

(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2008-501332

(P2008-501332A)

(43) 公表日 平成20年1月24日(2008.1.24)

(51) Int. Cl.	F I	テーマコード (参考)
<b>C12Q 1/68 (2006.01)</b>	C12Q 1/68 ZNAA	4B024
<b>C12Q 1/02 (2006.01)</b>	C12Q 1/02	4B063
<b>C12N 15/09 (2006.01)</b>	C12N 15/00 F	
<b>GO1N 33/574 (2006.01)</b>	C12N 15/00 A	
<b>GO1N 33/53 (2006.01)</b>	GO1N 33/574 A	

審査請求 未請求 予備審査請求 未請求 (全 80 頁) 最終頁に続く

(21) 出願番号	特願2007-514130 (P2007-514130)	(71) 出願人	506402562 ダイアゲニック エイエス ノルウェー エヌー0661 オスロ オ ステンスジョヴァイエン 15ビー
(86) (22) 出願日	平成17年6月2日(2005.6.2)	(74) 代理人	100081994 弁理士 鈴木 俊一郎
(85) 翻訳文提出日	平成19年2月5日(2007.2.5)	(74) 代理人	100103218 弁理士 牧村 浩次
(86) 国際出願番号	PCT/GB2005/002180	(74) 代理人	100115392 弁理士 八本 佳子
(87) 国際公開番号	W02005/118851	(72) 発明者	シャルマ ブラヴェーン ノルウェー エヌー0173 オスロ ヴ ァルデマーズ トランス ゲート 62 ディー
(87) 国際公開日	平成17年12月15日(2005.12.15)		
(31) 優先権主張番号	0412301.4		
(32) 優先日	平成16年6月2日(2004.6.2)		
(33) 優先権主張国	英国 (GB)		

最終頁に続く

(54) 【発明の名称】 癌診断のためのオリゴヌクレオチド

## (57) 【要約】

本発明は、分析法、特に診断法に使用することができる、細胞中の遺伝子転写レベルの評価用オリゴヌクレオチドプローブに関する。本発明は、タンパク質合成及び/又は安定性に関与するタンパク質をコードする遺伝子、又は特に癌診断用転写パターンを作成するのに使用される防御の調節及び/又はクロマチンリモデリングに関与するタンパク質をコードする遺伝子に相当するオリゴヌクレオチドセットを提供する。そうしたプローブは、キットの形態で提供されるのが都合がよい。異なる型のプローブを、遺伝子発現パターンの作成、及び異なる癌又は癌の病期の同定、診断又はモニタリングの技術に使用できる。また、本発明は、このようなセット及びこのようなセットを含むキット、ならびに特徴的な発現プロファイルを得るための前記遺伝子によりコードされているマーカーポリペプチドの分析を用いた関係方法を提供する。

## 【特許請求の範囲】

## 【請求項 1】

生物における癌又は癌の病期に特徴的な標準遺伝子転写パターンを作成する方法であって、少なくとも：

a) 必要に応じて cDNA に逆転写してもよい、その癌又は癌の病期を有する一種以上の生物の試料の細胞から mRNA を単離する工程と；

b) 工程 (a) の mRNA 又は cDNA を、調査中の生物及び試料に相当する生物及び試料中の該癌又は癌の病期に特異的な一組のオリゴヌクレオチドプローブにハイブリダイズさせる工程であって、前記の一組が少なくとも 10 個のオリゴヌクレオチドを含み、各オリゴヌクレオチドが、

タンパク質合成及び / 又は安定性に関連しているタンパク質をコードしているファミリー (i) 遺伝子；又は

防御及び / 又はクロマチンリモデリングの調節に関与しているタンパク質をコードしているファミリー (ii) 遺伝子；からの遺伝子配列に相当するオリゴヌクレオチドから選択されたものであるか、

又はそのような配列に由来するものであるか、又は相補的配列を有するオリゴヌクレオチドか、又は機能的に等価のオリゴヌクレオチドである、工程と；

c) 該プローブの各々にハイブリダイズしている mRNA 又は cDNA の量を評価して、該癌又は癌の病期についての該試料において、該オリゴヌクレオチドが結合する遺伝子の遺伝子発現のレベルを示す特徴的なパターンを作成する工程と；

