abstract #217, 32-33, (1999). B511S and B533S are two cDNA fragments isolated using cDNA library subtraction approach (manuscript in preparation) while the B726P cDNA fragment was derived from PCR subtraction (Jiang et al., *Proceedings of the Amer Assoc Cancer Res.* 40:Abstract #216, 32 (1999); Xu et al., *Proceedings of the Amer Assoc Cancer Res.* 40:Abstract #2115, 319 (1999); and Molesh et al., *Proceedings of Amer Assoc Cancer Res.* 41:Abstract #4330, 681 (2000).

Three of the novel genes, B311D, B533S and B726P, showed identical breast tissue expression profile by quantitative PCR analysis. These genes were evaluated in quantitative PCR on two different panels consisting of (a) breast tumor, breast normal and normal tissues and (b) panel of breast tumor metastases (primarily lymph nodes). Primers and probes used are shown in Table 1. The data for panel (a) is shown in Figure 2 for all three genes. Overall, the expression profiles are comparable and are in the same rank order, however, the levels of expression are considerably different. B311D in general was expressed at lower levels than B533S and both less than B726P but all three were restricted to breast tissue. All three sequences were used to search against the Genbank database. Both B311D and B533S sequences contain different repetitive sequences and an ORF has not been identified for either. B726P is a novel gene, with mRNA splicing yielding several different putative ORF's.

The quantitative PCR confirmed there was a differential mRNA expression between normal breast tissue and breast tumors, with approximately 50% of breast tumors overexpressed these genes. When tested on a panel of distant metastases derived from breast cancers all three genes reacted with 14/21 metastases and presented similar profiles (data not shown). Interestingly, when tested on a prostate cancer panel, all three genes identified the same 3/24 prostate tumors but at much lower expression

WO 01/75171

PCT/US01/10631

54

levels than in breast. This group of genes exhibited increasing levels of expression as a function of pathological stage of the tumor as shown for B533S.

More rigorous analysis of mammaglobin gene expression by microarray, and quantitative PCR showed expression at significant levels in skin and salivary gland and much lower levels in esophagus and trachea. B511S had a slightly different profile of reactivity on breast tumors and normal breast tissue when compared to mammaglobin, yet reacted with a similar subset of normal tissues as mammaglobin. Mammaglobin detected 14/24 of distant metastatic breast tumors, 31/42 breast tumors and exhibited ten-fold over-expression in tumors and metastases as compared to normal breast tissue. There was at least 300-fold over-expression of mammaglobin in normal breast tissue versus other negative normal tissues and tumors tested. B511S detected 33/42 breast tumors and 14/24 distant metastases. A combination of B511S with mammaglobin would be predicted to detect 38/42 breast tumors and 17/24 metastatic lesions. The quantitative level of expression of B511S and mammaglobin were also in similar ranges, in concordance with the microarray profiles observed for these two genes.

Certain genes complemented mammaglobin's expression profile, *i.e.* were shown to express in tumors that mammaglobin did not. B305D was highly over-expressed in breast tumors, prostate tumors, normal prostate tissue and testis compared to normal tissues including normal breast tissue. Different splice variants of B305D were identified with the forms A and C being the most abundant. All forms tested had similar tissue profiles in quantitative PCR. The A and C forms contain ORF's of 320 and 385 aa, respectively. A known gene shown to be complementary with B305D, in breast tumors, was GABA π . This tissue expression profile is in contrast to other GABA π receptors that typically have appreciable expression in neuronal tissues. An additional observation was that tissue expression profiling of this gene showed it to be over-expressed in breast tumors in an inverse relationship to the B305D gene (Table 3). GABA π detected 15/25 tumors and 6/21 metastases including 4 tumors and 5 metastases missed by mammaglobin. In contrast, B305D detected 13/25 breast tumors and 8/21 metastases again including 3 tumors and 2 metastases missed by mammaglobin. A combination of just B305D and the GABA π would be predicted to

WO 01/75171

PCT/US01/10631

55

identify 22/25 breast tumors and 14/21 metastases. This combination detected 20/21 of the breast metastases as well as 25/25 breast tumors that were evaluated on the same panels for all three genes. The one breast metastasis that was negative for these three genes was strongly positive for B726P.

5 The use of microarray analysis followed by quantitative PCR provided a methodology to accurately determine the expression of breast cancer genes both in breast tissues (tumor and normal) as well as in normal tissues and to assess their diagnostic and therapeutic potential. Five novel genes and two known genes were evaluated using these techniques. Three of these genes B311D, B533S and B726P
10 exhibited concordant mRNA expression and collectively the data is consistent with coordinated expression of these three loci at the level of transcription control. All three genes showed differential expression in breast tumors versus normal breast tissue and the level of overexpression appeared related to the pathological stage of the tumor. In the case of mammaglobin, expression was found in other tissues apart from breast
15 tissue. Expression was seen in skin, salivary gland and to a much lesser degree in trachea.

 Expression of GABA π in breast tumors was also a novel observation. While the expression of several genes complemented that seen with mammaglobin, two genes in particular, B305D and GABA π added to the diagnostic sensitivity of
20 mammaglobin detection. A combination of these three genes detected 45/46 (97.8%) breast tumors and metastases evaluated. Inclusion of B726P enabled the detection of all 25 of the breast tumors and 21 distant metastases.

EXAMPLE 8

25 ENRICHMENT OF CIRCULATING BREAST CANCER CELLS BY IMMUNOCAPTURE

 This example discloses the enhanced sensitivity achieved by use of the immunocapture cell capture methodology for enrichment of circulating breast cancer cells.

 To evaluate the presence of circulating tumor cells an immunocapture
30 method was adopted to first enrich for epithelial cells prior to RT-PCR analysis. Epithelial cells were enriched from blood samples with an immunomagnetic bead

WO 01/75171

PCT/US01/10631

56

separation method (DynaL A.S, Oslo, Norway) utilizing magnetic beads coated with monoclonal antibodies specific for glycopolypeptide antigens on the surface of human epithelial cells. (Exemplary suitable cell-surface antigens are described, for example, in Momburg, F. et al., *Cancer Res.*, 41:2883-91 (1997); Naume, B. et al., *Journal of Hemotherapy*, 6:103-113 (1997); Naume, B. et al., *Int J Cancer*, 78:556-60 (1998); Martin, V.M. et al., *Exp Hematol.*, 26:252-64 (1998); Hildebrandt, M. et al., *Exp Hematol.*, 25:57-65 (1997); Eaton, M.C. et al., *Biotechniques* 22:100-5 (1997); Brandt, B. et al., *Clin Exp Metastases* 14:399-408 (1996), each of which are incorporated herein by reference in their entirety. Cells isolated this way were lysed and the magnetic beads removed. The lysate was then processed for poly A⁺ mRNA isolation using magnetic beads (Dynabeads) coated with Oligo (dT)₂₅. After washing the beads in the kit buffer bead/polyA⁺RNA samples were finally suspended in 10mM Tris HCl pH 8 and subjected to reverse transcription. The RNA was then subjected to Real time PCR using gene specific primers and probes with reaction conditions as outlined herein above. β -Actin content was also determined and used for normalization. Samples with gene of interest copies/ng β -actin greater than the mean of the normal samples + 3 standard deviations were considered positive. Real time PCR on blood samples was performed exclusively using the TaqmanTM procedure but extending to 50 cycles.

Mammaglobin mRNA using enrichment procedures was found to be detectable at much lower levels than when direct isolation was used. Whole blood samples from patients with metastatic breast cancer were subsequently treated with the immunomagnetic beads, polyA⁺ RNA was then isolated, cDNA made and run in quantitative PCR using two gene specific primers to mammaglobin and a fluorescent probe (TaqmanTM). As observed in breast cancer tissues, complementation was also seen in the detection of circulating tumor cells derived from breast cancers. Again, mammaglobin PCR detected circulating tumor cells in a high percentage of bloods, albeit at low levels, from metastatic breast cancer (20/32) when compared to the normal blood samples. Several of the other genes tested to date could further increase this detection rate; this includes B726P, B305D, B311D, B533S and GABA π . A combination of all the genes tested indicates that 27/32 samples were positive by one or more of these genes.

WO 01/75171

PCT/US01/10631

57

EXAMPLE 9

MULTIPLEX DETECTION OF BREAST TUMORS

Additional Multiplex Real-time PCR assays were established in order to simultaneously detect the expression of four breast cancer-specific genes: LipophilinB, Gaba (B899P), B305D-C and B726P. In contrast to detection approaches relying on expression analysis of single breast cancer-specific genes, this Multiplex assay was able to detect all breast tumor samples tested.

This Multiplex assay was designed to detect LipophilinB expression instead of Mammaglobin. Due to their similar expression profiles, LipophilinB can replace Mammaglobin in this Multiplex PCR assay for breast cancer detection. The assay was carried out as follows: LipophilinB, B899P (Gaba), B305D, and B726P specific primers, and specific Taqman probes, were used to analyze their combined mRNA expression profile in breast tumors. The primers and probes are shown below:

15 LipophilinB: Forward Primer (SEQ ID NO: 33): 5'
TGCCCTCCGGAAGCT. Reverse Primer (SEQ ID NO:34): 5'
CGTTTCTGAAGGGACATCTGATC. Probe (SEQ ID NO: 35) (FAM-5' - 3'-
TAMRA): TTGCAGCCAAGTTAGGAGTGAAGAGATGCA.

GABA (B899P): Forward Primer (SEQ ID NO: 36): 5'
20 AAGCCTCAGAGTCCTTCCAGTATG. Reverse Primer (SEQ ID NO: 37): 5'
TTCAAATATAAGTGAAGAAAAATTAGTAGATCAA. Probe (SEQ ID NO: 38)
(FAM-5' - 3'-TAMRA):
AATCCATTGTATCTTAGAACCGAGGGATTGTTTAGA.

B305D (C form): Forward Primer (SEQ ID NO: 39): 5'
25 AAAGCAGATGGTGGTTGAGGTT. Reverse Primer (SEQ ID NO: 40): 5'
CCTGAGACCAAAATGGCTTCTC. Probe (SEQ ID NO: 41) (FAM-5' - 3'-TAMRA)
ATTCCATGCCGGCTGCTTCTCTG.

B726P: Forward Primer (SEQ ID NO: 42): 5'
TCTGGTTTCTCAITCTTTATTCATTTATT. Reverse Primer (SEQ ID NO: 43): 5'
30 TGCCAAGGAGCGGATTATCT. Probe (SEQ ID NO: 44) (FAM-5' - 3'-TAMRA):
CAACCACGTGACAAACACTGGAATTACAGG.

WO 01/75171

PCT/US01/10631

58

Actin: Forward Primer (SEQ ID NO: 45): 5'
 ACTGGAACGGTGAAGGTGACA. Reverse Primer (SEQ ID NO 46): 5'
 CGGCCACATTGTGAACTTTG. Probe (SEQ ID NO: 47): (FAM-5' – 3'-TAMRA):
 CAGTCGGTTGGAGCGAGCATCCC.

5 The assay conditions were:

Taqman protocol (7700 Perkin Elmer):

In 25 µl final volume: 1x Buffer A, 5mM MgCl, 0.2 mM dCTP, 0.2 mM
 dATP, 0.4 mM dUTP, 0.2 mM dGTP, 0.01 U/µl AmpErase UNG, 0.025 u/µl TaqGold,
 8% (v/v) Glycerol, 0.05% (v/v) Gelatin, 0.01% (v/v) Tween20, 4 pmol of each gene
 10 specific Taqman probe (LipophilinB + Gaba + B305D + B726P), 100 nM of B726P-F +
 B726P-R, 300 nM of Gaba-R, and 50 nM of LipophilinB-F + LipophilinB-R + B305D-
 R + Gaba-R, template cDNA (originating from 0.02 µg polyA + RNA).

LipophilinB expression was detected in 14 out of 27 breast tumor
 samples.

15 However, the Multiplex assay for LipophilinB, B899P, B305D-C and B726P detected
 an expression signal in 27 out of 27 tumors with the detection level above 10 mRNA
 copies/1000 pg actin in the majority of samples and above 100 mRNA copies/1000 pg
 actin in 5 out of the 27 samples tested (Figure 8).

20

EXAMPLE 10

MULTIPLEX DETECTION OPTIMIZATION

The Multiplex Real-time PCR assay described above was used to detect
 the expression of Mammaglobin (or LipophilinB), Gaba (B899P), B305D-C and B726P
 simultaneously. According to this Example, assay conditions and primer sequences
 25 were optimized to achieve parallel amplification of four PCR products with different
 lengths. Positive samples of this assay can be further characterized by gel
 electrophoresis and the expressed gene(s) of interest can be determined according to the
 detected amplicon size(s).

Mammaglobin (or LipophilinB), Gaba (B899P), B305D and B726P
 30 specific primers and specific Taqman probes were used to simultaneously detect their
 expression. The primers and probes used in this example are shown below.

WO 01/75171

PCT/US01/10631

59

- Mammaglobin: Forward Primer (SEQ ID NO: 48): 5' TGCCATAGATGAATTGAAGGAATG. Reverse Primer (SEQ ID NO: 49): 5' TGTCATATATTAATTGCATAAACACCTCA. Probe (SEQ ID NO: 50): (FAM-5' – 3'-TAMRA): TCTTAACCAAACGGATGAACTCTGAGCAATG.
- 5 GABA (B899P): Forward Primer (SEQ ID NO: 36): 5' AAGCCTCAGAGTCCTCCAGTATG. Reverse Primer (SEQ ID NO: 51): 5' ATCATTGAAAATTCAAATATAAGTGAAG. Probe (SEQ ID NO: 38) (FAM-5' – 3'-TAMRA) AATCCATTGTATCTTAGAACCGAGGGATTGTGTTAGA.
- B305D (C form): Forward Primer (SEQ ID NO: 39): 5' AAAGCAGATGGTGGTTGAGGTT. Reverse Primer (SEQ ID NO: 40): 5' CCTGAGACCAAATGGCTTCTTC. Probe (SEQ ID NO: 41): (FAM-5' – 3'-TAMRA): ATTCCATGCCGGCTGCTTCTCTG.
- B726P: Forward Primer (SEQ ID NO: 52): 5' GTAGTTGTGCATTGAAATAATTATCATTAT. Reverse Primer (SEQ ID NO: 43): 5' TGCCAAGGAGCGGATTATCT. Probe (SEQ ID NO: 44) (FAM-5' – 3'-TAMRA): CAACCACGTGACAAACACTGGAATTACAGG.

Primer locations and assay conditions were optimized to achieve parallel amplification of four PCR products with different sizes. The assay conditions were:

Taqman protocol (7700 Perkin Elmer):

- 20 In 25 µl final volume: 1x Buffer A, 5 mM MgCl, 0.2 mM dCTP, 0.2 mM dATP, 0.4 mM dUTP, 0.2 mM dGTP, 0.01 U/µl AmpErase UNG, 0.0375 U/µl TaqGold, 8% (v/v) Glycerol, 0.05% (v/v) Gelatin, 0.01% (v/v) Tween20, 4 pmol of each gene specific Taqman probe (Mammaglobin + Gaba + B305D + B726P), 300 nM of Gaba-R + Gaba-F, 100 nM of Mammaglobin-F + R; B726P-F + R, and 50 nM of
- 25 B305D-F + R template cDNA (originating from 0.02 (µg polyA + RNA).

PCR protocol:

50° for 2': x 1, 95° for 10': X 1, and 95° for 15'' / 60° for 1' / 68° for 1': x 50.

- Since each primer set in the multiplex assay results in a band of unique
- 30 length, expression signals of the four genes of interest can be measured individually by agarose gel analysis (*see*, Figure 9), or the combined expression signal of all four genes

WO 01/75171

PCT/US01/10631

60

can be measured in real-time on an ABI 7700 Prism sequence detection system (PE Biosystems, Foster City, CA). The expression of LipophilinB can also be detected instead of Mammaglobin. Although specific primers have been described herein, different primer sequences, different primer or probe labeling and different detection systems could be used to perform this multiplex assay. For example, a second fluorogenic reporter dye could be incorporated for parallel detection of a reference gene by real-time PCR. Or, for example a SYBR Green detection system could be used instead of the Taqman probe approach.

10

EXAMPLE 11

DESIGN AND USE OF GENOMIC DNA-EXCLUDING, INTRON-EXON BORDER SPANNING
PRIMER PAIRS FOR BREAST CANCER MULTIPLEX ASSAY

The Multiplex Real-time PCR assay described herein can detect the expression of Mammaglobin, Gaba (B899P), B305D-C and B726P simultaneously.

15 The combined expression levels of these genes is measured in real-time on an ABI 7700 Prism sequence detection system (PE Biosystems, Foster City, CA). Individually expressed genes can also be identified due to different amplicon sizes via gel electrophoresis. In order to use this assay with samples derived from non-DNase treated RNAs (e.g. lymph node cDNA) and to avoid DNase-treatment for small RNA-samples (e.g. from blood specimens, tumor and lymph node aspirates), intron-spanning primer pairs have been designed to exclude the amplification of genomic DNA and therefore to eliminate nonspecific and false positive signals. False positive signal is caused by genomic DNA contamination in cDNA specimens. The optimized Multiplex assay described herein excludes the amplification of genomic DNA and allows specific

25 detection of target gene expression without the necessity of prior DNase treatment of RNA samples. Moreover the genomic match and the location of the Intron-Exon border could be verified with these primer sets.

Mammaglobin, Gaba (B899P), B305D and B726P specific primers and specific Taqman probes were used to simultaneously detect their expression (Table 7).

30 Primer locations were optimized (Intron-Exon border spanning) to exclusively detect

WO 01/75171

PCT/US01/10631

61

cDNA and to exclude genomic DNA from amplification. The identity of the expressed gene(s) was determined by gel electrophoresis.

5 Table 7
Intron-Exon border Spanning Primer and Probe Sequences
for Breast Tumor Multiplex Assay

Gene	Forward Primer	Reverse Primer	Taqman probe (FAM-5' - 3'TAMRA)
Mammaglobin	tgccatagatgaattgaagga atg (SEQ ID NO:48)	tgctatattatgaataaacct ca (SEQ ID NO:49)	tcttaaccaacggatgaactctgagca atg (SEQ ID NO:50)
B899P	aagcctcagagtcctccagta tg (SEQ ID NO:36)	ttcaatataagtgagaaaaanfa gtatgataa (SEQ ID NO:37)	aatccatgtatcttagaaccgagagatt gttt (SEQ ID NO:62)
B305D	aaagcagatgggttgagg t (SEQ ID NO:39)	cctgagaccatggctcttc (SEQ ID NO:40)	attccatgccggctgtctcttg (SEQ ID NO:41)
B726P	tctggttttctctcttattcatt tatt (SEQ ID NO:42)	tgccnaggagcggattatct (SEQ ID NO:43)	caaccacgtgacaaactggaattaca gg (SEQ ID NO:44)
Actin	actggacgggtgaaggtgac a (SEQ ID NO:45)	cggccacattgtgaacttg (SEQ ID NO:46)	cagtcggttagagcagcatccc (SEQ ID NO:47)
B899P-INT	caattttgtggagaaccg (SEQ ID NO:53)	gctgctggaggtatagtg (SEQ ID NO:54)	catctcagaggtatcatggactacaca (SEQ ID NO:55)
B305D-INT	tctgataaagccgtacaa (SEQ ID NO:56)	tcacgacttgcgttttggctc (SEQ ID NO:57)	atcaaaaaaacatggcctcacacca ct (SEQ ID NO:58)
B726P-INT	gcaagtgccatgatcagagg (SEQ ID NO:59)	atatagactcaggtatacacact (SEQ ID NO:60)	tcctcatcagaatccaacagaggaga tg (SEQ ID NO:61)

Primer locations and assay conditions were optimized to achieve parallel amplification of the four PCR products. The assay conditions were as follows:

10 Taqman protocol (7700 Perkin Elmer)

In 25 μ l final volume: 1x Buffer A, 5 mM MgCl₂, 0.2 mM dCTP, 0.2 mM dATP, 0.4 mM dUTP, 0.2 mM dGTP, 0.01 U/AmpErase UNG, 8 % (v/v) Glycerol, 0.05 % (v/v) Gelatin, 0.01 % (v/v) Tween20, 4 pmol of each gene specific Taqman probe (Mammaglobin + B899P-INT + B305D-INT + B726P-INT), 300 nM of B305D-INT-F; B899P-INT-F, 100 nM of Mammaglobin-F + R; B726P-INT-F +R, 50 nM of B899P-INT-R; B305D-INT-R, template cDNA (originating from 0.02 μ g polyA+ RNA).

PCR cycling conditions

20 1 cycle at 50°C for 2 minutes, 1 cycle at 95°C for 10 minutes, 50 cycles of 95°C for 1 minute and 68°C for 1 minute.

Figure 10 shows a comparison of the multiplex assay using intron-exon border spanning primers (bottom panel) and the multiplex assay using non-optimized

WO 01/75171

PCT/US01/10631

62

primers (top panel), to detect breast cancer cells in a panel of lymph node tissues. This experiment shows that reduction in background resulting from genomic DNA contamination in samples is achieved using the intron-exon spanning primers of the present invention.

5

EXAMPLE 12

MULTIPLEX DETECTION OF METASTASIZED BREAST TUMOR CELLS IN
SENTINEL LYMPH NODE BIOPSY SAMPLES

Lymph node staging is important for determining appropriate adjuvant hormone and chemotherapy. In contrast to conventional axillary dissection a less invasive approach for staging of minimal residual disease is sentinel lymph node biopsy. Sentinel lymph node biopsy (SLNB) has the potential to improve detection of metastases and to provide prognostic values to lead to therapy with minimal morbidity associated with complete lymph node dissection. SLNB implements mapping of the one or two lymph nodes which primarily drain the tumor and therefore are most likely to harbor metastatic disease (the sentinel nodes). Routine pathological analysis of lymph nodes result in a high false-negative rate: one-third of women with pathologically negative lymph nodes develop recurrent disease [Bland: The Breast: Saunders 1991]. A more sensitive detection technique for tumor cells would be RT-PCR but its application is limited by lack of a single specific markers. The multimarker assay described above increases the likelihood of cancer detection across the population without producing false-positive results from normal lymph nodes.

As mentioned above, lymphatic afferents from a primary tumor drain into a single node, the sentinel lymph node, before drainage into the regional lymphatic basin occurs. Sentinel lymph nodes are located with dyes and/or radiolabelled colloid injected in the primary lesion site and sentinel lymph node biopsy allows pathological examination for micrometastatic deposits, staging of the axilla and therefore can avoid unnecessary axillary dissection. Nodal micrometastases can be located with staining (haematoxylin or eosin) or immunohistochemical analysis for cytokeratin proteins. Immunocytochemical staining techniques can produce frequent false-negative results by missing small metastatic foci due to inadequate sectioning of the node.

WO 01/75171

PCT/US01/10631

63

Immunohistochemistry can result in false-positive results due to illegitimate expression of cytokeratins (reticulum cells) or in false-negative results when using the antibody Ber-Ep4 which corresponding antigen is not expressed on all tumor cells.

The multiplex assay described herein could provide a more sensitive
5 detection tool for positive sentinel lymph nodes. Moreover the detection of breast cancer cells in bone marrow samples, peripheral blood and small needle aspiration samples is desirable for diagnosis and prognosis in breast cancer patients.

Twenty-two metastatic lymph node samples, in addition to 15 samples
also previously analyzed and shown in Figure 3A, were analyzed using the intron-exon
10 border spanning multiplex PCR assay described herein. The results from this analysis are summarized in Table 8. Twenty-seven primary tumors were also analyzed and the results shown in Table 9. Twenty normal lymph node samples tested using this assay were all negative.

15

WO 01/75171

PCT/US01/10631

64

Table 8.
Multiplex Real-time PCR Analysis of 37 Metastatic Lymph Nodes

breast metastatic lymph node samples	Mammaglobin	B305D	B899P	B726P	Multiplex
B.Met 317A	++	+		+	+++
B.Met 318A			++		+++
B.Met 595A	+			+	+++
B.Met 611A	+	+	+++		++
B.Met 612A	++	++		+	++
B.Met 614A		++		++	+++
B.Met 616A			+		++
B.Met 618A	+++	+			+++
B.Met 620A	++	++		++	+++
B.Met 621A	+	+++		+	+++
B.Met 624A			++		+++
B.Met 625A		++		++	+
B.Met 627A		+		+	+
B.Met 629A	++				+++
B.Met 631A	+		++		+
1255	+++	++		++	++
1257	+++	+	+	+	++
769	+++			+	++
1258	++	+	+		+
1259		++	++		+++
1250	+++	+		+	+++
1726	+++	+		+	+++
786	+++	+	+		+++
281-L1-r	+++				+++
289-L2	++	+			++
366-S	+				+
374-S+	+++	++			+++
376-S	++			+	++
381-S	+	+			+
383-Sx	+++	++			+++
496-M	+++	++			+++
591-S1-A	+	+			+
652-I		+	++		+++
772	+				+
777	+	+		++	++
778	+++				+++
779	+		++		++

Table 9
Multiplex Real-time PCR Analysis of 27 Primary Breast Tumors

breast primary tumor samples	Mammaglobin	B30SD	B899P	B726P	Multiplex
T443	+	++		+++	+++
T457		+	+		++
T395		++			++
T10A	+	+++		+++	+++
T446		+		++	++
T11C	+		+++		+++
T23B	+	++			+++
T207A		++			+
T437	+	+		++	+++
T391	+	++		+++	+++
T392	+	+			++
T876	+	++			+++
T483	++	+			+++
T81G	+	+	++	++	+++
T430	+		++		++
T465	+	+		+	++
T880			+		+
T469	+			+	+++
T467	+			++	+++
T439		+			+
T387	++		+	+	++
T318			+		++
T154A				+	+
T387A	+++		+	+	+++
T155A	+		++	+	+
T209A		++			++
T208A		+		+	++

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 01/75171

PCT/US01/10631

66

CLAIMS

We Claim:

1. A method for identifying one or more tissue-specific polynucleotides, said method comprising the steps of:

(a) performing a genetic subtraction to identify a pool of polynucleotides from a tissue of interest;

(b) performing a DNA microarray analysis to identify a first subset of said pool of polynucleotides of interest wherein each member polynucleotide of said first subset is at least two-fold over-expressed in said tissue of interest as compared to a control tissue; and

(c) performing a quantitative polymerase chain reaction (PCR) analysis on polynucleotides within said first subset to identify a second subset of polynucleotides that are at least two-fold over-expressed as compared to said control tissue;

wherein a polynucleotide is identified as tissue-specific if it is at least two-fold over-expressed by both microarray and quantitative PCR analyses.

2. The method of claim 1 wherein said genetic subtraction is selected from the group consisting of differential display and cDNA subtraction.

3. A method for identifying a subset of polynucleotides showing complementary tissue-specific expression profiles in a tissue of interest, said method comprising the steps of:

(a) performing a first expression analysis selected from the group consisting of DNA microarray and quantitative PCR to identify a first polynucleotide that is at least two-fold over-expressed in a first tissue sample of interest obtained from a first patient but not over-expressed in a second tissue sample of interest as compared to a control tissue; and

WO 01/75171

PCT/US01/10631

67

(b) performing a second expression analysis selected from the group consisting of DNA microarray and quantitative PCR to identify a second polynucleotide that is at least two-fold over-expressed in a second tissue sample of interest obtained from a second patient but not over-expressed in a first tissue sample of interest as compared to said control tissue;

wherein the first tissue sample and said second tissue sample are of the same tissue type, and wherein over-expression of said first polynucleotide in only said first tissue samples of interest and over-expression of said second polynucleotide in only said second tissue sample of interest indicates complementary tissue-specific expression of said first polynucleotide and said second polynucleotide.

4. A method for determining the presence of a cancer cell in a patient, said method comprising the steps of:

(a) obtaining a biological sample from said patient;

(b) contacting said biological sample with a first oligonucleotide that hybridizes to a first polynucleotide said first polynucleotide selected from the group consisting of polynucleotides depicted in SEQ ID NO:73, SEQ ID NO:74 and SEQ ID NO:76;

(c) contacting said biological sample with a second oligonucleotide that hybridizes to a second polynucleotide selected from the group consisting of SEQ ID NO: 1, 3, 5-7, 11, 13, 15, 17, 19-24, 30, 32, and 75;

(d) detecting in said sample an amount of a polynucleotide that hybridizes to at least one of said oligonucleotides; and

(e) comparing the amount of the polynucleotide that hybridizes to said oligonucleotide to a predetermined cut-off value, and therefrom determining the presence or absence of a cancer in the patient.

5. A method for determining the presence or absence of a cancer in a patient, said method comprising the steps of:

(a) obtaining a biological sample from said patient;

WO 01/75171

PCT/US01/10631

68

(b) contacting said biological sample with a first oligonucleotide that hybridizes to a first polynucleotide selected from the group consisting of polynucleotides depicted in SEQ ID NO:73, SEQ ID NO:74 and SEQ ID NO:76;

(c) contacting said biological sample with a second oligonucleotide that hybridizes to a second polynucleotide as depicted in SEQ ID NO:75;

(d) contacting said biological sample with a third oligonucleotide that hybridizes to a third polynucleotide selected from the group consisting of polynucleotides depicted in SEQ ID NO:5, SEQ ID NO:6 and SEQ ID NO:7;

(e) contacting said biological sample with a fourth oligonucleotide that hybridizes to a fourth polynucleotide selected from the group consisting of polynucleotides depicted in SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23 and SEQ ID NO:24;

(f) detecting in said biological sample an amount of a polynucleotide that hybridizes to at least one of said oligonucleotides; and

(g) comparing the amount of polynucleotide that hybridizes to the oligonucleotide to a predetermined cut-off value, and therefrom determining the presence or absence of a cancer in the patient.

6. A method for determining the presence or absence of a cancer in a patient, said method comprising the steps of:

(a) obtaining a biological sample from said patient;

(b) contacting said biological sample with an oligonucleotide that hybridizes to a tissue-specific polynucleotide;

(c) detecting in the sample a level of a polynucleotide that hybridizes to the oligonucleotide; and

(d) 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.

WO 01/75171

PCT/US01/10631

69

7. A method for monitoring the progression of a cancer in a patient, said method comprising the steps of:

- (a) obtaining a first biological sample from said patient;
- (b) contacting said biological sample with an oligonucleotide that hybridizes to a polynucleotide that encodes a breast tumor protein;
- (c) detecting in the sample an amount of said polynucleotide that hybridizes to said oligonucleotide;
- (d) repeating steps (b) and (c) using a second biological sample obtained from said patient at a subsequent point in time; and
- (e) comparing the amount of polynucleotide detected in step (d) with the amount detected in step (c) and therefrom monitoring the progression of the cancer in the patient.

8. The method any one of claim 6 and claim 7 wherein said polynucleotide encodes a breast tumor protein selected from the group consisting of mammaglobin, lipophilin B, GABA π (B899P), B726P, B511S, B533S, B305D and B311D.

9. A method for detecting the presence of a cancer cell in a patient, said method comprising the steps of:

- (a) obtaining a biological sample from said patient;
- (b) contacting said biological sample with a first oligonucleotide that hybridizes to a first polynucleotide selected from the group consisting of mammaglobin and lipophilin B;
- (c) contacting said biological sample with a second oligonucleotide that hybridizes to a second polynucleotide sequence selected from the group consisting of GABA π (B899P), B726P, B511S, B533S, B305D and B311D;
- (d) detecting in said biological sample an amount of a polynucleotide that hybridizes to at least one of the oligonucleotides; and

WO 01/75171

PCT/US01/10631

70

(e) comparing the amount of polynucleotide that hybridizes to the oligonucleotide to a predetermined cut-off value, and therefrom determining the presence or absence of a cancer in the patient.

10. A method for determining the presence of a cancer cell in a patient, said method comprising the steps of:

- (a) obtaining a biological sample from said patient;
- (b) contacting said biological sample with a first oligonucleotide that hybridizes to a first polynucleotide selected from the group consisting of a polynucleotide depicted in SEQ ID NO:73 and SEQ ID NO:74 or complement thereof;
- (c) contacting said biological sample with a second oligonucleotide that hybridizes to a second polynucleotide depicted in SEQ ID NO:75 or complement thereof;
- (d) contacting said biological sample with a third oligonucleotide that hybridizes to a third polynucleotide selected from the group consisting of a polynucleotide depicted in SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:6 and SEQ ID NO:7 or complement thereof;
- (e) contacting said biological sample with a fourth oligonucleotide that hybridizes to a fourth polynucleotide selected from the group consisting of a polynucleotide depicted in SEQ ID NO:11 or complement thereof;
- (f) contacting said biological sample with a fifth oligonucleotide that hybridizes to a fifth polynucleotide selected from the group consisting of a polynucleotide depicted in SEQ ID NO:13, 15 and 17 or complement thereof;
- (g) contacting said biological sample with a sixth oligonucleotide that hybridizes to a sixth polynucleotide selected from the group consisting of a polynucleotide depicted in SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23 and SEQ ID NO:24 or complement thereof;
- (h) contacting said biological sample with a seventh oligonucleotide that hybridizes to a seventh polynucleotide depicted in SEQ ID NO:30 or complement thereof;

WO 01/75171

PCT/US01/10631

71

(i) contacting said biological sample with an eighth oligonucleotide that hybridizes to a polynucleotide depicted in SEQ ID NO:32 or complement thereof;

(j) contacting said biological sample with a ninth oligonucleotide that hybridizes to a polynucleotide depicted in SEQ ID NO:76 or complement thereof;

(k) detecting in said biological sample a hybridized oligonucleotide of any one of steps (b) through (j) and comparing the amount of polynucleotide that hybridizes to the oligonucleotide to a predetermined cut-off value, wherein the presence of a hybridized oligonucleotide in any one of steps (b) through (j) in excess of the pre-determined cut-off value indicates the presence of a cancer cell in the biological sample of said patient.

11. A method for determining the presence of a cancer cell in a patient, said method comprising the steps of:

(a) obtaining a biological sample from said patient;

(b) contacting said biological sample with a first oligonucleotide and a second oligonucleotide;

i. wherein said first oligonucleotide and said second oligonucleotide hybridize to a first polynucleotide and a second polynucleotide, respectively;

ii. wherein said first polynucleotide and said second polynucleotide are selected from the group consisting of polynucleotides depicted in SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:30, SEQ ID NO:32, and SEQ ID NO:76; and

iii. wherein said first polynucleotide is unrelated in nucleotide sequence to said second polynucleotide;

WO 01/75171

PCT/US01/10631

72

(c) detecting in said biological sample said hybridized first oligonucleotide and said hybridized second hybridized oligonucleotide; and

(d) comparing the amount of said hybridized first oligonucleotide and said hybridized second hybridized oligonucleotide to a predetermined cut-off value; wherein an amount of said hybridized first oligonucleotide or said hybridized second oligonucleotide in excess of the pre-determined cut-off value indicates the presence of a cancer cell in the biological sample of said patient.

12. A method for determining the presence or absence of a cancer cell in a patient, said method comprising the steps of:

(a) obtaining a biological sample from said patient;

(b) contacting said biological sample with a first oligonucleotide and

a

second oligonucleotide;

i. wherein said first oligonucleotide and said second oligonucleotide hybridize to a first polynucleotide and a second polynucleotide, respectively;

ii. wherein said first polynucleotide and said second polynucleotide are both tissue-specific polynucleotides of the cancer cell to be detected; and

iii. wherein said first polynucleotide is unrelated in nucleotide sequence to said second polynucleotide;

(c) detecting in said biological sample said first hybridized oligonucleotide and said second hybridized oligonucleotide; and

(d) comparing the amount of polynucleotide that hybridizes to the oligonucleotide to a predetermined cut-off value, wherein the presence of a hybridized first oligonucleotide or a hybridized second oligonucleotide in excess of the pre-determined cut-off value indicates the presence of a cancer cell in the biological sample of said patient.

WO 01/75171

PCT/US01/10631

73

13. A method for detecting the presence of a cancer cell in a patient, said method comprising the steps of:

(a) obtaining a biological sample from said patient;

(b) contacting said biological sample with a first oligonucleotide pair said first pair comprising a first oligonucleotide and a second oligonucleotide wherein said first oligonucleotide and said second oligonucleotide hybridize to a first polynucleotide and the complement thereof, respectively;

(c) contacting said biological sample with a second oligonucleotide pair said second pair comprising a third oligonucleotide and a fourth oligonucleotide wherein said third and said fourth oligonucleotide hybridize to a second polynucleotide and the complement thereof, respectively, and wherein said first polynucleotide is unrelated in nucleotide sequence to said second polynucleotide;

(d) amplifying said first polynucleotide and said second polynucleotide;

and

(e) detecting said amplified first polynucleotide and said amplified second polynucleotide;

wherein the presence of said amplified first polynucleotide or said amplified second polynucleotide indicates the presence of a cancer cell in said patient.

14. The method of any one of claims 4-7 and 9-13 wherein said biological sample is selected from the group consisting of blood, serum, lymph node, bone marrow, sputum, urine and tumor biopsy sample.

15. The method of claim 14 wherein said biological sample is selected from the group consisting of blood, a lymph node and bone marrow.

16. The method of claim 15 wherein said lymph node is a sentinel lymph node.

WO 01/75171

PCT/US01/10631

74

17. The method of any one of claims 4-7 and 9-13 wherein said cancer is selected from the group consisting of prostate cancer, breast cancer, colon cancer, ovarian cancer, lung cancer head & neck cancer, lymphoma, leukemia, melanoma, liver cancer, gastric cancer, kidney cancer, bladder cancer, pancreatic cancer and endometrial cancer.

18. The method of any one of claims 12 and 13 wherein said first polynucleotide and said second polynucleotide are selected from the group consisting of mammaglobin B, GABA π (B899P), B726P, B511S, B533S, B305D and B311D.

19. The method of any one of claims 12 and 13 wherein said first polynucleotide and said second polynucleotide are selected from the group consisting of polynucleotide depicted in SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:30, SEQ ID NO:32, and SEQ ID NO:76.