を含む方法。

## 【請求項 2】

試験遺伝子転写パターンの作成方法であって、少なくとも：

a) 必要に応じて cDNA に逆転写してもよい、前記試験生物の試料の細胞から mRNA を単離する工程と；

b) 工程 (a) の mRNA 又は cDNA を、調査中の生物及び試料に相当する生物及び試料中の該癌又は癌の病期に特異的な請求項 1 に記載の一組のオリゴヌクレオチドにハイブリダイズさせる工程と；

c) 前記プローブの各々にハイブリダイズしている mRNA 又は cDNA の量を評価して、該試験試料において、該オリゴヌクレオチドが結合する遺伝子の遺伝子発現のレベルを示す前記パターンを作成する工程と；

を含む方法。

## 【請求項 3】

生物における癌又は癌の病期を診断又は同定又はモニタリングする方法であって：

a) 必要に応じて cDNA に逆転写してもよい、該生物の試料の細胞から mRNA を単離する工程と；

b) 工程 (a) の mRNA 又は cDNA を、調査中の生物及び試料に相当する生物及び試料中の該癌又は癌の病期に特異的な請求項 1 に記載の一組のオリゴヌクレオチドにハイブリダイズさせる工程と；

c) 前記プローブの各々にハイブリダイズしている mRNA 又は cDNA の量を評価して、前記試験試料において、該オリゴヌクレオチドが結合する遺伝子の遺伝子発現のレベルを示す特徴的なパターンを作成する工程と；

d) 前記パターンを、前記調査中の生物及び試料に相当する生物からの試料を用いて請求項 1 に記載の方法により作成した標準的な診断パターンと比較して、前記調査中の生物に該癌又は癌の病期の存在の有無を決定する工程と；

を含む方法。

## 【請求項 4】

タンパク質合成及び / 又は安定性に関与するタンパク質をコードしている前記ファミリー (i) 遺伝子が、

(a) リボソームタンパク質及びリボソーム活性化タンパク質、好ましくはリボソームタ

10

20

30

40

50

ンパク質 L 1 - L 5 6、L 7 A、L 1 0 A、L 1 3 A、L 1 8 A、L 2 3 A、L 2 7 A、L 3 5 A、L 3 6 A、L 3 7 A、P 0、P 1、P 2、S 2 - S 2 9、S 3 1、S 3 3 - S 3 6、S 3 A、S 1 5 A、S 1 8 A、S 1 8 B、S 1 8 C、S 2 7 A、6 3、1 1 5 ( 及び偽遺伝子)、リボソームタンパク質キナーゼ、リボヌクレアーゼ、推定 S 1 R N A 結合ドメインタンパク質、真核生物翻訳開始因子及びグアニンヌクレオチド結合タンパク質 G をコードしている遺伝子と；

( b ) 翻訳阻害因子及び翻訳開始因子、好ましくは真核生物翻訳延長因子、t R N A 合成酵素、R N A 結合タンパク質、ポリアデニル化要素結合タンパク質、チロシンフォスターゼ、真核生物翻訳開始因子及び R N A ポリメラーゼ I、I I I 転写因子をコードしている遺伝子と；

10

( c ) 転写又は翻訳の他のモジュレーター、好ましくはサイクリン D 型結合タンパク質及びグアニンヌクレオチド結合タンパク質をコードしている遺伝子と；

からなる群より選択されたものである、請求項 1 ~ 3 のいずれか一項に記載の方法。

#### 【請求項 5】

防御及び / 又はクロマチンリモデリングファミリーの調節に關与するタンパク質をコードしている前記ファミリー ( i i ) 遺伝子が、

( a ) 免疫応答関連タンパク質、好ましくは T 細胞受容体及び会合成分、種々のサイトカイン、インターフェロン調節因子、オンコスタチン M、白血病阻害因子、ケモカインリガンド及び受容体ファミリー、補体成分、インターフェロン刺激因子、M H C クラス I 又は II ( 又は関連成分)、接着タンパク質、B 細胞におけるカップポリペプチド遺伝子エンハンサーの核因子、ミエリン塩基性タンパク質、カテプシン、トール様受容体、プロテオソームサブユニット、フェリチン、タンパク質キナーゼ又はフォスターゼ、ならびにそれらの活性剤及び阻害剤、白血球免疫グロブリン様受容体、免疫グロブリン成分、デフェンシン、オキシトシン、S 1 0 0 カルシウム結合タンパク質、レクチン及びその受容体及びスーパーファミリー、レプチン、ホスホリパーゼ及び成長因子をコードしている遺伝子と；

20

( b ) T N F 誘発タンパク質、好ましくは T N F アルファ誘発タンパク質 8、インテグリン、B 細胞中のカップ軽ポリペプチド遺伝子エンハンサーの阻害剤、T N F 会合因子 2、5、B 細胞中のカップ軽ポリペプチド遺伝子エンハンサーの核因子、M A P キナーゼ、プロテインキナーゼ C、ユビキタスキナーゼ、カドヘリン、カスパーゼ、サイクリン D 1、スーパーオキシドジスムターゼ及びインターロイキンをコードしている遺伝子と；

30

( c ) 低酸素症誘発タンパク質、好ましくはセストリン、E 1 A 結合タンパク質 p 3 0 0、エンドセリン、毛細血管拡張性運動失調症及び R a d<sub>3</sub> 関連タンパク質、ヘキソキナーゼ 2、T E K チロシンキナーゼ、D N A 断片化因子、カスパーゼ、プラスミノゲン活性化因子、低酸素症誘発性因子 1 及びグルコースホスフェートイソメラーゼをコードしている遺伝子と；

( d ) 酸化ストレスタンパク質、好ましくはスーパーオキシドジスムターゼ、グルタチオンシンテターゼ、カタラーゼ、ラクトペルオキシダーゼ、甲状腺ペルオキシダーゼ、ミエロペルオキシダーゼ、好酸球ペルオキシダーゼ、耐酸化性 1、ペルオキシレドキシニン、チトクローム P 4 5 0、スカベンジャー受容体、パラオキシナーゼ、グルタチオンレダクターゼ、N A D ( P ) H デヒドロゲナーゼ、グルタチオン S - トランスフェラーゼ、カテニン、グルタレドキシニン、熱ショックタンパク質、ミトジェン活性化タンパク質キナーゼ、エノラーゼ、チオレドキシニンレダクターゼ及びペルオキシレドキシニンをコードしている遺伝子と；

40

( e ) クロマチンリモデリング、好ましくはヒストン代替タンパク質に關与しているタンパク質をコードしている遺伝子と、

からなる群より選択されたものである、請求項 1 ~ 4 のいずれか 1 項に記載の方法。

#### 【請求項 6】

( i ) 前記サイトカインは、インターロイキン又はそれらの受容体 ( 好ましくは I L - 1、2、3、4、5、6、7、8、9、1 0、1 1、1 2、1 3、1 5、1 7、1 8、

50

20、22又は24)又は腫瘍壊死因子又はその受容体又はそのスーパーファミリー(好ましくはTNFスーパーファミリーメンバー2、3、4、5、6、7、8、9、11、12、13、14又は15);及び/又は

(ii)前記接着タンパク質は、CD1A、CD1C、CD1D、CD3Z、6、8、11、14、18、24、27、28、29、40、44、50、54、59、74、79B、80、81、83、86、96又はICAM;及び/又は

(iii)前記免疫グロブリン成分は、重鎖又はFc断片、好ましくはIgG、IgE又はIgA又はそれらのスーパーファミリーの重鎖又はFc断片である;及び/又は

(iv)前記成長因子は、血管内皮細胞増殖因子又はエリスロポエチンである、請求項5に記載の方法。

10

【請求項7】

ファミリー(ii)遺伝子によりコードされている前記免疫応答タンパク質は、接着タンパク質、インターロイキン、それらの受容体もしくはスーパーファミリー、TNFもしくはその受容体もしくはスーパーファミリー、免疫グロブリン成分又はエリスロポエチンである、請求項1~6のいずれか1項に記載の方法。

【請求項8】

前記(i)ファミリーをコードしている遺伝子は癌対正常患者の面でダウンレギュレーションされ、(ii)ファミリーの場合では、コードしている遺伝子はアップレギュレーションされる、請求項1~7のいずれか1項に記載の方法。

【請求項9】

前記プローブは、前記癌又は癌の病期により全身疾患状態にある遺伝子に相当する、請求項1~8のいずれか1項に記載の方法。

20

【請求項10】

前記遺伝子は、構成的に中程度又は高度に発現するものである、請求項1~9のいずれか1項に記載の方法。

【請求項11】

前記の組が、ファミリー(i)及びファミリー(ii)からのオリゴヌクレオチドの組み合わせを含む、請求項1~9のいずれか1項に記載の方法。

【請求項12】

前記の組のオリゴヌクレオチドには、ファミリー(i)a、(ii)a及び(ii)eからのオリゴヌクレオチドが含まれている、請求項4~11のいずれか1項に記載の方法。

30

【請求項13】

前記の組には、一種以上のリボソームタンパク質をコードしている遺伝子からのオリゴヌクレオチドならびに必要に応じて一種以上のヒストン及び必要に応じてフェリチンを含んでもよい、請求項1~12のいずれか1項に記載の方法。