20. The method of any one of claims 12 and 13 wherein said oligonucleotides are selected from the group consisting of oligonucleotides depicted in SEQ ID NOs:33-35 and 63-72.

21. The method of any one of claims 12 and 13 wherein the step of detection of said first amplified polynucleotide and said second polynucleotide comprises a step selected from the group consisting of detecting a radiolabel and detecting a fluorophore.

22. The method of any one of claims 4-7 and 9-13 wherein said step of detection comprises a step of fractionation.

WO 01/75171

PCT/US01/10631

75

23. The method of any one of claims 12 and 13 wherein said first and said oligonucleotides are intron spanning oligonucleotides.

24. The method of claim 23 wherein said intron spanning oligonucleotides are selected from the group consisting of oligonucleotides depicted in SEQ ID NOs:36-62.

25. The method of claim 13 wherein detection of said amplified first or said second polynucleotide comprises contacting said amplified first or said second polynucleotide with a labeled oligonucleotide probe that hybridizes, under moderately stringent conditions, to said first or said second polynucleotide.

26. The method of claim 13 wherein said labeled oligonucleotide probe comprises a detectable moiety selected from the group consisting of a radiolabel and a fluorophore.

27. The method of any one of claims 4-7 and 9-13 further comprising a step of enriching said cancer cell from said biological sample prior to hybridizing said oligonucleotide primer(s).

28. The method of claim 27 wherein said step of enriching said cancer cell from said biological sample is achieved by a methodology selected from the group consisting of cell capture and cell depletion.

29. The method of claim 28 wherein cell capture is achieved by immunocapture, said immunocapture comprising the steps of:

- (a) adsorbing an antibody to the surface of said cancer cells; and
- (b) separating said antibody adsorbed cancer cells from the remainder of

said biological sample.

WO 01/75171

PCT/US01/10631

76

30. The method of claim 29 wherein said antibody is directed to an antigen selected from the group consisting of CD2, CD3, CD4, CD5, CD8, CD10, CD11b, CD14, CD15, CD16, CD19, CD20, CD24, CD25, CD29, CD33, CD34, CD36, CD38, CD41, CD45, CD45RA, CD45RO, CD56, CD66B, CD66e, HLA-DR, IgE and TCR $\alpha\beta$.

31. The method of claim 29 wherein said antibody is directed to a breast tumor antigen.

32. The method of any one of claims 29-31 wherein said antibody is a monoclonal antibody.

33. The method of claim 29 wherein said antibody is conjugated to magnetic beads.

34. The method of claim 29 wherein said antibody is formulated in a tetrameric antibody complex.

35. The method of claim 28 wherein cell depletion is achieved by a method comprising the steps of:

- (a) cross-linking red cells and white cells, and
- (b) fractionating said cross-linked red and white cells from the remainder of said biological sample.

36. The method of claim 13 wherein said step of amplifying is achieved by a polynucleotide amplification methodology selected from the group consisting of reverse transcription polymerase chain reaction (RT-PCR), inverse PCR, RACE, ligase chain reaction (LCR), Qbeta Replicase, isothermal amplification, strand displacement amplification (SDA), rolling chain reaction (RCR), cyclic probe reaction (CPR), transcription-based amplification systems (TAS), nucleic acid sequence based amplification (NASBA) and 3SR.

WO 01/75171

PCT/US01/10631

77

37. A composition for detecting a cancer cell in a biological sample of a patient, said composition comprising:

- (a) a first oligonucleotide; and
- (b) a second oligonucleotide;

wherein said first oligonucleotide and said second oligonucleotide hybridize to a first polynucleotide and to a second polynucleotide, respectively; wherein said first polynucleotide is unrelated in nucleotide sequence from said second polynucleotide; and wherein said first polynucleotide and said second polynucleotide are tissue-specific polynucleotides of the cancer cell to be detected.

38. The composition of claim 37 wherein said first polynucleotide and said second polynucleotide are complementary tissue-specific polynucleotides of the tissue-type of said cancer cell.

39. The composition of any one of claim 37 and claim 38 wherein said first polynucleotide and said second polynucleotide are selected from the group consisting of the polynucleotides depicted in SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:30, SEQ ID NO:32, and SEQ ID NO:76.

40. The composition of any one of claim 37 and claim 38 wherein said oligonucleotides are selected from the group consisting of oligonucleotides as disclosed in SEQ ID NOs: 33-72.

41. A composition for detecting a cancer cell in a biological sample of a patient, said composition comprising:

- (a) a first oligonucleotide pair; and
- (b) a second oligonucleotide pair;

WO 01/75171

PCT/US01/10631

78

wherein said first oligonucleotide pair and said second oligonucleotide pair hybridize to a first polynucleotide (or complement thereof) and to a second polynucleotide (or complement thereof), respectively; wherein said first polynucleotide is unrelated in nucleotide sequence from said second polynucleotide; and wherein said first polynucleotide and said second polynucleotide are tissue-specific polynucleotides of the cancer cell to be detected.

42. The composition of claim 41 wherein said first polynucleotide and said second polynucleotide are complementary tissue-specific polynucleotides of the tissue-type of said cancer cell.

43. The composition of any one of claim 41 and claim 42 wherein said first polynucleotide and said second polynucleotide are selected from the group consisting of the polynucleotides depicted in SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:30, SEQ ID NO:32, and SEQ ID NO:76.

44. The composition of any one of claim 41 and claim 42 wherein said oligonucleotides are selected from the group consisting of oligonucleotides as disclosed in SEQ ID NOs: 33-72.

45. A composition comprising an oligonucleotide primer or probe of between 15 and 100 nucleotides that comprises an oligonucleotide selected from the group consisting of oligonucleotides depicted in SEQ ID NOs:33-72.

46. The composition of claim 45 comprising an oligonucleotide primer or probe selected from the group consisting of oligonucleotides depicted in SEQ ID NOs:33-72.



WO 01/75171

PCT/US01/10631

2/9

B 533S vs Tumor Stage

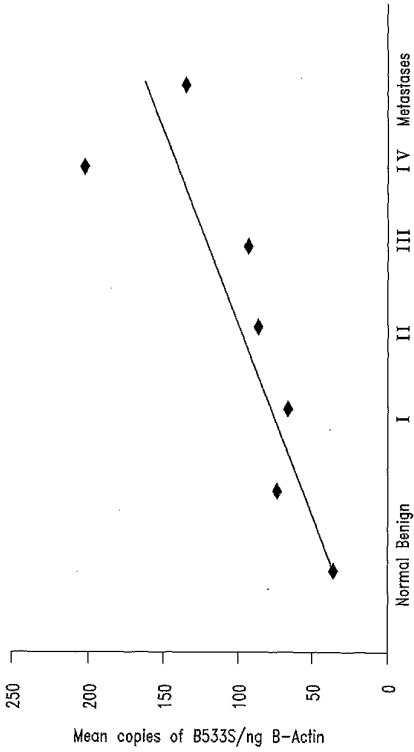


Fig. 2

WO 01/75171

PCT/US01/10631

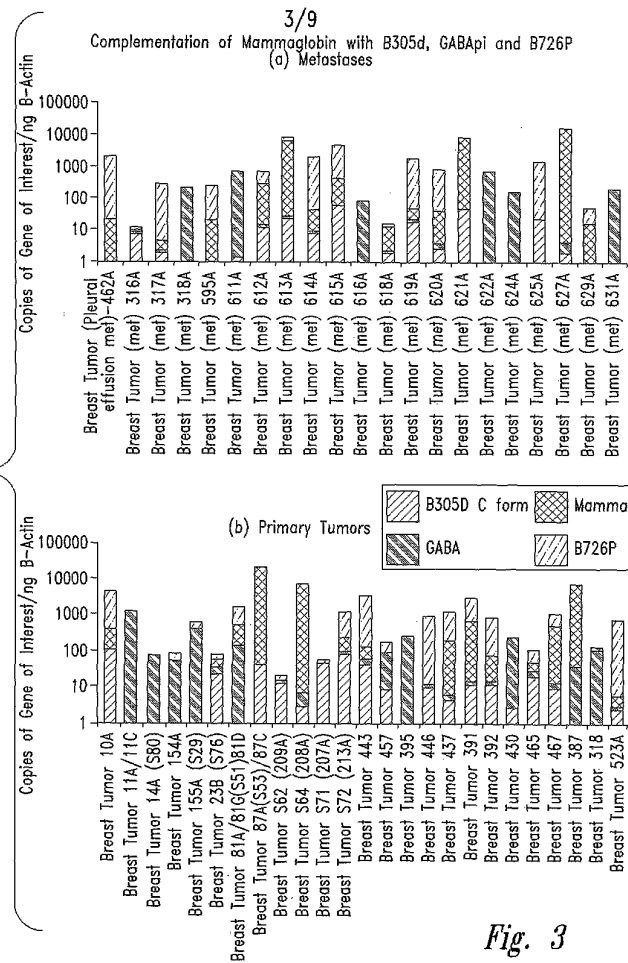


Fig. 3

SUBSTITUTE SHEET (RULE26)

WO 01/75171

PCT/US01/10631

4/9

GACAGGGGTTCTTGATCTTGGCACCGGSACTGAACCGACAGCAG 50
CAGCCTCACCATGAAGTTGCTGATGGTCTCATGCTGGGSGCCCTCTCCC 100
AGCACTGCTACGCAAGGCTCTGGCTGGCCCTTATTGGAGAATGTGATTCC 150
AAGACAATCAATCCACAGTGTCTAAGACTGAATACAAAGAACTTCTTCA 200
AGAGTTTCATAGAGACAATGCCACTACAAATGCCATAGATGAATGAAGG 250
AATGTTTCTTAACCAACGGATGAACCTCTGAGCAATGTTGAGGTGTTT 300
CTGCAATTAATATATGACAGCAGTCTTTGTGATTTATTTAACTTTCTGC 350
AAGACTTTGGCTCACAGAACTGCAGGATGAGTGAAGAACTCACTACGG 400
ATTGCTGCAACCAACACTTCTCTTCTTATGTCCTTTTACTACAACTA 450
CAAGACAATTGTTGAAACCTGCTATACATGTTTATTTAATAAATTGATG 500
GCA 503

Fig. 4

CACCTGCTACGCAAGGCTCTGGCTGCCCTTATTGGAGAATGTGATTTCAA 50
GACAATCAATCCACAAGTGTCTAAGACTGAATACAAAGAACTTCTTCAAG 100
AGTTTCATAGAGACAATGCCACTACAAATGCCATAGATGAATGAAGGAA 150
TGTTTCTTAACCAACGGATGAACCTCTGAGCAATGTTGAGGTGTTTAT 200
GCAATTAATATATGACAGCAGTCTTTGTGATTTATTTGGGGCCATCACC 250
ATCACCATCACTAAGGTCCCGAGCTCGAATTCCTGCAGATATCCATCACAC 300
T 301

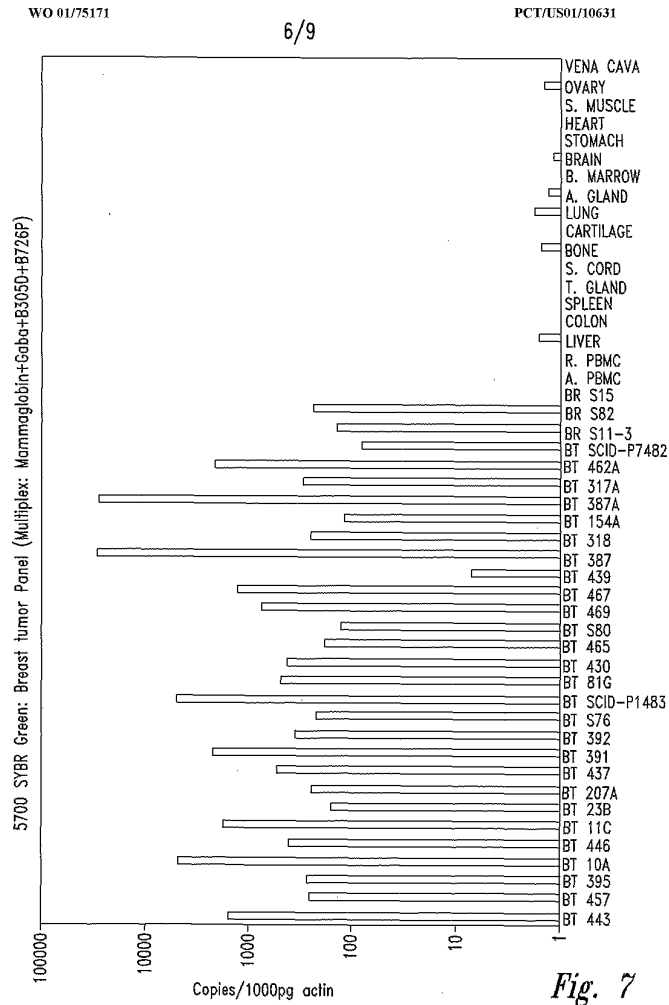
Fig. 5

SUBSTITUTE SHEET (RULE26)

[illegible]

Fig. 6

SUBSTITUTE SHEET (RULE26)



SUBSTITUTE SHEET (RULE26)

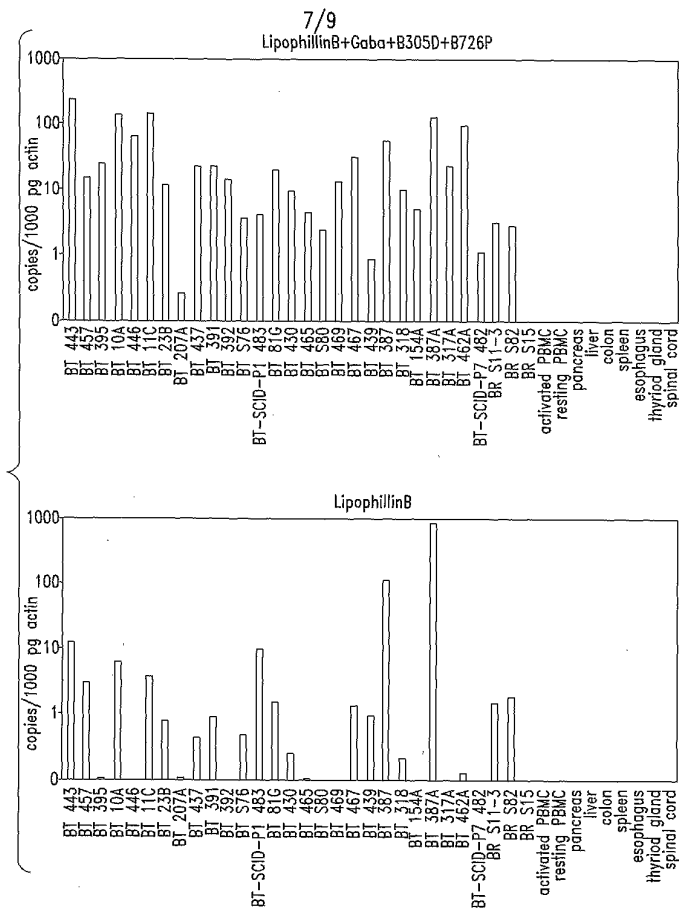


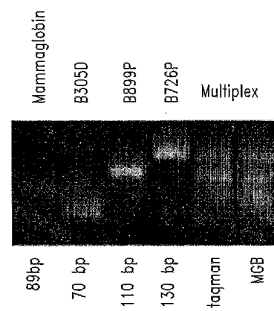
Fig. 8

WO 01/75171

PCT/US01/10631

8/9

Multiplex PCR assay:
Gene determination by amplicon size

*Fig. 9*

SUBSTITUTE SHEET (RULE26)

WO 01/75171

PCT/US01/10631

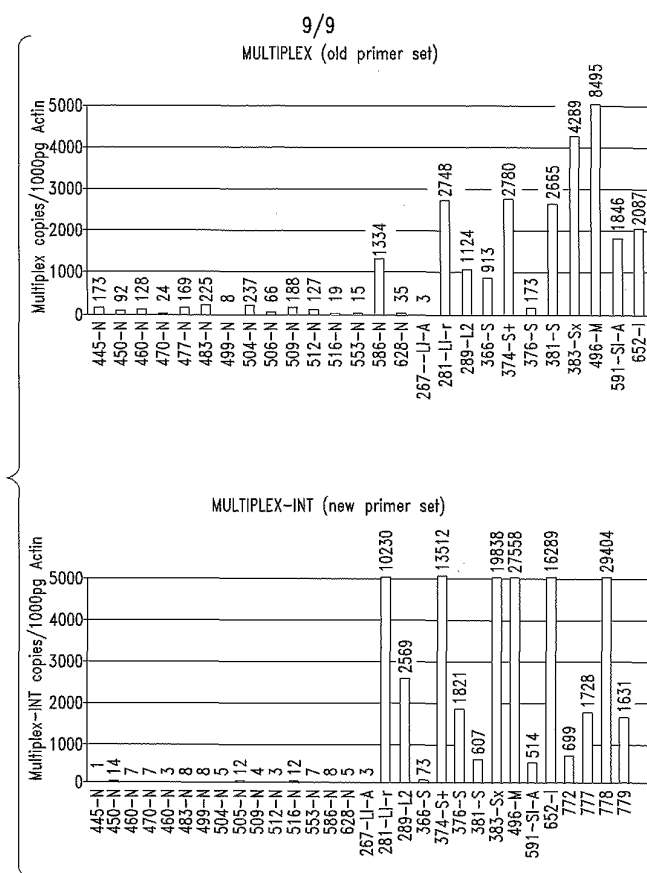


Fig. 10

SUBSTITUTE SHEET (RULE26)

WO 01/75171

PCT/US01/10631

1

SEQUENCE LISTING

<110> Corixa Corporation
Houghton, Raymond L.
Dillon, Davin C.
Molesh, David A.
Xu, Jiangchun
Zehentner, Barbara
Persing, David H.

<120> METHODS, COMPOSITIONS AND KITS FOR THE DETECTION
AND MONITORING OF BREAST CANCER

<130> 210121.513PC

<140> PCT

<141> 2001-04-02

<160> 77

<170> FastSEQ for Windows Version 3.0

<210> 1

<211> 1851

<212> DNA

<213> Homo sapien

<400> 1

tcataccat	tgccagcagc	ggcaccgtta	gtcagggttt	ctgggaatcc	cacatgagta	60
cttcggtgtt	cttcattett	cttcaatagc	cataaatctt	ctagctctgg	ctggctgttt	120
tcacttcttt	taagcctttg	tgactcttcc	tctgaatgca	gctttaagtc	ttgtcttgga	180
ttgctgtttt	cagaagagat	ttttaacatc	tgtttttctt	tgtagtcaga	aagtaactgg	240
caaatcacat	gatgatgact	agaaacagca	tactctctgg	cgtctcttcc	agatcttgag	300
aagatacatc	aacatttttg	tcaagtagag	ggctgaciat	acttgctgat	ccacaacata	360
cagaagtat	gagagcagtt	cttccatc	tatccagcgc	atttaaatc	gctttttct	420
tgattaaaa	tttccacact	tgctgttttt	gtccatgtat	accaagtgc	agtgggtgta	480
ggccatgtct	gttttttgat	tccatctcag	caccgtataa	gagcagtgt	ttggccatta	540
atttatcttc	attgtagaca	gcatagtgta	gagtggtatt	tccatactca	tctggcaaat	600
ttggatcagt	gccatgttcc	agcaacatta	acgcacattc	atcttctctg	cattgtacgg	660
cctttgtcag	agctgtcttc	ttttgttgt	caaggacatt	aagtgtacat	cgtctgtcca	720
gcacgagttt	tactacttct	gaattcccat	tgccagagcc	cagatgtaga	gcagtctctc	780
tttgcttgtc	cctcttggtc	acatccgtgt	ccttgagcat	gacgatgaga	tcctttcttg	840
ggactttacc	ccaccaggca	gctctgtgga	gcttgccag	atcttctcca	tggaactggt	900
acctgggatac	catgaaggcg	ctgtcatcgt	agtctcccca	agcgaccacg	ttgctcttgc	960
cgctccctcg	cagcagggga	agcagtggca	gcaccaattg	cacctcttgc	tcocaaagct	1020
cttcacagag	gagtctgtgt	ggtctccaga	agtgcccaag	ttgctcttgc	cgctccctct	1080
gtccatccag	ggaggaagaa	atgcaggaaa	tgaagagatg	atgcaagatg	gtataactct	1140
cagcatcaaa	acttctggac	agcaggtcac	ttccagcaag	gtggagaag	ctgtccaccc	1200
acagaggatg	agatccagaa	accacaatat	ccattcacaa	acaaacactt	ttcagccaga	1260
cacaggtact	gaatcatgt	catctcgccg	aacatggtgg	aacctaccca	atcacacatc	1320
aagagatgaa	gacactgcag	tatatctgca	caactgaata	ctcttcatcc	ataacaaaat	1380
aataataatt	tcctctggag	ccatatggat	gaactatgaa	ggaagaactc	ccgaagaag	1440
ccagtgcgag	agaagccaca	ctgaagctct	gtctccagcc	atcagcgcca	cggacaggag	1500
tgtgtttctt	ccccagtgat	gcagctccaa	gttatccga	agctgcgcga	gcacacggtg	1560
gtcctcgaga	aacaccccg	ctcttccggt	ctaacacag	caagtcaata	aattgataa	1620

WO 01/75171

PCT/US01/10631

2

```

tcacataaac agaattaasa gcaaaagtcac ataagcatct caacagacac agaaaaggca 1680
ttttgcaaaa tccagcatcc ttgtatttat tgttgacgtt ctcagaggaa atgcttttaa 1740
cttttcccc tttagtatta tgitggctgt gggcttgcca tagtggtttt ttattacttt 1800
aaggatatgc ctttatatgc ctgttttgct gaggggttitta attctatgta c 1851

```

```

<210> 2
<211> 329
<212> PRT
<213> Homo sapien

```

```

<400> 2
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 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 Glu 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> 3
<211> 1852
<212> DNA
<213> Homo sapiens

```

WO 01/75171

PCT/US01/10631

3

```

<400> 3
ggcagcgagaa ttaaaacccct cagcaaaaaca ggcatagaag ggcataacct taaagtaata 60
aaaaccacct atgacaagcc caoagccaac ataatactaa atggggaaaa gttagaagca 120
tttcctctga gaactgcaac aataaataca aggatgctgg attitgtcaa atgcttttc 180
tgtgtctgtt gagatgctta tgtgactttg cttttaattc tgttatgtg attatcacat 240
ttattgactt gactgtgcta gaccggaaga gctgggggtt ttctcaggag caacogtggt 300
ctgoggcagc ttogggataa cttgaggctg catcactggg gaagaaacac aytctgtcc 360
gtggcgctga tggctgagga cagagcttca gtgtggcttc tctcgaactg gcttcttcgg 420
ggagtctctc ctccatagtt catccatagt gctccagagg aaaaattat taitttgtta 480
tggatgaaga gtattacgtt gtgcagatat actgcaaggt cticactctc tgaatgttga 540
ttgggttagt tccacatagt tgcgcagat gacatgattt cagtaacctg gtcctggctga 600
aaagtgttgy ttgtgtgaat gatactgtgg ttcttggttc tcatctctg tgggtgaca 660
gctttctcca ccttgctgga agtgacctgc tgtccagaag ttgatggct gaggatata 720
ccatctgca tgcactcttc atttctgca ttctctctc cctggatgga caggggagc 780
ggcaagagca acgtggggc acgtggggac cacaacgact cctctgtgaa gacgcttgg 840
agcaagaggt gcaagtgtg ctgccactgc ttcccctgct gcagggggag cggcaagagc 900
aacgtgtctg cttggggaga ctacgatgac agcgcttcca tggatccag gtaccacgtc 960
catggagaag atctggacaa gctccacaga gctgcctggt ggggtaaagt cccagaaag 1020
gatctcatcg tcatgctcag ggacacggat gtgaacaaga gggacaagca aaagaggact 1080
gctctacatc tggcctctgc caatgggaat tcagaagtag taaaactcgt gctgcacaga 1140
ogatgtcaac ttaatgtcct tgacaaceaa aagaggacag cctctgacaa ggcgtacaa 1200
tgccaggaag atgaatgtgc gttaatgtg ctgggaacatg gcactgaccc aaatattcca 1260
gatgagtatg gaaataccac tctacactat gctgtctaca atgaagataa attaatggcc 1320
aaagcactgc tottataagg tgotgatata gaatcaaaaa acagcatyg cctcacacca 1380
ctgctacttg gtatacatga gcaaaaacag caagtgtgga aatttttaat caaaaaaaa 1440
gcgaatttaa atgogctgga tagatatgga agaactgctc tcatactgc tgtatgtgt 1500
ggatcagcaa gtatagtcag cctctctact gacgaatg ttgatgtatc ttctcaagat 1560
ctggaagagc ggccagagag tatgtgttt ctagtcatca tcatgtatc tggcagttac 1620
tttctgacta caaagaaaaa cagatgttaa aaatctcttc tgaacaacag aatccagaac 1680
aagacttaaa gctgacatca gaggaagagt cacaaggtt taaagggaag gaaaacagcc 1740
agccagagct agaagattta tggctattga agaagaatga agaaccagga agtactcatg 1800
tgggatctcc agaaaacctg actaacgggt cgcgtgctgg caatgggtgat ga 1852

```

```

<210> 4
<211> 292
<212> FRN
<213> Homo sapiens

```

```

<400> 4
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

```


WO 01/75171

PCT/US01/10631

4

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 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
 225 230 235 240
 Asn Ala Leu Asp Arg Tyr Gly Arg Thr Ala Leu Ile Leu Ala Val Cys
 245 250 255
 Cys Gly Ser Ala Ser Ile Val Ser Pro Leu Leu Glu Gln Asn Val Asp
 260 265 270
 Val Ser Ser Gln Asp Leu Glu Arg Arg Pro Glu Ser Met Leu Phe Leu
 275 280 285
 Val Ile Ile Met
 290

<210> 5
 <211> 1155
 <212> DNA
 <213> Homo sapien

<400> 5
 atggtggttg aggttgattc catgccggct gcctctcttg tgaagaagcc atttggcttc 60
 agggagcaaga tgggcaagtg gtgctgccgt tgcttccctt gctgcaggga gagcggcaag 120
 agcaacgttg gcacttcttg agaccacgac gactctgcta tgaagacact caggagcaag 180
 atgggcaast ggtgcggcca ctgcttcccc tctctcaggg gtagtggcaa gacacactg 240
 ggctctcttg gagaccaaga cgaactctgt atgaagacac tcaggaaaca gatggcaag 300
 tggctctgac actgcttccc ctgctgcagg gggagcgcca agagcaaggt gggcgcttg 360
 ggagactacg atgacagtgc ctctcatggg cccaggtacc acgtccgttg agaagactg 420
 gacaagctcc acagagctgc ctggtggggt aaagtcccca gaaaggatct catcgtcatg 480
 ctacaggaca ctgacgtgaa caagaaggac aagcaaaaga ggaactgctt acatctggcc 540
 tctgccaatg ggaattcaga agtagtaaaa ctctgctg acagacgatg tcaacttaat 600
 gtctctgaca acaaaaagag gacagctctg ataaaggccg tacaatgcca gsaagatgaa 660
 tgtcggttaa tgttgctgga acatggcact gatccaaata ttccagatga gtatgaaat 720
 accactctgc actacgctat ctataatgaa gataaattaa tggccaaagc actgctctta 780

WO 01/75171

PCT/US01/10631

5

```

tatggtgctg atatcgaaac aaaaaaacag catggcctca caacctgtt acttgggtga 840
catgagcaaa aacagcaagt cgtgaaattt ttaatacaaga aaaaagcgaa ttaaatgca 900
ctggatagat atggaaggac tgctctcata cttgctgtat gtttgggato agcaagtata 960
gtcagccttc tacttgagca aaatattgat gtatctcttc aagatctatc tggacagacg 1020
gccagagagt atgctgtttc tagtcacat catgtaattt gccagttact ttctgactac 1080
aaagaaaaac agatgctaaa aatctcttct gaaaacagca atccagaaaa tgtctcaaga 1140
accagaaata aataa 1155

```

```

<210> 6
<211> 2000
<212> DNA
<213> Homo sapien

```

```

<400> 6
atggtggttg aggttgatcc catgccggct gcctcttctg tgaagaagcc atttggcttc 60
aggagcaaga tgggcaagtg gtgctgccgt tgcttccctt gctgcaggga gagcggaag 120
agcaacgtgg gcacttctgg agaccacgac gactctgcta tgaagacact caggagcaag 180
atgggcaagt ggtgcgcca ctgcttccc tgctgcaggg ggaaggcaca gagcaacgtg 240
ggcgcttctg gagaccagca cgaactctgt atgaagacac tcaggaaaca gatgggcaag 300
tggtgctgcc actgcttccc ctgctgcagg gggagcgcca agagcaaggt gggcgcttgg 360
ggagactacg atgacagtgc ctctcatggg cccaggtacc acgtccgtgg agaagatctg 420
gacaaagctcc acagagctgc ctgggtgggt aaagtcccca gaaaggatct catctctatg 480
ctcaggggaca ctgacgtgaa caagaaggac aagcaaaaga ggaactgctc acatctggcc 540
tctgcaaatg ggaattcaga agtagtaaaa ctctctgctg acagacgatg tcaacttaat 600
gtccttgaca acaaaaagag gacagctctg ataaaggcgg tacaatgcca ggaagatgaa 660
tgtgtgttaa tgttgcgtga acatggcact gatccaaata ttccagatga gtatggaaat 720
accactctgc actacgtctat ctataatgaa gataaattaa tggccaaagc actgctotta 780
tatggtgctg atatcgaaac aaaaaaacag catggcctca caacctgtt acttgggtga 840
catgagcaaa aacagcaagt cgtgaaattt ttaatacaaga aaaaagcgaa ttaaatgca 900
ctggatagat atggaaggac tgctctcata cttgctgtat gtttgggato agcaagtata 960
gtcagccttc tacttgagca aaatattgat gtatctcttc aagatctatc tggacagacg 1020
gccagagagt atgctgtttc tagtcacat catgtaattt gccagttact ttctgactac 1080
aaagaaaaac agatgctaaa aatctcttct gaaaacagca atccagaaaa agacttaag 1140
ctgacatcag aggaagagtc acaaaagttc aaaggcagtg aaaaagcaca gccagagaaa 1200
atgtctcaag acccagaat aataaaggat ggtgatagag aggttgaaga agaatgaag 1260
aagcatgaaa gtaataatgt gggattacta gaaaacctga ctaatgggtg caactgtggc 1320
aatgttgata atggattaat tctctaaagg aagagcagaa caactgaaaa tcagcaattt 1380
ctgcaaacg aagtgaaaga gtatcaaga atttgggaat tagtttctga ctacaagaa 1440
aaacagatgc caaatactc ttctgaaaaa agcaacccag aacagactt aaagctgaca 1500
tcagagggaag agtcacaaag gcttgagggc agtgaataag gccagccaga gctagaaaaat 1560
tttatggcta tcgaagaaat gaagaagcac ggaagtactc atgtcggatt cccagaaaaac 1620
ctgactaatg gtgccactgc tggcaatggt atgatggat taattctccc aaggaagagc 1680
agaacacctg aaagcaagca atttctctgac actgagaatg aagagtatca cagtgcgcaa 1740
caaaatgata ctcaagaagca attttgtgaa gaacagaaca ctgggaatatt acacgatgag 1800
attctgattc atgaagaaaa gcagatagaa gtggttgaaa aatgaattc tgagctttct 1860
cttatgttga agaaagaaaa agacatcttg catgaaaaata gtacgttgcg ggaagaaatt 1920
gccatgctaa gactggagct agacacaatg aaacatcaga gccagctaaa aaaaaaaaaa 1980
aaaaaaaaaa aaaaaaaaaa 2000

```

```

<210> 7
<211> 2040
<212> DNA
<213> Homo sapien

```

```

<400> 7
atggtggttg aggttgatcc catgccggct gcctcttctg tgaagaagcc atttggcttc 60
aggagcaaga tgggcaagtg gtgctgccgt tgcttccctt gctgcaggga gagcggaag 120
agcaacgtgg gcacttctgg agaccacgac gactctgcta tgaagacact caggagcaag 180

```

WO 01/75171

PCT/US01/10631

6

```

atggggcaagt ggtgcgcgca ctgcttcccc tgetgcaggg ggaagtggcaa gagcaacgtg 240
ggcgctcttg gagaccacga cgcctctgct atgaagacac tcaggaaacaa gatgggcaag 300
tgggtgctgc actgottccc ctgctgcaggg ggaagcggca agagcaaggt gggcgcttgg 360
ggagactacg atgacagtgc ctctcatggag cccaggtacc acgtccgtgg agaagatctg 420
gacaagctcc acagagctgc ctggtgggggt aaagtccccc gaaaggatct catcgtcatg 480
ctcagggaca ctgacgtgaa caagaaggac aagcaaaaga ggaactgctct acatctggcc 540
tctgccaatg ggaattcaga agtagtaaaa ctctgtctgg acagacgatg tcaacttaat 600
gtccttgaca acaaaaagag gacagctctg ataaaggccg tacaatgcca ggaagatgaa 660
tgtcggttaa tgttgcctga acatggcact gatccaaata ttccagatga gtatggaaat 720
accactctgc actacgctat ctataatgaa gataaattaa tggccaaagc actgctctta 780
tatgtgtgtg atatogaatc aaaaaacaag catggcctca caccactgtt acttggtgta 840
catgagcaaa aacacgaagt cgtgaattt ttaatcaaga aaaaagcgaa ttaaatgca 900
ctgagatgat atggaaggac tgcctccta ctgtctgtat gttgtggatc agcaagata 960
ctcagccttc tacttgagca aaatattgat gtatcttctc agatctctat tggacagacg 1020
gocagagggt atgctgttct tagtcatcat catgtaattt gccagttact ttctgactac 1080
aaagaaaaac agatgctaaa aatctcttct gaaaacagca atccagaaca agacttaag 1140
ctgacatcag aggaagagtc acaagggttc aaaggcagtg aaatagacca gccagagaaa 1200
atgtctcaag aaccagaaat aaataaggat ggtgatagag aggttgaaaga agaaatgaag 1260
aagcatgaaa gtaataatgt gggattacta gaaaacctga ctaatgggtg cactgctggc 1320
aatggtgata atggattaat tcttcaaaag aagagcagaa cactgaaaa tcagcaattt 1380
ctcagcaacg aaagtgaaga gtatcacaga atttgcgaat tagtttctga ctacaagaa 1440
aaacagatgc caaaatactc ttctgaaaac agcaaccagc acaagactt aaagctgaca 1500
tcagaggaag agtcacaaag gcttgagggc agtgaanaat gccagccaga gaaaagatct 1560
caagaaccag aataaataaa ggaagggtgt agagagctag aaaaatttat ggtatctgaa 1620
gaaatgaaga agcagcgaag tactcatgtc ggaattccag aaaaactgac taatggtgcc 1680
actgctggca atggtgatga tggattaatt cctccaaaga agagcagaac acctgaaagc 1740
cagcaatttc ctgcactga gaatgaagag tatccacagt acgaacaaaa tgatactcag 1800
aagcaatttt gtgaagaaca gaacactgga atattacacg atgagattct gattcatgaa 1860
gaaaagaca tgaagtggt tgaaaaaatg aattctgagc ttctctttag ttgtaagaaa 1920
gaaaagaca tcttgcatga aaatagtagc ttgcgggaag aaattgccat gctaagactg 1980
gagctagaca caatgaaaca tcagagccag ctaaaaaaaa aaaaaaaa 2040

```

```

<210> 8
<211> 384
<212> PRT
<213> Homo sapien

```

```

<400> 8
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

```

WO 01/75171

PCT/US01/10631

7

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 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> 9
<211> 656
<212> PRT
<213> Homo sapien

<400> 9
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

WO 01/75171

PCT/US01/10631

8

```

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
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
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
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
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
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
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
Ala Met Leu Arg Leu Glu Leu Asp Thr Met Lys His Gln Ser Gln Leu
640

```

WO 01/75171

PCT/US01/10631

9

645 650 655

<210> 10
<211> 671
<212> PRT
<213> Homo sapien

<400> 10

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
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

WO 01/75171

PCT/US01/10631

10

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
 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> 11
 <211> 800
 <212> DNA
 <213> Homo sapien

<400> 11
 atkagcttcc gctctgaca acactagaga tccctcccct cccctcagggt atggccctcc 60
 acttcatttt tggtagataa catctttata ggacaggggt aaaatcccaa tactaacagg 120
 agaatgctta ggactctaac aggtttttga gaatgtgttg gtaagggccca ctcaatccaa 180
 tttttcttgg tccctcttgt ggtctaggag gacagggcaag ggtgcagatt tcaagaatg 240
 catcagtaag ggcactaaa tccgaccttc ctctgtcttc cttgtgtctt gggaggaaaa 300
 ctagtgtttc tgttgctgtg tcagttagca caactattcc gatcagcagg gtccagggac 360
 cactgcaggt tcttgggcag ggggagaaac aaacaaaacc aaaacctgg gcrgttttgt 420
 ctttcagatg ggaacactc aggcataaac aggcatacct ttgaatgca tccaaagcca 480
 atggggacaaa ttgacccac aaaccttggg aaaaagaggtg gctcattttt ttgcactat 540
 ggcttggccc caacattctc tctctgatgg ggaataatgg ccacctgaug gaagtacaga 600
 ttcaataact atcctgagc ttgacctttt ctgtgaaggg gaaggcaaat gtagtgaat 660
 acctatgtc caagctttct ttctattgaa ggagaataca ctatgcaag ctgaaattt 720
 acatcccaaa ggaggacctc tcagcttacc cccatattct agcctccota tagctccct 780
 tctattagt gataagctc 800