【請求項14】

各オリゴヌクレオチドプローブは、表2又は表3に記載されているオリゴヌクレオチドから選択されたか、あるいは表2又は表3に記載の配列もしくはその相補的配列に由来するオリゴヌクレオチドである、請求項1~13のいずれか1項に記載の方法。

40

【請求項15】

各オリゴヌクレオチドプローブは、表2に記載されているオリゴヌクレオチドから選択されたか、あるいは表2に記載の配列もしくはその相補的配列に由来するオリゴヌクレオチドである、請求項14に記載の方法。

【請求項16】

各オリゴヌクレオチドプローブは、表3に記載されているオリゴヌクレオチドから選択されたか、あるいは表3に記載の配列もしくはその相補的配列に由来のものである、請求項14に記載の方法。

【請求項17】

前記の組は、表4に記載のオリゴヌクレオチドから選択されたか、あるいは表4に記

50

載の配列もしくはその相補的配列に由来するものである、一種以上のオリゴヌクレオチドプローブをさらに含む、請求項 14 ~ 16 のいずれか 1 項に記載の方法。

【請求項 18】

表 2、表 3 又は表 4 の前記由来のオリゴヌクレオチドは、それぞれ表 2、表 5 又は表 6 に記載の受託番号により示されている遺伝子の一部分又はその相補的配列である、請求項 14 ~ 17 のいずれか 1 項に記載の方法。

【請求項 19】

前記の組が、10 ~ 500 個のプローブからなる、請求項 1 ~ 18 のいずれか 1 項に記載の方法。

【請求項 20】

前記の組のプローブが、一種以上の固体担体に固定化されている、請求項 1 ~ 19 のいずれか 1 項に記載の方法。

10

【請求項 21】

前記細胞は、疾病細胞でも、疾病細胞に接触したものでも、疾病部位由来のものでもない、請求項 1 ~ 20 のいずれか 1 項に記載の方法。

【請求項 22】

前記試料は、疾病部位から離れた部位から得られたものである、請求項 1 ~ 21 のいずれか 1 項に記載の方法。

【請求項 23】

前記試料は、組織、体液又は体内老廃物で表される、請求項 1 ~ 22 のいずれか 1 項に記載の方法。

20

【請求項 24】

前記試料は、末梢血である、請求項 23 に記載の方法。

【請求項 25】

前記癌は、胃癌、肺癌、乳癌、前立腺癌、腸癌、皮膚癌、結腸癌又は卵巣癌、好ましくは乳癌である、請求項 1 ~ 24 のいずれか 1 項に記載の方法。

【請求項 26】

前記生物が、哺乳動物、好ましくはヒトである、請求項 1 ~ 25 のいずれか 1 項に記載の方法。

【請求項 27】

前記の組における前記プローブの少なくとも一つが、前記癌又は癌の病期の少なくとも 2 つの診断、同定又はモニタリングに好適である、請求項 1 ~ 26 のいずれか 1 項に記載の方法。

30

【請求項 28】

生物における 2 種以上の癌又は癌の病期の診断、同定又はモニタリングのために、前記診断方法の工程 c) で作成した試験パターンを、工程 d) で、請求項 1 又は 4 ~ 27 のいずれか 1 項に記載したようにして作成された少なくとも 2 つの標準診断パターンと比較するが、各標準診断パターンは異種の癌又は癌の病期について作成されたものであることを特徴とする、請求項 3 ~ 27 のいずれか 1 項に記載の診断又は同定又はモニタリングの方法。

40

【請求項 29】

請求項 1 ~ 28 のいずれか 1 項に記載の一组のオリゴヌクレオチドプローブ。

【請求項 30】

一種以上の固体担体上に固定化された、請求項 29 に記載の一组のオリゴヌクレオチドプローブを含む、請求項 1 ~ 28 のいずれか 1 項に記載の方法を実施するためのキット。

【請求項 31】

前記方法をどのように実施するかを詳述した添付文書をさらに含む、請求項 30 に記載のキット。

【請求項 32】

50

前記オリゴヌクレオチドプローブが結合する遺伝子の遺伝子発現のレベルを反映した細胞の遺伝子発現パターンを決定するために、請求項 29～31 のいずれか 1 項に記載の一群のオリゴヌクレオチドプローブ又はキットの使用であって、少なくとも

a) 必要に応じて cDNA に逆転写してもよい、前記細胞から mRNA を単離する工程と；

b) 工程 (a) の mRNA 又は cDNA を、請求項 29～31 のいずれか 1 項に記載の一群のオリゴヌクレオチドプローブ又はキットにハイブリダイズさせる工程と；

c) 該プローブの各々にハイブリダイズしている mRNA 又は cDNA の量を評価して、前記パターンを作成する工程と；

を含む使用。

10

【請求項 33】

生物における癌又は癌の病期に特徴的な標準遺伝子転写パターンの作成方法であって、少なくとも

a) 癌又は癌の病期を有する一種以上の生物の試料から標的ポリペプチドを放出する工程と；

b) 該標的ポリペプチドを、一つ以上の結合パートナーであって、各結合パートナーが請求項 1～27 のいずれか 1 項に記載のオリゴヌクレオチドが結合する遺伝子によりコードされているマーカーポリペプチド (又はその断片) に特異的である結合パートナーと接触させて、該結合パートナーを該標的ポリペプチドに結合させる工程であって、該マーカーポリペプチドは、調査中の生物及びその試料に対応する生物及びその試料中の該癌に特異的である工程と；

20

c) 該結合パートナーに結合している該標的ポリペプチドを評価して、該癌又は癌の病期を有する該試料において前記マーカーポリペプチドを発現する遺伝子の遺伝子発現のレベルを反映した特徴的なパターンを作成する工程と；

を含む方法。

【請求項 34】

試験遺伝子転写パターンの作成方法であって、少なくとも

a) 前記試験生物の試料から標的ポリペプチドを放出する工程と；

b) 該標的ポリペプチドを、一つ以上の結合パートナーであって、各結合パートナーが請求項 1～27 のいずれか 1 項に記載のオリゴヌクレオチドが結合する遺伝子によりコードされているマーカーポリペプチド (又はその断片) に特異的である結合パートナーと接触させて、該結合パートナーを該標的ポリペプチドに結合させる工程であって、該マーカーポリペプチドは、調査中の生物及びその試料に対応する生物及びその試料中の該癌に特異的である工程と；

30

c) 該結合パートナーに結合している該標的ポリペプチドを評価して、該試料において該マーカーポリペプチドを発現する遺伝子の遺伝子発現のレベルを反映した特徴的なパターンを作成する工程と；

を含む方法。

【請求項 35】

生物における癌又は癌の病期の診断又は同定又はモニタリング方法であって、

40

a) 該生物の試料から標的ポリペプチドを放出する工程と；

b) 該標的ポリペプチドを、一つ以上の結合パートナーであって、各結合パートナーが請求項 1～27 のいずれか 1 項に記載のオリゴヌクレオチドが結合する遺伝子によりコードされているマーカーポリペプチド (又はその断片) に特異的である結合パートナーと接触させて、該結合パートナーを該標的ポリペプチドに結合させる工程であって、該マーカーポリペプチドは、調査中の生物及びその試料に対応する生物及びその試料中の該癌に特異的である工程と；

c) 該結合パートナーに結合している該標的ポリペプチドを評価して、該試料において該マーカーポリペプチドを発現する遺伝子の遺伝子発現のレベルを反映した特徴的なパターンを作成する工程と；

50

d) 該パターンを、調査中の生物及び試料に対応する生物からの試料を用いて請求項 33の方法により作成した標準診断パターンと比較して、調査中の該生物における該癌又は癌の病期の存在を示す相関度を求める工程と；  
を含む方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、分析法、特に診断法に使用することができる、細胞中の遺伝子転写レベルの評価用オリゴヌクレオチドプローブに関する。そうしたプローブは、キットの形態で提供されるのが都合がよい。異なる型のプローブを、遺伝子発現パターンの作成、及び異なる癌又は癌の病期の同定、診断又はモニタリングの技術に使用できる。

10

【背景技術】

【0002】

例えば、診断用途に用いられる試料分析を迅速かつ容易におこなう方法を確定することは、依然として数多くの研究者の目標とされている。エンドユーザーは、経済的であり、統計的に有意な結果が得られ、高度の熟練者を必要とすることなく日常的に実施できる方法を求めている。

【0003】

細胞内遺伝子発現の解析が、これらの細胞の状態及び重要なこととして細胞が由来する個体の状態についての情報を提供するのに用いられてきた。細胞における種々の遺伝子の相対的な発現が、体内の特定の状態を反映するものとして同定されてきた。例えば、癌細胞は、種々のタンパク質の発現変化を示すことが知られており、したがって、転写物又は発現タンパク質は、疾病状態のマーカーとして使用できる。

20

【0004】

したがって、バイオプシー組織を分析して、これらのマーカーが存在するかどうかを調べることができ、疾病部位由来の細胞は、該マーカーの存在により体の他の組織又は体液で同定できる。さらに、変化した発現の産物は血流中に放出されるかも知れず、これらの産物を分析してもよい。さらに、疾病細胞と接触した細胞は、これらの細胞と直接接触することにより影響を受けて遺伝子発現の変化を生じ、それらの発現又は発現産物を同様に分析してもよい。