<210> 12
 <211> 102

WO 01/75171

PCT/US01/10631

11

```
<212> PRT
<213> Homo sapien

<220>
<221> VARIANT
<222> (1)...(102)
<223> Xaa = Any Amino Acid

<400> 12
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
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> 13
<211> 1206
<212> DNA
<213> Homo sapien

<400> 13
ggcagcaggga agttttgtgt actgaaaaag aaactgtcag aagcaaaaga aataaaatca 60
cagtttagaga accaaaaagt taaatgggaa caagagctct gcagtgtgag gttttccaca 120
ctcatgaaaa tgaaaattat ctcttacatg aaatgtcat gttgaaaaag gaaattgccca 180
tgctaaaact ggaatagacc acactgaaac accaatacca ggaaggaa aataaatact 240
ttgaggacat taagatttta aaagaaaaaga atgctgaact tcagatgacc ctaaaactga 300
aagaggaaac attaacaaaa agggcatttc aatatagttg gcagottaaa gtctgatag 360
ctggagaacac aatgctcaat tctaaattga agyaaaaca agacaagaa atactagag 420
cagaaattga atcacaccat cctagactgg ctctgctgtt acaagaocat gatcaattg 480
tgacatcaag aaaaagtcaa gaactgctt tccacattgc agagatgct tgtttgcaaa 540
gaaaaatgaa tgttgatgtg agtatacga tatataacaa tgaggtgct catcaaccac 600
tttctgagc tcaaggaaa tccaaaagcc taataathea tctcaattat gccggagatg 660
ctctaaagaa aatacattg gtttcagaac atgcacaaag agaccaactg gaacacagt 720
gtcaaatgaa ggaagctgaa cacatgtatc aaaacgaaca agataatgtg aacaaacaca 780
ctgaacagca ggaagtctcta gatcagaat tatttcaact acaagcaaa aatatgtggc 840
ttcaacagca attagtctat gcacataaga aagctgacaa caaagcaag ataacaattg 900
atattcattt tcttgagggg aaatgcaac atcatctcct aaaagagaaa aatgaggaga 960
tatttaatta caataaccat ttaaaaaacc gtatatatca atatgaaaaa gagaagcag 1020
aaacagaagt tatataatag tataacactg ccaaggagcg gattatotca tcttcatcct 1080
gtaattcaag tgtttgtcac gtggttggtt aataaatgaa taaagaatga gaaaaccaga 1140
agctctgata cataatcata atgataatta tttoaatgca caactacggg tgggtgctgct 1200
cgtgcc
```


WO 01/75171

PCT/US01/10631

12

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

<210> 15

<211> 1665

<212> DNA

<213> Homo sapien

<400> 15

gcaaaactttc	aagcagagcc	tcccagaag	ccatctgcct	togagcctgc	cattgaaatg	60
caaaagtctg	ttccaaataa	agccttggaa	ttgaagaatg	aacaacatt	gagagcagat	120
cagatgttcc	cttcagaatc	aaaacaaaag	aaggttgaag	aaaattcttg	ggattctyag	180
agtctccttg	agactgtttc	acagaagat	gttgtgtac	ccaaggtac	acatcaaaaa	240
gaatggata	saataagtgg	aaattagaa	gattcaacta	gootatcaaa	aatcttggat	300
acagttcatt	cttgtgaaag	agcaaggga	cttcaaaaag	atcactgtga	acaactaca	360
ggaaaaatgg	aacaaatgaa	aaagaagttt	tgtgtactga	aaaagaaact	gtcagaagca	420
aaagaaataa	aatcacagtt	agagaacca	aaagttaaat	gggaacaaga	gctctgcagt	480
gtgaggtttc	tcacactcat	gaaaatgaaa	attatctctt	acatgaaaat	tgcatgttga	540
aaaaggaaat	tgccatgcta	aaactggaaa	tagccacact	gaacaccaa	taccaggaaa	600
aggaaaataa	atactttgag	gacattaaga	ttttaaaaga	aaagaatgct	gaacttcaga	660

WO 01/75171

PCT/US01/10631

13

```

tgaccctaaa actgaagag gaactaataa ctaaaagggc atctcaatat agtggggaag 720
ttaaagtctt gatagctgag aacscaatgc tcaactctaa attgaaggaa aaacaagaca 780
aagaataact agaggcagaa attgaatcac accatcctag actggcttct gctgtacaag 840
accatgatca aattgtgaca tcaagaanaa gtcaagaacc tgctttccac attgcaggag 900
atgcttggtt gcaagaaaa atgaatgttg atgtgagtag taagatatat aacaatgagg 960
tgctccatca accactttct gaagctcaaa ggaatccaa aagcctaata ataatctca 1020
attatgccgg agatgtctta agagaaaata cattggtttc agaactatga caaagagacc 1080
aacgtgaac acagtgtcaa atgaaggaag ctgaacacat gtatcaaaac gaacaagata 1140
atgtgaacaa acacactgaa cagcaggagt ctctagatca gaattatttt caactacaaa 1200
gcaaaaatat gtggcttcaa cagcaattag ttcatgcaca taagaaagct gacaacaaaa 1260
gcaagataac aattgatatt cattttcttg agaggaaaat gcaacatcat ctctaaaaag 1320
agaaaaatga ggagatattt aattacaata accatttaas aaacggtata tatcaatatg 1380
aaaaagagaa agcagaacaa gaanaactcat gagagacaaag cagtaagaaa ctctctttgg 1440
agaaacaca gacagatctt ttactcaaaa ctcatgctag gagggcagtc ctgacattac 1500
cttatgttga aaatcttacc aatagtcgtg gtcaacagaa tacttttttt agaagaaaaa 1560
ttcatgattt ctctctgaag cotggcgagc agagcgagac tctgtctcaa aaaaaaaaaa 1620
aaaaaaagaa agaaagaat gctgtgctt acttcgcttc ccagg 1665

```

<210> 16
 <211> 179
 <212> PRT
 <213> Homo sapien

```

<400> 16
Ala Asn Phe Gln Ala Glu Pro Pro Glu Lys Pro Ser Ala Phe Glu Pro
1 5 10 15
Ala Ile Glu Met Gln Lys Ser Val Pro Asn Lys Ala Leu Glu Leu Lys
20 25 30
Asn Glu Gln Thr Leu Arg Ala Asp Gln Met Phe Pro Ser Glu Ser Lys
35 40 45
Gln Lys Lys Val Glu Glu Asn Ser Trp Asp Ser Glu Ser Leu Arg Glu
50 55 60
Thr Val Ser Gln Lys Asp Val Cys Val Pro Lys Ala Thr His Gln Lys
65 70 75 80
Glu Met Asp Lys Ile Ser Gly Lys Leu Glu Asp Ser Thr Ser Leu Ser
85 90 95
Lys Ile Leu Asp Thr Val His Ser Cys Glu Arg Ala Arg Glu Leu Gln
100 105 110
Lys Asp His Cys Glu Gln Arg Thr Gly Lys Met Glu Gln Met Lys Lys
115 120 125
Lys Phe Cys Val Leu Lys Lys Lys Leu Ser Glu Ala Lys Glu Ile Lys
130 135 140
Ser Gln Leu Glu Asn Gln Lys Val Lys Trp Glu Gln Glu Leu Cys Ser
145 150 155 160
Val Arg Phe Leu Thr Leu Met Lys Met Lys Ile Ile Ser Tyr Met Lys
165 170 175
Ile Ala Cys

```

<210> 17
 <211> 1681
 <212> DNA
 <213> Homo sapien

```

<400> 17
gatacagtc tttctgtgaa agagcaagg aacttcaaaa agatcaactgt gaacaacgta 60
caggaaaaat ggaacaaaatg aaaaagaagt ttgtgtact gaaaaagaaa ctgtcagaag 120
caaaagaaat aaaatcacag ttagagaaco aaaaagttaa atgggaacaa gagctctgca 180

```

WO 01/75171

PCT/US01/10631

14

```

gtgtgagatt gactttaaac caagaagaag agaagagaag aaatgcgat atattaaatg 240
aaaaaatag ggaagaatta ggaagaatcg aagagcagca taggaagag ttagaagtga 300
aacaacaact tgaacaggct ctcgaatac asgatataga attgaagagt gtagaagtga 360
atttgantca ggtttctcac actcatgaaa atgaaaatta tctcttcaat gaaaattgca 420
tggtgaaaaa ggaattgcc atgctaaaaa tggaaaatgc cacactgaaa caccaatacc 480
aggaaaagga aataaatac tttgaggaca ttaagatttt aaaagaaaag aatgctgaac 540
ttcagatgac cctaaaactg aaagagggaat cattaactaa aagggcactc caatatagtg 600
ggcagcttaa agttctgata gctgagaaca caatgctcac ttctaattg aaggaanaac 660
aagcaaaaga aatactagag gcagaaattg aatcacacca tcttagactg gcttctgctg 720
tacaagacca tgatcaaatt gtgacatcaa gaaaagtca agaactgct ttccacattg 780
caggagatgc ttgtttgcaa agaaaaatga atgttgatg tagtagtaag atataataca 840
atgaggtgct ccatacaacca ctttctgaag ctcaaaaggaa atccaaaagc ctaaaaatta 900
atctcaattt tgcgggagat gctctaaag aaatacaatt ggtttcagaa catgcacaaa 960
ggagccnagc tgaacacnag tctcaaatga aggaagctga acacatgctt caaaagaaac 1020
nagataatgt gaacaaacac actgaacagc agtagctctc agatcagaa ttatttcaac 1080
tacaagcaaa aatatgtgtg cttcaacagc aattagtcca tgcacataag aaagctgaca 1140
acaaaagcaa gataacaatt gatattcatt tctttgagag gaaaatgcaa catcatctcc 1200
taaaagagaa aatgaggag atatttaatt acaataacca tttaaaaaac cgtatatatc 1260
aatatgaaaa agagaagcaa gaacagaaaa actcatgaga gacacagagt aagaaacttc 1320
ttttggagaa acaacagacc agatctttac tcacaactca tgctaggagc ccagtcctag 1380
cattacctta tgttgaaaaa tcttccaat agtctgtgtc aacagaatc ttattttaga 1440
agaaaatttc atgattttct cctgaagcct acagacataa aataacagt tgaagaatta 1500
ctgtttcacg aattgcataa aagctgcccc ggatttccat ctaccctgga tgatgcogga 1560
gacatcattc aatccaacca gaattctgct ctgtcactca ggctggagtg cagtgggcgc 1620
aatctggctc caactgcaact ctgctctcca ggttcaagcc attctctggc acagctcccc 1680
g 1681

```

<210> 18
 <211> 432
 <212> PRT
 <213> Homo sapien

<400> 18
 Asp Thr Val His Ser Cys Glu Arg Ala Arg Glu Leu Gln Lys Asp His
 1 5 10 15
 Cys Glu Gln Arg Thr Gly Lys Met Glu Gln Met Lys Lys Phe Cys
 20 25 30
 Val Leu Lys Lys Lys Leu Ser Glu Ala Lys Glu Ile Lys Ser Gln Leu
 35 40 45
 Glu Asn Gln Lys Val Lys Trp Glu Gln Glu Leu Cys Ser Val Arg Leu
 50 55 60
 Thr Leu Asn Gln Glu Glu Lys Arg Arg Asn Ala Asp Ile Leu Asn
 65 70 75 80
 Glu Lys Ile Arg Glu Glu Leu Gly Arg Ile Glu Glu Gln His Arg Lys
 85 90 95
 Glu Leu Glu Val Lys Gln Gln Leu Glu Gln Ala Leu Arg Ile Gln Asp
 100 105 110
 Ile Glu Leu Lys Ser Val Glu Ser Asn Leu Asn Gln Val Ser His Thr
 115 120 125
 His Glu Asn Glu Asn Tyr Leu Leu His Glu Asn Cys Met Leu Lys Lys
 130 135 140
 Glu Ile Ala Met Leu Lys Leu Glu Ile Ala Thr Leu Lys His Gln Tyr
 145 150 155 160
 Gln Glu Lys Glu Asn Lys Tyr Phe Glu Asp Ile Lys Ile Leu Lys Glu
 165 170 175
 Lys Asn Ala Glu Leu Gln Met Thr Leu Lys Leu Lys Glu Glu Ser Leu
 180 185 190
 Thr Lys Arg Ala Ser Gln Tyr Ser Gly Gln Leu Lys Val Leu Ile Ala

WO 01/75171

PCT/US01/10631

15

195 200 205
Glu Asn Thr Met Leu Thr Ser Lys Leu Lys Glu Lys Gln Asp Lys Glu
210 215 220
Ile Leu Glu Ala Glu Ile Glu Ser His His Pro Arg Leu Ala Ser Ala
225 230 235 240
Val Gln Asp His Asp Gln Ile Val Thr Ser Arg Lys Ser Gln Glu Pro
245 250 255
Ala Phe His Ile Ala Gly Asp Ala Cys Leu Gln Arg Lys Met Asn Val
260 265 270
Asp Val Ser Ser Thr Ile Tyr Asn Asn Glu Val Leu His Gln Pro Leu
275 280 285
Ser Glu Ala Gln Arg Lys Ser Lys Ser Leu Lys Ile Asn Leu Asn Tyr
290 295 300
Ala Gly Asp Ala Leu Arg Glu Asn Thr Leu Val Ser Glu His Ala Gln
305 310 315 320
Arg Asp Gln Arg Glu Thr Gln Cys Gln Met Lys Glu Ala Glu His Met
325 330 335
Tyr Gln Asn Glu Gln Asp Asn Val Asn Lys His Thr Glu Gln Gln Glu
340 345 350
Ser Leu Asp Gln Lys Leu Phe Gln Leu Gln Ser Lys Asn Met Trp Leu
355 360 365
Gln Gln Gln Leu Val His Ala His Lys Lys Ala Asp Asn Lys Ser Lys
370 375 380
Ile Thr Ile Asp Ile His Phe Leu Glu Arg Lys Met Gln His His Leu
385 390 395 400
Leu Lys Glu Lys Asn Glu Glu Ile Phe Asn Tyr Asn Asn His Leu Lys
405 410 415
Asn Arg Ile Tyr Gln Tyr Glu Lys Glu Lys Ala Glu Thr Glu Asn Ser
420 425 430

<210> 19
<211> 3681
<212> DNA
<213> Homo sapiens

<400> 19
tcagagctga ttacagacac caaggaagat gctgtaaaga gtcagcagcc acagccoctgg 60
ctagctggcc ctgtgggcat ttattagtaa agttttaatg acaaaagctt tgaatcaaca 120
cacccgtggg taattaaact ggtcatcccc accctggaga gccatcctgc coactgggtga 180
tcaaaagaag aacatctgca gaaacacctg atgaggctgc acccttgccg gaaagaacac 240
ctgacacagc tgaagcttg gtgaaaaaaa caccctgata ggctgcacc ttggtggaaa 300
gaacacctga cagcgctgaa agcttggtgg aaaaaacacc tgaatgagct gcatccttgg 360
tgagggggac atctgacaaa attcaatgtt tggagaagac gacatctgga aagtctgaac 420
agtcagcaga agaaacacct agggaaatta cgagctctgc aaaaagaaaca totgagaat 480
ttacgtggcc agcaaaagga agacctagga agatcgcatg ggagaaaaaa gaagcacacac 540
ctagggaat tatgagctcc gcaaaagaaa catctgagaa attacgtgg gcagcaaaag 600
gaagacctag gaagatcgca tgggagaaaa aagaacacc tgtaaagact ggaatcgctg 660
caagagtaac atctaataaa actaaagtgt tggaaaaagg aagatctaag atgattgcat 720
gtcctacaaa agaatacatc ncaaaagcaa gtgccaatga tcagagggtc coactcagaat 780
ccaaacaaga ggaagatgaa gaatatcttt gtgattctog gagtctcttt gagagttctg 840
caaaagattca agtctgtata cctgagctta tatatacaaa agtaatggag atcaaatagag 900
aagtagaaga gctctctaag aagccatctg ccttcaagcc tgcacattga atgcaaaact 960
ctgttccaaa taaagccttt gaattgaaga atgaacaacac attgagagca gatccgagt 1020
tccacacaga atccaaacaa aaggactatg aaaaaattc ttgggattct gagagttctt 1080
gtgagactgt ttccacagaag gatgtgtgtt taccacaagg tacacatcaa aaagaaatag 1140
ataaaataaa tggaaaatta gaagagctct ctaataaaga tggctctctg aaggctacct 1200
gcggaatgaa agtttctatt ccaactaaag cottagaatt gaaggacatg caaactttca 1260
aagcagagcc tccggggaag ccactctgct toagagctgc caetgaaatg caaaagtctg 1320

WO 01/75171

PCT/US01/10631

16

```

tcccaataa agccttggaa ttgaaaaatg aacaaacatt gagagcagat gagatactcc 1380
catcagaatc caaacaaaag gactatgaag aaagtctctg ggattctgag agtctctgtg 1440
agactgtttc acagaaggat gtgtgtttac ccaaggctro rcatacaaaa gaaatagata 1500
aaataaatgg aaaattagaa ygtgtctctg ttaaagatgg tcttctgag gctaactcgg 1560
gaatgaaagt ttctattcca actaaagcct tagaattgat ggacatgcaa actttcaaa 1620
cagagcctcc cgagaagcca tctgctctgc agcctgccat tgaatgcaa aagtctgttc 1680
caataaagc cttggaattg aagaatgaac aaacattgag agcagatgag ataactccat 1740
cagaatccaa acaaaaggac tatgaagaaa gttcttggga ttctgagagt ctctgtgaga 1800
ctgtttcaca gaaggatgtg tgtttaccga aggcctrcra tcasaaagaa atagataaaa 1860
taaatggaaa attagaagag tctcctgata atgatgtgtt tctgaaggct cctgcagaaa 1920
tgaagtttc tattccaaact aaagccttag aattgatgga catgcaaaact tcaaaagcag 1980
agccttcccg gaagccatct gccttcgagc ctgcctatga aatgcasaaag tctgttccaa 2040
ataagcctt ggaattgaa atgaacaaa catctgagag agatcagatg ttcccttcag 2100
aatcaaaaaa aaagaasgtt gaagaaaatt cttggatctc tgaagatctc cgtgagatg 2160
tttcacagaa ggatgtgtgt gtacccaagg ctacacatca aaaaagaatg gataaataa 2220
gtggaaaatt agaagattca actagcctat caaaaatott ggatacagtt cattctgtg 2280
aaagagcaag ggaacttcaa aaagatcact gtgaacaacg tacaggaaaa atggaacaaa 2340
tgaaaaagaa gttttgtgta ctgaaaaaga aactgtcaga agcaaaagaa ataaaaatca 2400
agtttagagaa ccaaaaagtt aaatgggaac aagagctctg cagtgtgagg ttctccacac 2460
tcatgaaat gaaaattatc tcttcatga aaattgcag ttgaaaaagg aaattgcact 2520
gctaaactg gaaatagcca cactgaacaa ccaataccag gaaaaggaaa ataactactt 2580
tgaggacatt aagattttta aagaaaagaa tgcagacatt cagatgaccc taaaactgaa 2640
agaggaaatca ttaactaaaa gggcatctca atatagtggg cagcttaaa tctgtatagc 2700
tgagaacaca atgctcactt cttaattgaa ggaasaaaca gacaaagaaa tactagaagg 2760
agaaattgaa tcaacacatc ctgactggc ttctgtctga caagaccatg atcaaatgtt 2820
gacatcaaga aaagtcag agcctgtctt caaatgtga ggaatgctt gtttgcaag 2880
aaaaatgaat gttgatgta gtatcagat atataacaa gaggtgtctc atcaacacat 2940
ttctgaagct caaggaat caaaagcct aaaaatfaat ctcaattatg cmggagatgc 3000
tctaagaaa aatacattg tttcagaaca tgcacaaaga gaccaacgtg aaacacagt 3060
tcaaatgaag gaagctgaac acatgtatca aaacgaaca gataatgtg acaaacacac 3120
tgaaacagcag gagtctctag atcagaatatt atttcaacta caaagcaaaa atatgtggct 3180
tcaacagcaa ttagttcatg cacataagaa agctgacac aaaaagcaaga taacaattga 3240
tattcatttt cttgagagga aaatgcaaca tcatctccta aaagagaaaa atgagagat 3300
atttaattac aataaccatt taaaaaacg tatatatca tatgaaaaag aagaagcaga 3360
aacagaaaaa tcatgagaga caagcagtaa gaaactctt ttggagaac aacagacag 3420
atctttactc acaactcatg ctaggaggcc agtctatgca tcaacttatg ttgaaatct 3480
taocaaatagt ctgtgtcaac agaatactta ttttagaaga aaatttcag atttctctct 3540
gaagcctaca gacataaaat aacagtgtga agaattcatt gllcaagaa tgataaagc 3600
tgacacagat tccatctac cctgatgatg cagcagacat catcaatcc aaccagaatc 3660
tcgtctctgt actcaggtg g 3681