30

【0005】

しかしながら、これらの方法にはいくつかの限界がある。例えば、癌を同定するために特定の腫瘍マーカーを使用することには、特異性又は感度の欠如、マーカーが特異型の癌の他に病態と会合すること、及び自覚のない個体で検出するのが困難であることといった種々の欠点がある。

【0006】

一種又は二種のマーカー転写物又はマーカータンパク質の分析の他に、より最近では、遺伝子発現パターンが分析されてきた。疾病診断における大規模遺伝子発現分析を含む研究のほとんどは、異常組織又は細胞由来の臨床試料を用いてきた。例えば、遺伝子発現データが類似する癌種間を区別するのに使用され得ることを示すいくつかの最近の刊行物では、異常組織又は細胞からは得た臨床試料を使用している（非特許文献1；非特許文献2；非特許文献3；非特許文献4）。

40

【0007】

しかしながら、これらの方法は、異常細胞又はこれらの細胞の産物又は異常細胞に接触した細胞を含有する試料を分析するものである。このような試料の分析では、疾病の存在とその場所についての情報が必要である。これは、自覚のない患者では困難なことがある。さらに、試料は、疾病部位から、例えば、脳の疾病では必ずしも採取できない。

【0008】

本発明者らは、甚だ重要な知見において、細胞を得た生物体の状態に関連する情報を得るために、体内のすべてについて、細胞のこれまで利用されてこなかった可能性につい

50

て確認した。特許文献1は、疾病部位から離れた細胞の遺伝子発現、例えば、癌部位から離れたところで採取した末梢血の分析について記載している。特許文献2（引用することにより本明細書に組み込まれる）は、乳癌及びアルツハイマー病の診断用である特定プローブについて記載し、この目的及び他の疾病を診断するのに好適な他のプローブを同定するためのプロトコルについて検討している。

【0009】

この知見は、生物の体の異なる部分が互いに動的相互作用するという前提に基づいている。疾病が体の一部分に影響を及ぼすと、体の他の部分も影響される。この相互作用は、罹患部から放出される広範囲の生化学的シグナルから生じ、体の他の部分に影響を及ぼす。放出されたシグナルにより誘発された生化学的変化及び生理学的変化の性質は、体の部分ごとに異なるけれども、これらの変化は、遺伝子発現のレベルで測定でき、診断に使用できる。

10

【0010】

生物における細胞の生理学的状態は、体内で遺伝子が発現するパターンにより決定される。このパターンは、前記細胞に対する内的及び外的な生物学的刺激に依存し、これらの刺激の程度又は性質が変化すると、パターンが変化して、その中で異なる遺伝子が細胞中で発現する。生物試料における細胞の遺伝子発現パターンの系統的变化を分析することにより、前記細胞に作用している生物学的刺激の種類及び性質についての情報を得ることができるという理解がますます受け入れられるようになってきている。したがって、例えば、試験試料における細胞の多数の遺伝子の発現をモニタリングすることにより、それらの遺伝子が特定の疾病、状態又はその病期について特徴的なパターンで発現するかどうかを測定できる。したがって、例えば、組織又は体液から得た細胞の遺伝子活性の変化を測定することが、疾病を診断するための強力なツールとなる。

20

【0011】

このような方法には、種々の利点がある。しばしば罹患した体内の一定領域から臨床試料を得ることが困難なことがあり、かつ体における望ましくない浸潤が生じることがあり、例えば、バイオプシーを利用して癌についての試料を得ることがしばしばある。アルツハイマー病といったある症例においては、罹患した脳の試料は、死後でしか得ることができない。さらに、得られる組織試料は、不均質であることがよくあり、異常細胞と非異常細胞の両方の混合物を含むことがあり、得られる遺伝子発現データの解析が複雑かつ困難になる。

30

【0012】

腫瘍の形態的な外観に関して発病的に均質と思われる腫瘍組織のプールは、分子レベルでは極めて不均質なことがあり（非特許文献3）、ならびに実際に本質的に異なる疾病の腫瘍を含むことがあることが示唆された（非特許文献3、非特許文献2）。