```

<210> 20
 <211> 1424
 <212> DNA
 <213> Homo sapiens

```

<400> 20
tcagagctga ttacagacac caaggaagat gctgtaaga gtcagcagcc acagccctgg 60
ctagctggcc ctgtgggcat ttattagtaa agtthtaag acaaaagctt tgaatcaaca 120
cacccgtggg taattcaact ggtcactccc accctggaga gccatctctg ccatgggtga 180
tcaaaagaag aacatctgca ggaacaactg atgaggtgc acccttggc gaagaacac 240
ctgacacagc tgaagcttg gtggaasaaa oactgatga ggcctgaccc ttgttgaaa 300
gaacacctga cacggtgaa agcttgggtg aaaaaacac tgaagagct gcatctctgg 360
tgagaggaa atctgacaaa attcaatgtt tggagaagc gacatctgga aagtctgaac 420
agtcagcaga agaaacacat agggaaatta cgaactctgc aaaaagaaa tctgagaat 480
ttacgtggcc agcaaaagga agacatagga agatcgcatg ggaagaaaaa gaagacacac 540
ctagggaat tatgagtcac gcaaaagaaa catctgagaa atttactgg gcaagcaaa 600
gaagacctag gaagatcgca tgggagaaaa aagaacaccc tgtaaagact ggaatcggtg 660

```

WO 01/75171

PCT/US01/10631

17

```

caagagtaac atctaataaa actaaagttt tggaaaaagg aagatctaag atgaitcoat 720
gtcctacaaa agaatacatct acaaaagcaa gtgccaatga tcaagagttc coatacagaat 780
ccaacaaga ggaagatgaa gaataattct gtgattctcg gagtctcttt gagagttctg 840
caaaagattca agtgtgtata cctgagctta tatatacaaa agtaattggag ataaatagag 900
aagtagaaga gctcctaag aagccatctg ccttcaagcc tgcattgaa atgcaaaact 960
ctgttccaaa taaagccttt gaattgaaga atgaacaac attgagagca gatccgatgt 1020
tcccaccaga atccaaacaa aaggactatg aagaaaattc ttgggattct gagagtctct 1080
gtgagactgt ttccacaag gatgtgtgtt tacccaagcc tacacatcaa aaagaatatg 1140
ataaaataaa tggaaaatta gaaggtaga accgtttttt atttaaaat cagttgaccg 1200
aatatttctc taaactgatg agggggata tccctagta gctgaagaaa attactcctc 1260
aatgcaaac catgaaaaa aagagaagtg caatggctgt aagttgtatg tctcatcagg 1320
tgttggcaac agactatatt gagagtgtgt aaaaaggagt gaattattag ttgaaattca 1380
agatattgca agacctgaga gaaaaaaaaa aaaaaaaa aaaa 1424

```

<210> 21

<211> 674

<212> DNA

<213> Homo sapiens

<400> 21

```

attccgagct gattacagac accaaggaag atgctgtaaa gagtacgag ccacagccct 60
ggctagctgg cctgtgggc atttattagt aaagttttaa tgacaaaagc tttagtcaa 120
cacaccctgg ggttaattac ctggctatcc ccaccctgga gagccatcct gccatgggt 180
galtcaagaa ggaacatctg caggaaaccc tgatgaggtt gaccccttgg cygaagaagc 240
acttgacaca gctgaagct tggtagaaa aacacctgat gaggtgcaac ccttggtaga 300
aagacacact gacacggctg aaagcttggg ggaaaaaaca cctgatgagg ctgcatcctt 360
gttggaggga acatctgaca aaattcaatg tttaggaaa ggcacatctg gaaagtoga 420
acagtcagca gaagaacac ctagggaatg tacgagtctt gcaaaagaaa catctgaga 480
atttacgttg ccagcaaaag gaagacctag gaagatcgca tgggagaaaa aagatgactc 540
agttaagcca aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 600
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 660
aaaaaaaaaa aaaa 674

```

<210> 22

<211> 1729

<212> DNA

<213> Homo sapiens

<220>

<221> unsure

<222> (11)

<223> n=A,T,C or G

<221> unsure

<222> (1128)

<223> n=A,T,C or G

<400> 22

```

gaaagtctga ncagtcagca gaagaacac ctagggaat taagagtctt gcaaaagaaa 60
catctgagaa atttaogtgg ccagcaaaag gaagacctag gaagatcgca tgggagaaaa 120
aagaagacac acctagggaa attatgagtc cggcaaaaga aacatctgag aaatttaagt 180
gggcagcaaa aggaagacct aggaagatcg catgggagaa aaagaacaca cctgtaaaga 240
ctggatcggt ggcaagagta acatctaata aaactaaagt tttagaaaaa ggaagatcta 300
agatgattgc atgtcctaca aaagaatcat ctacaaaagc aagtgcocat gatcagaggt 360
tcccatcaga atccaaacaa gaggaagatg aagataatc ttgtgattct cggagtctct 420
ttgagagttc tgcaaaagatt caagtgtgta taccctgagtc tatatatcaa aaagtaattg 480
agataaatag agaagtagaa gaggctccta agaagccatc tgccttcaag cotgccattg 540
aaatgcaaaa ctctgttcca aataaagcct ttgaattgaa gaatgaacaa acattgagag 600
cagatccgat gttccaccca gaatccaaac aaaaggacta tgaagaaat tottgggatt 660

```

WO 01/75171

PCT/US01/10631

18

```

ctgagagctct ctgtgagact gtttcacaga aggatgtgtg tttaacccaag gctacacato 720
aaaaagaaat agataaaata aatggaaaat tagaagagtc tccataataa gatggctctc 780
tgaagggtac ctgoggaatg aaagtgttota ttccaaactaa agccttagaa ttgaaggaca 840
tgcaaaacttt caaagcagag cctcogggga agccatctgc cticgagcct gccactgaaa 900
tgcaaaagtc tgtcccaaat aaagccttgg aattgaaaaa tgaacaaaca ttgagagcag 960
atgagatact cccatcagaa tccaaacaaa aggaactatga agaaattctt tgggatactg 1020
agagtctctg tgagactgtt tcacagaagg atgtgtgttt acccaaggct ggcacataaa 1080
aagaataaga taaaataaat ggaataatag aagggtctcc tggtaaanat ggtcttctga 1140
aggctaaactg cgaatgaaa gtttctatcc caactaaagc cttagaattg atggacatgc 1200
aaactttcaa agcagagcct cccgagaagc catctgctct cgagcctgcc attgaaatgc 1260
aaagtctgtt tccaaataaa gctctgggat tgaagatga acaaacattg agagcagatg 1320
agataactccc atcagaatcc aaacaaaagg actatgaaga aagttcttgg gattctgaga 1380
gtctctgtga gactgtttca cagaaggatg tgtgtttacc caaggctgcg catcaaaaag 1440
aaatagataa aataaatgga aaattagaag gtaagaacog tttttatatt aaaaatcatt 1500
tgaccaataa tttctctaaa ttgatgagga aggatctcct ctatagctg agaaaaatta 1560
octcctaagt gcaaacatcg gaaaaaaga gaagtgaat ggtcataagc tatgtgtctc 1620
atcaggcatt ggcaacagac tatattgtga gtgctgaaga ggagctgaat tactagttaa 1680
aattcaagat attccaagac gtgaggaaaa tgagaaaaaa aaaaaaaaaa 1729

```

```

<210> 23
<211> 1337
<212> DNA
<213> Homo sapiens

```

```

<400> 23
aaaaagaaat agataaaata aatggaaaat tagaagggtc tctgtttaa gatggctctc 60
tgaagggtac ctgoggaatg aaagtgttota ttccaaactaa agccttagaa ttgatggaca 120
tgcaaaacttt caaagcagag cctcogggga agccatctgc cticgagcct gccactgaaa 180
tgcaaaagtc tgttccaaat aaagccttgg aattgaaaga tgaacaaaca ttgagagcag 240
atgagatact cccatcagaa tccaaacaaa aggaactatga agaaattctt tgggattctg 300
agagtctctg tgagactgtt tcacagaagg atgtgtgttt acccaaggct ggcacataaa 360
aagaataaga taaaataaat ggaataatag aaggtctccc tgataatgat ggtttcttga 420
aggctcctcg cagaatgaaa gtttctatcc caactaaagc cttagaattg atggacatgc 480
aaactttcaa agcagagcct cccgagaagc catctgctct cgagcctgcc attgaaatgc 540
aaagtctgtt tccaaataaa gctctgggat tgaagatga acaaacattg agagcagatc 600
aaagtctctc ttcaaatcaa aaacaaaaga aggttgaaga aaattcttgg gattctgaga 660
gtctcctgta gactgtttca cagaaggatg tgtgttacc caaggctaca catcaaaaag 720
aaatggataa aataagtgga aaattagaag ttccaaacta cctatcaaaa atcttgata 780
cagtctatcc ttgtgaaaga gcaagggaac ttcaaaaaga tcactgtgaa caactgacag 840
gaaaaatgga acaaatgaaa aagaagtttt gtgtactgaa aaagaactg tcagaagcaa 900
aagaataaaa atcaacagta gagaacccaa aagttaaatg ggaacagag ctctgcagtg 960
tgagattgac tttaaaccaa gaagaagaga agagaagaaa tgccgatata ttaaatgaaa 1020
aaattaggga agaattagga aagaatcgaag agcagatag gaagagtta gaagtgaac 1080
aacaacttga acaggtctct agaatacaag atatagaatt gaagagtga gaagtgaatt 1140
tgaatcaggt ttctcactc catgaaaatg aaattatct ctacactgaa aattgcattg 1200
tgaaaaagga aattgccatg ctaaaactgg aatatgccac actgaaacac caataccagg 1260
aaagggaaaa taaatacttt gaggacatta agattttaaa agaaaagaat gctgaacttc 1320
agatgacccc togtgcc 1337

```

```

<210> 24
<211> 2307
<212> DNA
<213> Homo sapiens

```

```

<400> 24
attgagagca gatgagatc tccatcaga atccaaacaa aaggactatg aagaaagttc 60
ttgggattct gagagtctc gtgagactgt ttacagaag gatgtgtgtt taccgaaggc 120
tacacatcaa aaagaatatg ataaaataaa tggaaaatta gaagggtctc ctgttaaaga 180

```

WO 01/75171

PCT/US01/10631

19

```

tggtcttctg aaggctaact gcgggaatgaa agtttctatt ccaactaaag ccttagaatt 240
gatggacatg caaactttca aagcagagcc tcccgagaag coactctgct tcgagcctgc 300
cattgaaatg caaaagtctg ttccaaataa agccttggaa ttgaagaatg acaaacatt 360
gagagcagat gagatactcc catcagaatc caaacaasaag gactatgaag aaagtctctg 420
ggattctgag agtctctgtg agactgtttc acagaaggaat gtgtgtttac ccaaggctac 480
acatcaaaaa gaatatagata aaataaatgg aaaattagaa gagtctctctg ataagtatgg 540
ttttctgaag tctccctgca gaatgaaagt ttctattcca actaaagcct tagaattgat 600
ggacatgcaa acotttcaag cagagcctcc cgagaagcca tctgctctcg agcctgccat 660
tgaatgcaa aagtcctgttc caaataaagc ctgtgaattg aagaatgaac aaacattgag 720
agcagatcag atgttccott cagaatcaaa acaaaagaac gttgaagaaa attcttggga 780
ttctgagagt ctccgtgaga ctgtttcaca gaaggaatgt gtgtaccca aggtacaca 840
tcaaaagaa atggataaaa taagtggaaa attgaagat tcaactagcc tatcaaaaat 900
cttggataca gtctattctt gtgaasagc aagggaactt caaanagatc actgtgaaca 960
acgtacagga aaaaatggaac aaatgaaaa gaagttttgt gtactgaaaa agaaactgtc 1020
agaaacaaa gaataaaaat cacagttaga gaacaaaaaa gttaaatggg aacagagact 1080
ctgcagtggt aggtttctca cactcatgaa aatgaaatc attctttaca tgaaaattgc 1140
atgttgaaaa aggaatttgc catgctaaaa ctggaatatg ccacactgaa acaccaatc 1200
caggaaaagg aaaaataata ctttgaggac attaaattt taaaagaaaa gaatgtctga 1260
cttcagatga cctcaaaact gaagaggaaa tcaataacta aaagggcacc tcaatatagt 1320
gggcagctta aagttctgat agctgagaa acaatgtcca ctctaaatt gaaggaaaaa 1380
caagacaaag aactactaga gccagaaatt gaatcacacc atcctagact ggcttctgct 1440
gtacaagacc atgatcaaat tgtgacatca agaaaagtc aagaacctgc ttccacatt 1500
gcaggagatg cttgttttga aagaaaaatg aatgttatg tgagtgtac gatataaac 1560
aatgaggtgc tccatcaacc actttctgaa gctcaaaaga aatccaaaag cctaaaaatt 1620
aatccaaat atgagagaga tgcctcaaga gaaatacat tggtttcaga acatgcaca 1680
agagaccaac gtgaacaca gtgtcaaatg aagggaactg aacacatgta tcaaacgaa 1740
cagataatg tgaacaaaca cactgaacag caggagtctc tagatcagaa attatttcaa 1800
ctacaagca aaaaatgtg gcttcaacag caattagtct atgcacataa gaaagtgcac 1860
aacaagaaga agataacaat tgatattcat ttcttgaga gaaaatgca acatcatctc 1920
ctaaaagaga aaaaatgaga gatatttcat tacaataacc atttaaaaaa cgttatatat 1980
caatatgaaa aagagaagc agaaacagaa aactcatgag agacaagcag taagaaactt 2040
cttttgaga aacaacagac cagatcttta ctcaaacctc atgtaggag gccagtctca 2100
gcatacctt atgttgaaaa ttattaccaat agtctgtgtc aacagaatac ttattttaga 2160
agaaaaatc atgatttctt cctgaagcct acagacataa aataacagtg tgaagaatta 2220
cttgttcacg aattgcataa agctgcacag gattcccatc taccctgatg atgcagcaga 2280
catcatcaca tcaaccaga atctgc 2307

```

<210> 25
 <211> 650
 <212> PRT
 <213> Homo sapiens

<220>
 <221> unsure
 <222> (310)
 <223> Xaa = Any Amino Acid
 <221> unsure
 <222> (429)
 <223> Xaa = Any Amino Acid
 <221> unsure
 <222> (522)
 <223> Xaa = Any Amino Acid

<400> 25
 Met Ser Pro Ala Lys Glu Thr Ser Glu Lys Phe Thr Trp Ala Ala Lys
 5 10 15

Gly Arg Pro Arg Lys Ile Ala Trp Glu Lys Lys Glu Thr Pro Val Lys

WO 01/75171

PCT/US01/10631

20

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

WO 01/75171

PCT/US01/10631

21

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

WO 01/75171

PCT/US01/10631

22

Lys Ile Ile Ser Tyr Met Lys Ile Ala Cys
645 650

<210> 26
<211> 228
<212> PRT
<213> Homo sapiens

<400> 26
Met Ser Pro Ala Lys Glu Thr Ser Glu Lys Phe Thr Trp Ala Ala Lys
5 10 15
Gly Arg Pro Arg Lys Ile Ala Trp Glu Lys Lys Glu Thr Pro Val Lys
20 25 30
Thr Gly Cys Val Ala Arg Val Thr Ser Asn Lys Thr Lys Val Leu Glu
35 40 45
Lys Gly Arg Ser Lys Met Ile Ala Cys Pro Thr Lys Glu Ser Ser Thr
50 55 60
Lys Ala Ser Ala Asn Asp Gln Arg Phe Pro Ser Glu Ser Lys Gln Glu
65 70 75 80
Glu Asp Glu Glu Tyr Ser Cys Asp Ser Arg Ser Leu Phe Glu Ser Ser
85 90 95
Ala Lys Ile Gln Val Cys Ile Pro Glu Ser Ile Tyr Gln Lys Val Met
100 105 110
Glu Ile Asn Arg Glu Val Glu Glu Pro Pro Lys Lys Pro Ser Ala Phe
115 120 125
Lys Pro Ala Ile Glu Met Gln Asn Ser Val Pro Asn Lys Ala Phe Glu
130 135 140
Leu Lys Asn Glu Gln Thr Leu Arg Ala Asp Pro Met Phe Pro Pro Glu
145 150 155 160
Ser Lys Gln Lys Asp Tyr Glu Glu Asn Ser Trp Asp Ser Glu Ser Leu
165 170 175
Cys Glu Thr Val Ser Gln Lys Asp Val Cys Leu Pro Lys Ala Thr His
180 185 190
Gln Lys Glu Ile Asp Lys Ile Asn Gly Lys Leu Glu Gly Lys Asn Arg
195 200 205
Phe Leu Phe Lys Asn Gln Leu Thr Glu Tyr Phe Ser Lys Leu Met Arg
210 215 220
Arg Asp Ile Leu
225