【特許文献1】WO98/49342

【特許文献2】PCT/GB03/005102

【非特許文献1】Alone et al. 1999、PNAS、96、P6745-6750

【非特許文献2】Golub et al. 1999、Science、286、P531-537

40

【非特許文献3】Alizadeh et al. 2000、Nature、403、p503-511

【非特許文献4】Bittner et al. 2000、Nature、406、p536-540

【発明の開示】

【発明が解決しようとする課題】

【0013】

疾病、その状態又はその病期を特定する目的のためには、複数の種類の細胞の均質混合物である臨床試料が体内において容易にアクセスできる領域から得ることができること

50

から、罹患した組織又は細胞から直接臨床試料を得ることを必要としない方法が極めて望ましい。

【課題を解決するための手段】

【0014】

本発明者らは、このたび癌、特に乳癌を特定するのに驚くべき有用性がある一組のプローブを誘導できる配列群を特定した。したがって、本発明者らは、癌患者からの血液試料の細胞において発現が変わり、それを癌又は癌の病期を同定、診断又はモニタリングする方法に使用されるプローブを作製するのに使用できる、遺伝子ファミリーについて説明する。

【0015】

本発明に至るまでの研究において、本発明者らは、多数の遺伝子の発現レベルを、正常患者と癌患者との比較において調べた。多数の遺伝子の発現が変化したことが判明しただけではなく、発現が変化したものは、機能の面で特別の遺伝子ファミリーに属することが分かった。これらの遺伝子自体から、対応するプローブを作製することができるプールが得られ、これをまとめて使用して個体における、該遺伝子の発現のフィンガープリントを得ることができる。これらの遺伝子の発現は、癌患者において変化し、したがって、癌の状態を調べるのに有益であると考えられるので、プローブ収集物から得たフィンガープリントは、正常状態に対して疾病状態を示している。

【0016】

癌患者において示差的に発現していることが確認された遺伝子ファミリーは、以下のよう

にまとめることができる：

(i) タンパク質の合成及び/又は安定性に関与しているタンパク質をコードしている遺伝子；

(ii) 防御及び/又はクロマチンリモデリングの調節に関与しているタンパク質をコードしている遺伝子。

【0017】

ファミリー(i)には、以下の遺伝子が含まれる：

(a) リボソームタンパク質及びリボソーム活性化タンパク質(すなわち、リボソームタンパク質の成分又はそれらの機能の変更に

関与している成分を含むタンパク質であって、癌患者においてダウンレギュレーションされていることが分かったタンパク質)をコードしている遺伝子。これらのコードされているタンパク質には、リボソームタンパク質L1-L56、L7A、L10A、L13A、L18A、L23A、L27A、L35A、L36A、L37A、P0、P1、P2、S2-S29、S31、S33-S36、S3A、S15A、S18A、S18B、S18C、S27A、63、115(及び偽遺伝子)、リボソームタンパク質キナーゼ(例えば、S6キナーゼ)、リボヌクレアーゼ、推定S1RNA結合ドメインタンパク質、真核生物翻訳開始因子及びグアニンヌクレオチド結合タンパク質Gなどがある；

(b) 翻訳阻害因子及び翻訳開始因子(すなわち、タンパク質産物へのmRNAの翻訳に関与しているタンパク質であって、癌患者においてダウンレギュレーションされていることが分かったタンパク質)をコードしている遺伝子。これらのコードされているタンパク質には、真核生物翻訳延長因子、tRNA合成酵素、RNA結合タンパク質、ポリアダニル化要素結合タンパク質、チロシンフォスターゼ、真核生物翻訳開始因子及びRNAポリメラーゼI、III転写因子などがある；