<210> 27
<211> 154

WO 01/75171

PCT/US01/10631

23

```

<212> PRT
<213> Homo sapiens

<220>
<221> unsure
<222> (148)
<223> Xaa = Any Amino Acid

<400> 27
Met Arg Leu His Pro Trp Arg Lys Glu His Leu Thr Gln Leu Lys Ala
      5              10              15

Trp Trp Lys Lys His Leu Met Arg Leu His Pro Trp Trp Lys Glu His
      20              25              30

Leu Thr Arg Leu Lys Ala Trp Trp Lys Lys His Leu Met Arg Leu His
      35              40              45

Pro Trp Trp Arg Glu His Leu Thr Lys Phe Asn Val Trp Arg Lys Arg
      50              55              60

His Leu Glu Ser Ser Asn Ser Gln Gln Lys Lys His Leu Gly Lys Leu
      65              70              75              80

Arg Val Leu Gln Lys Lys His Leu Arg Asn Leu Arg Gly Gln Gln Lys
      85              90              95

Glu Asp Leu Gly Arg Ser His Gly Arg Lys Lys Met Thr Gln Leu Arg
      100             105             110

Gln Lys Lys Lys Lys Lys Lys Lys Lys Lys Lys Lys Lys Lys Lys
      115             120             125

Lys Lys Lys Lys Lys Lys Lys Lys Lys Lys Lys Lys Lys Lys Lys
      130             135             140

Lys Lys Lys Xaa Lys Lys Lys Lys Lys
      145             150

<210> 28
<211> 466
<212> PRT
<213> Homo sapiens

<220>
<221> unsure
<222> (329)
<223> Xaa = Any Amino Acid

<400> 28
Met Ser Pro Ala Lys Glu Thr Ser Glu Lys Phe Thr Trp Ala Ala Lys
      5              10              15

Gly Arg Pro Arg Lys Ile Ala Trp Glu Lys Lys Glu Thr Pro Val Lys
      20              25              30

Thr Gly Cys Val Ala Arg Val Thr Ser Asn Lys Thr Lys Val Leu Glu

```

WO 01/75171

PCT/US01/10631

24

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

WO 01/75171

PCT/US01/10631

25

Gln Thr Phe Lys Ala Glu Pro Pro Glu Lys Pro Ser Ala Phe Glu Pro
355 360 365
Ala Ile Glu Met Gln Lys Ser Val Pro Asn Lys Ala Leu Glu Leu Lys
370 375 380
Asn Glu Gln Thr Leu Arg Ala Asp Glu Ile Leu Pro Ser Glu Ser Lys
385 390 395 400
Gln Lys Asp Tyr Glu Glu Ser Ser Trp Asp Ser Glu Ser Leu Cys Glu
405 410 415
Thr Val Ser Gln Lys Asp Val Cys Leu Pro Lys Ala Ala His Gln Lys
420 425 430
Glu Ile Asp Lys Ile Asn Gly Lys Leu Glu Gly Lys Asn Arg Phe Leu
435 440 445
Phe Lys Asn His Leu Thr Lys Tyr Phe Ser Lys Leu Met Arg Lys Asp
450 455 460
Ile Leu
465

<210> 29
<211> 445
<212> PRT
<213> Homo sapiens

<400> 29
Lys Glu Ile Asp Lys Ile Asn Gly Lys Leu Glu Gly Ser Pro Val Lys
5 10 15
Asp Gly Leu Leu Lys Ala Asn Cys Gly Met Lys Val Ser Ile Pro Thr
20 25 30
Lys Ala Leu Glu Leu Met Asp Met Gln Thr Phe Lys Ala Glu Pro Pro
35 40 45
Glu Lys Pro Ser Ala Phe Glu Pro Ala Ile Glu Met Gln Lys Ser Val
50 55 60
Pro Asn Lys Ala Leu Glu Lys Asn Glu Gln Thr Leu Arg Ala Asp
65 70 75 80
Glu Ile Leu Pro Ser Glu Ser Lys Gln Lys Asp Tyr Glu Glu Ser Ser
85 90 95
Trp Asp Ser Glu Ser Leu Cys Glu Thr Val Ser Gln Lys Asp Val Cys
100 105 110
Leu Pro Lys Ala Ala His Gln Lys Glu Ile Asp Lys Ile Asn Gly Lys
115 120 125
Leu Glu Glu Ser Pro Asp Asn Asp Gly Phe Leu Lys Ala Pro Cys Arg
130 135 140

WO 01/75171

PCT/US01/10631

26

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

WO 01/75171

PCT/US01/10631

27

```
<210> 30
<211> 578
<212> DNA
<213> Human

<400> 30
cttgctctct cttaggcttt gaagcatttt tgtctgtgct cctgatctt caggtoacca 60
ccatgaagtt cttagcagtc ctggactctt tgggagtttc catctttctg gtctctgccc 120
agaatccgac aacagctgct ccagctgaca cgtatccagc tactggtcct gctgatgatg 180
aagccctcga tgcctgaacc actgctgctg caaccactgc gaccactgct gctcctacca 240
ctgcaaccac cgtgcttctt accactgctc gtaagacat tccagtttta ccccaatggg 300
ttgggagctt ccgaatggg agagctgtgc cctgagatgg aatcagcttg agtctctctg 360
aatgggtcac aactattcat gtttctgtg atttcatcaa aactctaac ttgctctcga 420
tatccctctt atctctaato agtttatttt ctttcaata aaaataaact atgagcaaca 480
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 540
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 578

<210> 31
<211> 90
<212> PRT
<213> Homo sapien

<400> 31
Met Lys Phe Leu Ala Val Leu Val Leu Leu Gly Val Ser Ile Phe Leu 1
1 5
Val Ser Ala Gln Asn Pro Thr Thr Ala Ala Pro Ala Asp Thr Tyr Pro 15
20
Ala Thr Gly Pro Ala Asp Asp Glu Ala Pro Asp Ala Glu Thr Thr Ala 25
30
Ala Ala Thr Thr Ala Thr Thr Ala Ala Pro Thr Thr Ala Thr Thr Ala 35
40
Ala Ser Thr Thr Ala Arg Lys Asp Ile Pro Val Leu Pro Lys Trp Val 45
50
Gly Asp Leu Pro Asn Gly Arg Val Cys Pro 55
60
85
90

<210> 32
<211> 3101
<212> DNA
<213> Homo sapien

<400> 32
tgttggggcc tcagcctccc aagtagctgg gactacaggt gccggccacc acgcccagct 60
aattttttgt atatttttta gtagagacgg ggtttcaccg tgggtotcaat ctctgacct 120
cgtgatctgc cagccttggc ctcccaaggt gtattctctt tttattatta ttattatttt 180
tgagatggag tctgtctctg togccaggc tggagtgcag tgggtgcgac totgctcact 240
gcaagctccg cctcctgggt tcatgccatt ctctcgcttc agcctccoga gtactgggga 300
ctacaggccc ctgccaccac acccggttaa ttttttgtat ttttagtaga gacagggttt 360
cccatgttta gccagggtyg tctctatctt ctgaactcgt gatccgcttg cctcagcttc 420
tcaagtgctt gggattacag gctgagacca ccgggaccaa ccaactattg ctgtttattt 480
ttaaatataa tttaagaaaa caattagatt tgtttctctt ctcaattctt tactctaac 540
ctcatgttat gtataattat atttgtgttt tctattacct ttctccttt tactgtattg 600
gactataata attgtgctca ctaatttctg ttcaataata ttatcagott agataatact 660
ttaattttta acttatataa tgagtattaa attgacagtt tttatttgta attatctatc 720
ttcgccttgg ctgaatataa cttcttaagc ttataacttc ttgtttcttc catgttattt 780
ttttcttttt tttaattgat tgaatttctt ctgacaactca ttctagtaac tttttctcgc 840
gtgtgcaacg taagtataa tttgtttctc agatttgaga tctgccataa gtttgaggct 900
ttattttttt tttttatttg ctttatggca agtcggacaa cctgcattgga tttggcatca 960
```


WO 01/75171

PCT/US01/10631

28

```

atgtagtcac ccatatctaa gacgagcact tgcctcttag catgatgagt tgtttctgga 1020
ttgtttcttt attttactta ttttctgtgt agattcttat attttccctt aaactctatt 1080
cagcatttta ggaattctta ggactttctg agaattttag cttttctgtat taaatgtttt 1140
taatgagtat tgcattttct caaagacac aaatatcaat agtgtacaca tgaaggaaac 1200
tatatatata tctgtttgca gatgacagca tctcataaca aaatcctagt tacttcattt 1260
aaaagacagc tctcctccaa tatactatga ggtaacaaaa atttgtatgt tgtaattttt 1320
ttaatattag aaactctato ttacattgtg cacaaatttc tgaagtata atacttcaat 1380
gtttttctat agaagtaact taatattggc aaaattactt atttgaattt aggttttggc 1440
tttcatcata tacttctca ttaacatttc cctcaatcca taaatgcaat ctcatgttga 1500
atcttccatt taacccagaa gtttaatttt aaaaccttaa taaatgtga atgtagctag 1560
atattatttg ttggttaaat attagtcaat aatttatatt acttacaatg atcagaaaa 1620
atgatctgaa ttctctgtgt cataaattca ataacttatt ttaggcctaa accttccat 1680
ttcaatctt ttggtctgtt aattgaaat aatcattatc ttgtgtttt ttggccaaaa 1740
tgctgcccat ttatttctat ccttaattag tcaaaatttc taaataaagt atttaacgtt 1800
aatgatgttt atttctgtgt tgtatactaa aaccttagt ttctataatt taaatgtcac 1860
ctaatatgag tgaaaatgtg tcagagcgtg ggaagaaatg tggatggaga aagggaaagt 1920
gttgatcaaa agtaccctaa gtctcagtta cacagaggcg atgagattga tctagtccaa 1980
aaaatgatga gtataataaa taataatgca ctgtatat ttgaaattgtc aaaagttagt 2040
ttaaaattga ttacataaat attttacata ttataaaagc acatgcaata tgttgttaca 2100
tgtatagaat gtgcaacgat caagtccagg tatctgtgtt atccaccact ttgagcattt 2160
atgatctcta tatgtcagga acatttcaag ttatctgttc tagcaaggaa atataaaaa 2220
cattatagtt aactatggcc tatctacagt gcaactaaac actagatttt attcctttcc 2280
aactgtgggt ttgtattcat ttaccaacct cttttcattc cttttctcac ccacacactg 2340
tgccggggct caggcatata ctattctact gtctgtctct gtaaggatta tcaatttagc 2400
ttccacatat gagagaatgc atgcaaaagt ttcttttcca tgtctggott atttcaatta 2460
acaaatgao ctccgtctcc atccatgta ttcatattac ccaatagtg tcaataaat 2520
atatacacac atataacca cattgcattt gbaaattat tcaatgaag aaactggita 2580
atgttatatc gttgtattg tgaatagtgc tgcataaac acgcaagtgc ggaataaat 2640
tgaagagttt ttgtgtgat gtltacata aattttaaga ttgttttgc tatgtttgtg 2700
aaatggcgt tagtatttcc atagagattg cattgaatct gtatagttct ttgggttaagt 2760
atggttattt tgaatgtatt aatttttcca ttccatgaag atgagatgct ttccatttg 2820
tttgtgtct ctacatttcc ttccatcaaa gttttgtgtt atttttgaag tagatgtatt 2880
tcacattata gatcaagtgt attccctaaa tatttttatt ttgtagctat tgtatgtaa 2940
attgctctct cgatttcttt ttcaactaat tcaatttagt tgtatgaaa tgttatggat 3000
ttttatttgt tggtttttaa tcaaaaactg tattaacctt agagtttttt gtggagtttt 3060
taagtttttc tagatataag atcatgacat ctacaaaaaa a 3101

```

<210> 33
 <211> 16
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> PCR primer

<400> 33
 tgccctccg gaagct 16

<210> 34
 <211> 23
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> PCR primer

WO 01/75171

PCT/US01/10631

29

<400> 34 cgtttctgaa gggacatctg atc	23
<210> 35 <211> 30 <212> DNA <213> Artificial Sequence	
<220> <223> PCR primer	
<400> 35 ttgcagccaa gttaggagtg aagagatgca	30
<210> 36 <211> 24 <212> DNA <213> Artificial Sequence	
<220> <223> PCR primer	
<400> 36 aagcctcaga gtccttcag tatg	24
<210> 37 <211> 35 <212> DNA <213> Artificial Sequence	
<220> <223> PCR primer	
<400> 37 ttcaaatata agtgaagaaa aaattagtag atcaa	35
<210> 38 <211> 37 <212> DNA <213> Artificial Sequence	
<220> <223> PCR primer	
<400> 38 aatccattgt atcttagaac cgagggattt gtttaga	37
<210> 39 <211> 22 <212> DNA <213> Artificial Sequence	
<220> <223> PCR primer	
<400> 39 aaagcagatg gtggttgagg tt	22

WO 01/75171

PCT/US01/10631

30

<210> 40
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR primer

<400> 40
octgagacca aatggottct to 22

<210> 41
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR primer

<400> 41
attccatgcc ggctgcttct tctg 24

<210> 42
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR primer

<400> 42
tctggtttcc tcattcttta ttcatttatt 30

<210> 43
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR primer

<400> 43
tgccaaggag cggattatct 20

<210> 44
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR primer

<400> 44
caaccaagtg acaaacactg gaattacagg 30

<210> 45
<211> 21
<212> DNA

WO 01/75171

PCT/US01/10631

31

<213> Artificial Sequence
<220>
<223> PCR primer
<400> 45
actggaacgg tgaaggtgac a 21
<210> 46
<211> 20
<212> DNA
<213> Artificial Sequence
<220>
<223> PCR primer
<400> 46
cgccacatt gtgaacttg 20
<210> 47
<211> 23
<212> DNA
<213> Artificial Sequence
<220>
<223> PCR primer
<400> 47
cagtcggttg gagcgagcat occ 23
<210> 48
<211> 24
<212> DNA
<213> Artificial Sequence
<220>
<223> PCR primer
<400> 48
tgccatagat gaattgaagg aatg 24
<210> 49
<211> 29
<212> DNA
<213> Artificial Sequence
<220>
<223> PCR primer
<400> 49
tgtcatatat taattgcata aacacctca 29
<210> 50
<211> 32
<212> DNA
<213> Artificial Sequence
<220>

WO 01/75171

PCT/US01/10631

32

<223> PCR primer
<400> 50
tcttaaccaa acggatgaaa ctctgagcaa tg 32
<210> 51
<211> 28
<212> DNA
<213> Artificial Sequence
<220>
<223> PCR primer
<400> 51
atcattgaaa attcaaatat aagtgaag 28
<210> 52
<211> 30
<212> DNA
<213> Artificial Sequence
<220>
<223> PCR primer
<400> 52
gtagttgtgc attgaaataa ttatcattat 30
<210> 53
<211> 20
<212> DNA
<213> Artificial Sequence
<220>
<223> PCR Primer
<400> 53
caattttggt ggagaacccg 20
<210> 54
<211> 20
<212> DNA
<213> Artificial Sequence
<220>
<223> PCR Primer
<400> 54
gctgtcggag gtatatggtg 20
<210> 55
<211> 28
<212> DNA
<213> Artificial Sequence
<220>
<223> PCR Primer

WO 01/75171

PCT/US01/10631

33

<400> 55 catttcagag agtaacatgg actacaca	28
<210> 56 <211> 21 <212> DNA <213> Artificial Sequence	
<220> <223> PCR Primer	
<400> 56 tctgataaag gcgtacaat g	21
<210> 57 <211> 22 <212> DNA <213> Artificial Sequence	
<220> <223> PCR Primer	
<400> 57 tcacgacttg ctgtttttgc tc	22
<210> 58 <211> 30 <212> DNA <213> Artificial Sequence	
<220> <223> PCR Primer	
<400> 58 atcaaaaaac aagcatggcc tcacaccact	30
<210> 59 <211> 21 <212> DNA <213> Artificial Sequence	
<220> <223> PCR Primer	
<400> 59 gcaagtgccca atgatcagag g	21
<210> 60 <211> 23 <212> DNA <213> Artificial Sequence	
<220> <223> PCR Primer	
<400> 60 atatagactc aggtatacac act	23

WO 01/75171

PCT/US01/10631

34

<210> 61
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR Primer

<400> 61
tcccatcaga atccaaacaa gaggaagatg 30

<210> 62
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR Primer

<400> 62
aatccattgt atcttagaac cgagggatgt gttt 34

<210> 63
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR Primer

<400> 63
cagcttctga caacactaga gatc 24

<210> 64
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR Primer

<400> 64
octataaaga ttttatgtac caaaaatgaa gt 32

<210> 65
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR Primer

<400> 65
ccctccctc aggtatggc cc 22

<210> 66
<211> 22

WO 01/75171

PCT/US01/10631

35

<212> DNA
<213> Artificial Sequence
<220>
<223> PCR Primer
<400> 66
ccctttctca cccacacact gt 22
<210> 67
<211> 24
<212> DNA
<213> Artificial Sequence
<220>
<223> PCR Primer
<400> 67
tgcattctct catatgtgga agct 24
<210> 68
<211> 33
<212> DNA
<213> Artificial Sequence
<220>
<223> PCR Primer
<400> 68
ccgggcctca ggcataact attctactgt ctg 33
<210> 69
<211> 24
<212> DNA
<213> Artificial Sequence
<220>
<223> PCR Primer
<400> 69
gacattccag ttttaccacaa atgg 24
<210> 70
<211> 23
<212> DNA
<213> Artificial Sequence
<220>
<223> PCR Primer
<400> 70
tgcagaagac tcaagctgat tcc 23
<210> 71
<211> 28
<212> DNA
<213> Artificial Sequence

WO 01/75171

PCT/US01/10631

36

```

<220>
<223> PCR Primer

<400> 71
tctcaggagc acactctacc attcggga      28

<210> 72
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR Primer

<400> 72
aaatataagt gaagaaaaaa attagtagat      30

<210> 73
<211> 503
<212> DNA
<213> Homo sapiens

<400> 73
gacagcggct tcttgatcc ttgccaccg cgaactgaac ccgacagcag cagctcacc      60
atgaagttgc tgaigtgctt catgtggcg gccctctccc apcactgcta cgcaggctct      120
ggctgcccc ttatggagaa tgtgatttcc aagacaatcc atcccaagt gtctaagact      180
gaatacaaaag aaatttctca agagtctcta gaagacaatg ccactacaaa tgcctatagat      240
gaattgaagc aatgttttct taaccacaac gatgaactc tgaacaatgt tgaggtgttt      300
ctgcaattaa tatatgacag cagtcittgt gatttatttt aactttctgc aagacctttg      360
gtccacagaa ctgcagggtta tggtagagaa ccaactacgg attgctgcaa accacacctt      420
ctctttctta tgtcttttta ctacaaaacta caagacaatt gttgaaacct gctatacatg      480
tttatttttaa taaattgatg gca      503

<210> 74
<211> 301
<212> DNA
<213> Homo sapiens

<400> 74
cactgctacg caggctctgg ctgcccttta ttggagaatg tgatttccaa gacaatcaat      60
ccacaagtgt ctaagactga atcaaaagaa ctcttcaag agttcataga cgaatgccc      120
actacaaaag ccatagatga attgaagaa tgttttctta accaaaaggga tgaacctctg      180
agcaatgttg aggtgtttat gcaatttaata tatgacagca gtctttgtga tttatttggc      240
ggccatcacc atcaccatca ctaaggtccc gagctcgaat tctgcagata tccatcacac      300
t      301

<210> 75
<211> 3282
<212> DNA
<213> Homo sapiens

<400> 75
gggacagggc tgaggatgag gagaaccctg gggaccocga agaccgtgcc ttgcccgga      60
gtcctgctgc taggcctgaa ggacttgccc taacagagcc toaacaacta cctggtgatt      120
cctacttcag ccccttggtg tgagcagctt ctcaacatga actacagcct ccacttgccc      180
ttcgtgtgtc tgagtctctt caotgagagg atgtgcatcc aggggagtcg gtccaacgtc      240
gaggtcggca gaagtgcaca gctttccctg cctggctttg agaaacctac agcaggatat      300
aacaatttc tcaggcccaa ttttggtgga gaaccctac agatagcgct gactctggac      360

```

WO 01/75171

PCT/US01/10631

37

```

attgcaagta tctctagcat ttcagagagt aacatggact acacagccac catataoctc 420
cgacagcgct ggatggacca ggggctgggt ttggaaggca acaagagctt cactctggat 480
gcccgcctcg tggagttcct ctgggtgcca gatactfaca ttgtggatc caagaagtcc 540
ttcctccatg aagtcactgt gggaaacagg ctcatccgcc tctctoccaa tggcaccgtc 600
ctgtatgccc tcagaatcac gacaactgtt goactgaaca tggatctgtc taaatccccc 660
atggacacac agacatgcaa gttgcagctg gaaagctggg gctatgatgg aatgatgtg 720
gagttcacct ggctgagagg gaacgactct gtgctgggac tggaaacacct gcggcttgct 780
cagtacacca tagaggggta ttccacctta gtcaccagat cgcagcagga gacaggaat 840
tacctagat tggctctaca gtttgagctt cggaggaatg ttctgtattt cattttgaa 900
acctacgttc ctccacttt cctgggtggt ttgtcctggg ttctattttg gatctctctc 960
gattcagttc ctgcaagaac ctgcattgga gtgacgacgg tgttatcaat gaccacactg 1020
atgatcggtt cccgcacttc tcttcccaac accaactgct tcatcaaggc catgatgtg 1080
tacctgggga tctgctttag ctctgtgttt ggggcttgc tagaatatgc agttgtcac 1140
taactgtcct tacagagat ggcagccaaa gataggggga caacaaggga agtagaagaa 1200
gtcagttatta ctaatacat caacagctcc atctccagct ttaaacggaa gatcagcttt 1260
ggcagcattg aaatttccag cgacaaogtt gactacagt acttgacaat gaaaacacgg 1320
gacaagtcca agtttgtctt ccgagaaaaa atgggocagga ttgttgatta ttccacaatt 1380
caaaacccca gtaatgttga tcaactattc aaactactgt tctcttggat ttttatgcta 1440
gccaatgtat tttactgggc atactacatg tatttttgag toaatgttaa atttcttgca 1500
tgccataggt ctccaacagg acaagataat gatgtaaat gtatttttag ccaagtgtgc 1560
accacatccc aatggtgcta caagtactg aaataatatt tgagtctttc tgcataaga 1620
atgaagctcc aaccattgtt ctaagtctgt tagaagtct agcattatag gatcttgyaa 1680
tagaaacatc agtccattcc tctttcatct taatcaagga cattcccatg gagcccaaga 1740
ttcaaatgtt actcagggct gtttattcgg tggctccctg gtttgcaatt acctatata 1800
aagaatggga aggagacatc tgggtaaccc tcaagtgtca gaagtgtttt cttaaagtaac 1860
tatcatgttt ttttactaaa tctctgcagt gcttataaaa tacttggttg cctatttagg 1920
gagtaaacatt ttctagtttt tgtttctgtg taaaatgaaa tatgggctta tgtcaattca 1980
ttggaagtca atgcactaac tcaatcccaa gatgagtttt taaataatga atattattta 2040
ataccaaac agaatatcc ocaatttcca ataagtctta tcaatgaaa ttcaaatata 2100
agtgaagaaa aaattagtag atcaacaatc taaacaatc cctcgttctt aagatacaat 2160
ggattcccca tactggaagg actctgaggg tttattccca cactatgcat atcttatcat 2220
tttattatta tacacacatc catctaaac tatactaaag cctttttccc atgcatggat 2280
ggaaatggaa gatttttttg taactgttc tagaagtctt aatatgggct gttgcaatga 2340
aggcttgcaa aattgactcc atttcttagc tgcctttatt cacatagtga tggggtacta 2400
aaagtactgg gttgactcag agagtctgtg tcatctgtc attgctgcta ctctaacatc 2460
gagcaacact ctccagtggt cagatccccc gtatcatccc aagaggagca tcatccctt 2520
tgctctaatg atcaggaatg atgcttatta gaaaacaaac tggttgaccc aggaacaaatg 2580
ggcttagctt aagtaaacctt ggccttgctc agatccctga tccctccagg tggctgctc 2640
tgagttgctt atcccgatg agcagagagg tgcctggcct gagtaactga ctctctgagt 2700
aacaatgaga caggttacag aacctatgtt caggttgagg gtgagctgcc ctctocaaat 2760
ccagccagag atgcacatc ctccggcagt ctacgccaac agtacccaaa gtgattttt 2820
agtggtccag ggtaaaggct tccagttcag cctcagttat tttagacaaat ctcccatct 2880
ttaatttctt agcttccgtt tctaataaat gcacggcttt acctttctg tcaaaaataa 2940
accaaggctc taaaagatga ttcccttctt gtaactccct agagccacag gttctcatc 3000
cttttcccat tatacttctc acaactcagt ttctatgagt ttgatccact gatttttta 3060
acaaaatatt tctaacggga atgggtggga gtgctggtga aaagagatga aatgtggtg 3120
tatgagccaa tcatatttgt gattttttaa aaaaagtta aaaggaataa tctgtctctg 3180
aaccccaatt aagcattgtt ttatatataa acaaatgata aagatgtgaa ctgtgaaata 3240
aatataccat attagctaco ccccaaaaaa aaaaaaaa aa 3282

```

<210> 76

<211> 463

<212> DNA

<213> Homo sapiens

<400> 76

```

tagaattcag oggcogotta attotagaag tccaaatcac tcaattgttg tgaaagotga 60
gctcacagca aaacaagcca ccatgaagct gtcggtgtgt ctctctgtgg tcaocgtggc 120

```

WO 01/75171

PCT/US01/10631

38

```

cctctgtctgc taccaggcca atgccgagtt ctgcccagct cttgtttctg agctgttaga 180
cttcttcttc attagtgaac ctctgttcaa gtttagctct gccaaatttg atgcccctoc 240
ggaagctggtt gcagccaagt taggsagtga gagatgcac gatcagatgt ccttcagaa 300
acgaagcctc attgcggaag tcctggtgaa aatattgaag aatgtatgt tgtgacatgt 360
aaaaacttcc atctgtgttt ccactgtctt tcaatgacac cctgatcttc actgcagaat 420
gtaaagggtt caactcttg ctttaataaa tcacttgctc tac 463

```

<210> 77

<211> 90

<212> PRT

<213> Homo sapiens

<400> 77

```

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

```

【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
11 October 2001 (11.10.2001)

PCT

(10) International Publication Number
WO 01/075171 A3(51) International Patent Classification: C12Q 1/68, (74) Agents: POTTER, Jane, E., R. et al.; Seed Intellectual
G01N 33/543, 33/574 // C12N 5/08 Property Law Group PLLC, Suite 6300, 701 Fifth Avenue,
Seattle, WA 98104-7092 (US).

(21) International Application Number: PCT/US01/10631

(22) International Filing Date: 2 April 2001 (02.04.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/194,241 3 April 2000 (03.04.2000) US
60/219,862 20 July 2000 (20.07.2000) US
60/221,300 27 July 2000 (27.07.2000) US
60/256,592 18 December 2000 (18.12.2000) US(81) Designated States (national): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,
CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GH, GI, GM,
HR, HU, ID, IL, IN, IS, JP, KG, KP, KR, KZ, LC, LK,
LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX,
MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL,
TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.(84) Designated States (regional): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,
IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF,
CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).(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): HOUGHTON, Ray-
mond, L. [US/US]; 2636 242nd Place S.E., Bothell, WA
98021 (US). DILLON, Davin, C. [US/US]; 18112 N.W.
Montreux Drive, Issaquah, WA 98027 (US). MOLESH,
David, Alan [US/US]; 12385 N.E. Klabo Road, Kingston,
WA 98346 (US). XU, Jiangchun [US/US]; 15806 S.E.
43rd Place, Bellevue, WA 98006 (US). ZEHENTNER,
Barbara [DE/US]; 4420 Bluff Lane, Bainbridge Island,
WA 98110 (US). PERSING, David, H. [US/US]; 22401
N.E. 25th Way, Redmond, WA 98053 (US).

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:
8 August 2002For two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.

WO 01/075171 A3

(54) Title: METHODS, COMPOSITIONS AND KITS FOR THE DETECTION AND MONITORING OF BREAST CANCER

(57) Abstract: Compositions and methods for the therapy and diagnosis of cancer, such as breast cancer, are disclosed. Composi-
tions may comprise one or more breast tumor proteins, immunogenic portions thereof, or polynucleotides that encode such portions.
Alternatively, a therapeutic composition may comprise an antigen presenting cell that expresses a breast tumor protein, or a T cell
that is specific for cells expressing such a protein. Such compositions may be used, for example, for the prevention and treatment of
diseases such as breast cancer. Diagnostic methods based on detecting a breast tumor protein, or mRNA encoding such a protein, in
a sample are also provided.

【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
11 October 2001 (11.10.2001)

PCT

(10) International Publication Number
WO 01/075171 A3

(51) International Patent Classification: C12Q 1/68, G01N 33/543, 33/574 // C12N 5/08 (74) Agents: POTTER, Jane, E., R. et al.; Seed Intellectual Property Law Group PLLC, Suite 6300, 701 Fifth Avenue, Seattle, WA 98104-7092 (US).

(21) International Application Number: PCT/US01/10631

(22) International Filing Date: 2 April 2001 (02.04.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
50/194,241 3 April 2000 (03.04.2000) US
60/219,852 20 July 2000 (20.07.2000) US
60/221,300 27 July 2000 (27.07.2000) US
60/256,592 18 December 2000 (18.12.2000) US

(81) Designated States (national): AF, AG, AI, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GI, GM, HR, HU, ID, IL, IN, IS, JP, KI, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MY, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

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

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

Published:
with international search report
with amended claims(72) Inventors; and
(75) Inventors/Applicants (for US only): HOUGHTON, Raymond, L. [US/US]; 2636 242nd Place S.E., Bothell, WA 98021 (US). DILLON, Davin, C. [US/US]; 18112 N.W. Montreux Drive, Issaquah, WA 98027 (US). MOLESH, David, Alan [US/US]; 12385 N.E. Klaho Road, Kingston, WA 98346 (US). XU, Jiangchun [US/US]; 15805 S.L. 43rd Place, Bellevue, WA 98006 (US). ZEHENTNER, Barbara [DE/US]; 4420 Bluff Lane, Bainbridge Island, WA 98110 (US). PERSING, David, H. [US/US]; 22401 N.E. 25th Way, Redmond, WA 98053 (US).

(88) Date of publication of the international search report: 8 August 2002

Date of publication of the amended claims: 6 February 2003

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 01/075171 A3

(54) Title: METHODS, COMPOSITIONS AND KITS FOR THE DETECTION AND MONITORING OF BREAST CANCER

(57) Abstract: Compositions and methods for the therapy and diagnosis of cancer, such as breast cancer, are disclosed. Compositions may comprise one or more breast tumor proteins, immunogenic portions thereof, or polynucleotides that encode such portions. Alternatively, a therapeutic composition may comprise an antigen presenting cell that expresses a breast tumor protein, or a T cell that is specific for cells expressing such a protein. Such compositions may be used, for example, for the prevention and treatment of diseases such as breast cancer. Diagnostic methods based on detecting a breast tumor protein, or mRNA encoding such a protein, in a sample are also provided.

【国際調査報告】

INTERNATIONAL SEARCH REPORT		International Application No. PCT/US 01/10631
A. CLASSIFICATION OF SUBJECT MATTER IPC 7 C12Q1/68 G01N33/543 G01N33/574 //C12N5/08		
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 C12Q G01N C12N C07K		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practical, search terms used)		
BIOSIS, MEDLINE		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 98 45328 A (CORIXA CORP) 15 October 1998 (1998-10-15) SEC ID NO: 292,296; ---	37-46
X	XU ET AL: "Identification of differentially expressed genes in human breast tumor using subtraction and microarray" CHEMICAL ABSTRACTS + INDEXES, AMERICAN CHEMICAL SOCIETY, COLUMBUS, US, vol. 40, March 1999 (1999-03), page 319 XP002160220 ISSN: 0009-2258 abstract --- -/--	37-46
A		1-36
<input checked="" type="checkbox"/> Further documents are listed in the continuation of box C. <input checked="" type="checkbox"/> Patent family members are listed in annex.		
* Special categories of cited documents: *A* document defining the general state of the art which is not considered to be of particular relevance *E* earlier document but published on or after the international filing date *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) *O* document referring to an oral disclosure, use, exhibition or other means *P* document published prior to the international filing date but later than the priority date claimed *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art *Z* document member of the same patent family		
Date of the actual completion of the international search		Date of mailing of the international search report
15 May 2002		31/05/2002
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentstraat 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax. (+31-70) 340-3016		Authorized officer van Klompenburg, W

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

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*		Relevant to claim No.
X	XU JIANGCHUN ET AL: "Identification of differentially expressed genes in human prostate cancer using subtraction and microarray." CANCER RESEARCH, vol. 60, no. 6, 15 March 2000 (2000-03-15), pages 1677-1682, XP002198982 ISSN: 0008-5472 the whole document	1-3,6,7, 9,14,17
A	---	
A	WO 98 04742 A (GENZYME CORP ;BINETTE FRANCOIS (US); HAUDENSCHILD DOMINIK (US)) 5 February 1998 (1998-02-05) page 4 -page 5; claims 1-15 ---	1-36
T	JIANG YUQIU ET AL: "Discovery of differentially expressed genes in human breast cancer using subtracted cDNA libraries and cDNA microarrays." ONCOGENE, vol. 21, no. 14, 2002, pages 2270-2282, XP001075060 28 March, 2002 ISSN: 0950-9232 the whole document	1-46

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

INTERNATIONAL SEARCH REPORT
Information on patent family members

 International Application No.
PCT/US 01/10631

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
WO 9845328	A	15-10-1998	US 6225054 B1	01-05-2001
			AU 6956098 A	30-10-1998
			BR 9808509 A	18-09-2001
			EP 1127893 A2	29-08-2001
			EP 0975666 A2	02-02-2000
			JP 2001521384 T	06-11-2001
			NO 994932 A	07-12-1999
			PL 336349 A1	19-06-2000
			TR 9903154 T2	21-08-2000
			WO 9845328 A2	15-10-1998
			ZA 9802968 A	27-10-1998
			HU 0001270 A2	28-07-2000
			US 6344550 B1	05-02-2002
WO 9804742	A	05-02-1998	AU 3895597 A	20-02-1998
			EP 0956364 A1	17-11-1999
			JP 2000515761 T	28-11-2000
			WO 9804742 A1	05-02-1998

フロントページの続き

(51)Int.Cl.⁷ F I テーマコード(参考)
G 0 1 N 33/58 G 0 1 N 33/58 A

(31)優先権主張番号 60/256,592

(32)優先日 平成12年12月18日(2000.12.18)

(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,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 ディロン, ダビン シー.
アメリカ合衆国 ワシントン 9 8 0 2 7, イザクア, エヌ.ダブリュー. モントリュー
ドライブ 1 8 1 1 2

(72)発明者 モレシュ, デイビッド アラン
アメリカ合衆国 ワシントン 9 8 3 4 6, キングストン, エヌ.イー. クラボ ロード
1 2 3 8 5

(72)発明者 スウ, ジャンチュン
アメリカ合衆国 ワシントン 9 8 0 0 6, ベルビュー, エス.イー. 4 3 アールディー
プレイス 1 5 8 0 5

(72)発明者 ツェーントナー, バーバラ
アメリカ合衆国 ワシントン 9 8 1 1 0, ベインブリッジ アイランド, ブラフ レーン
4 4 2 0

(72)発明者 パーシング, デイビッド エイチ.
アメリカ合衆国 ワシントン 9 8 0 5 3, レッドモンド, エヌ.イー. 2 5 ティーエイチ
ウェイ 2 2 4 0 1

F ターム(参考) 2G045 AA26 BA13 BB03 BB20 CA25 CA26 CB02 CB03 CB08 DA12
DA13 DA14 DA36 DA78 FA36 FB02 FB07 FB12
4B024 AA12 CA01 HA12
4B063 QA19 QQ08 QQ43 QR08 QR42 QR56 QS25 QS34 QX02 QX07

专利名称(译)	用于检测和监测乳腺癌的方法，组合物和试剂盒		
公开(公告)号	JP2004505611A	公开(公告)日	2004-02-26
申请号	JP2001573043	申请日	2001-04-02
[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	ヒュートンレイモンドエル ディロンダビンシー モレシュデイビッドアラン スウジャンチュン ツェントナーバーバラ パーシングデイビッドエイチ		
发明人	ヒュートン, レイモンド エル. ディロン, ダビン シー. モレシュ, デイビッド アラン スウ, ジャンチュン ツェントナー, バーバラ パーシング, デイビッド エイチ.		
IPC分类号	G01N33/53 C12N5/08 C12N15/09 C12Q1/68 G01N33/566 G01N33/574 G01N33/58		
CPC分类号	C12Q1/6844 C12Q1/6809 C12Q1/6851 C12Q1/6886 C12Q2600/16		
FI分类号	C12N15/00.ZNA.A C12Q1/68.A G01N33/53.M G01N33/566 G01N33/574.Z G01N33/58.A		
F-TERM分类号	2G045/AA26 2G045/BA13 2G045/BB03 2G045/BB20 2G045/CA25 2G045/CA26 2G045/CB02 2G045/CB03 2G045/CB08 2G045/DA12 2G045/DA13 2G045/DA14 2G045/DA36 2G045/DA78 2G045/FA36 2G045/FB02 2G045/FB07 2G045/FB12 4B024/AA12 4B024/CA01 4B024/HA12 4B063/QA19 4B063/QQ08 4B063/QQ43 4B063/QR08 4B063/QR42 4B063/QR56 4B063/QS25 4B063/QS34 4B063/QX02 4B063/QX07		
代理人(译)	夏木森下		
优先权	60/194241 2000-04-03 US 60/219862 2000-07-20 US 60/221300 2000-07-27 US 60/256592 2000-12-18 US		
外部链接	Espacenet		

摘要(译)

公开了用于治疗 and 诊断癌症（例如乳腺癌）的组合物和方法。组合物可包含一种或多种乳腺肿瘤蛋白，其免疫原性部分或编码这些部分的多核苷酸。或者，治疗组合物可包含表达乳腺肿瘤蛋白的抗原呈递细胞，或对表达这种蛋白的细胞特异的T细胞。此类组合物可用于例如预防和治疗诸如乳腺癌的疾病。还提供了基于在样品中检测乳腺肿瘤蛋白或编码这种蛋白的mRNA的诊断方法。