(c) 転写又は翻訳の他のモジュレータ、例えばサイクリンD型結合タンパク質及びグアニンヌクレオチド結合タンパク質をコードしている遺伝子。

【0018】

ファミリー(ii)には、以下の遺伝子が含まれる：

(a) 免疫応答関連タンパク質(すなわち、免疫刺激に

応答してアップレギュレーションされ、炎症に

応答するか又は炎症反応をおこす際にアップレギュレーションされるタンパク質を含み、癌患者においてアップレギュレーションされることが判明したタンパク

10

20

30

40

50

質)をコードしている遺伝子。これらのコードされているタンパク質には、T細胞受容体及び会合成分、例えば、タンパク質キナーゼ、種々のサイトカイン、例えば、インターロイキン及びそれらの受容体(例えば、IL-1、2、3、4、5、6、7、8、9、10、11、12、13、15、17、18、20、22、24)、腫瘍壊死因子及びその受容体及びそのスーパーファミリー(例えば、TNFスーパーファミリーメンバー2、3、4、5、6、7、8、9、11、12、13、14、15)、インターフェロン調節因子、オンコスタチンM、白血病阻害因子、ケモカインリガンド及び受容体ファミリー(例えば、No.1~28)、補体成分、インターフェロン刺激因子、例えば、転写因子、MHC(例えば、HLA)クラスI又はII(又は関連成分)(例えば、DQ、DR、DO、DP、DMアルファ又はベータ)、接着タンパク質(例えば、CD1A、CD1C、CD1D、CD3Z、6、8、11、14、18、24、27、28、29、40、44、50、54、59、74、79B、80、81、83、86、96、ICAM)、B細胞におけるカップポリペプチド遺伝子エンハンサーの核因子、ミエリン塩基性タンパク質、カテプシン、トール様受容体、プロテオソームサブユニット、フェリチン、タンパク質キナーゼ又はフォスファターゼ、ならびにそれらの活性剤及び阻害剤、白血球免疫グロブリン様受容体、免疫グロブリン成分、例えば、重鎖又はFc断片、例えば、IgG、IgE又はIgA又はそれらのスーパーファミリーの重鎖又はFc断片、デフェンシン、オキシトシン、S100カルシウム結合タンパク質、レクチン及びその受容体及びスーパーファミリー、レプチン、ホスホリパーゼ及び成長因子(例えば、内皮細胞成長因子又はエリスロポエチン)などがある;

10

20

(b) TNF誘発タンパク質(すなわち、TNFにさらされると個体中で誘発され、癌患者においてアップレギュレーションされることが分かったタンパク質)をコードしている遺伝子。これらのコードされているタンパク質には、TNFアルファ誘発タンパク質8、インテグリン、B細胞中のカップ軽ポリペプチド遺伝子エンハンサーの阻害剤、TNF会合因子2、5、B細胞中のカップ軽ポリペプチド遺伝子エンハンサーの核因子、MAPキナーゼ、プロテインキナーゼC、ユビキタスキナーゼ、カドヘリン、カスパーゼ、サイクリンD1、スーパーオキシドジスムターゼ及びインターロイキンなどがある;

(c) 低酸素症誘発タンパク質(すなわち、個体又はその一部分が低酸素状態にあるときに誘発され、癌患者においてアップレギュレーションされることが分かったタンパク質)をコードしている遺伝子。これらのコードされているタンパク質には、セストリン、E1A結合タンパク質p300、エンドセリン、毛細血管拡張性運動失調症及びRad3関連タンパク質、ヘキソキナーゼ2、TEKチロシinkinナーゼ、DNA断片化因子、カスパーゼ、プラスミノゲン活性化因子、低酸素症誘発性因子1及びグルコースホスフェートイソメラーゼなどがある;

30

(d) 酸化ストレスタンパク質(すなわち、酸化的ストレス下で個体又はその一部分で誘発され、癌患者においてアップレギュレーションされることが分かったタンパク質)をコードしている遺伝子。これらのコードされているタンパク質には、スーパーオキシドジスムターゼ、グルタチオンシンテターゼ、カタラーゼ、ラクトペルオキシダーゼ、甲状腺ペルオキシダーゼ、ミエロペルオキシダーゼ、好酸球ペルオキシダーゼ、耐酸化性1、ペルオキシレドキシシン、チトクロームP450、スカベンジャー受容体、パラオキシナーゼ、グルタチオンレダクターゼ、NAD(P)Hデヒドロゲナーゼ、グルタチオンS-トランスフェラーゼ、カテニン、グルタレドキシシン、熱ショックタンパク質(例えば、熱ショック転写因子)、ミトジェン活性化タンパク質キナーゼ、エノラーゼ、チオレドキシシンレダクターゼ及びペルオキシレドキシシンなどがある;

40

(e) クロマチンリモデリングに関与しているタンパク質(すなわち、クロマチン構造を維持し又は変更するのに役立ち、遺伝子調節に必須であることがあるタンパク質)をコードしている遺伝子。これらのコードされているタンパク質には、ヒストン代替タンパク質、例えば、H3.3A又はH3.3Bファミリーなどがある。

【0019】

上記ファミリーの範囲内に含まれる適切な遺伝子配列は、ノルウェーの国立バイオテ

50

クノロジー情報センター (National Centre for Biotechnology Information) の遺伝子又はタンパク質データベース上で、キーワードとしてファミリー名、例えば、「免疫応答」を用いて適切なデータベースを検索することにより同定できる。本明細書に記載の試験に用いられるオリゴヌクレオチドを開発するのにこのような遺伝子配列が有用であることを確認するために、特定の遺伝子配列の発現を、試験癌患者と正常患者とを比較して評価することができる。発現の変化が対照レベルを超えるか又は対照レベルよりも小さい場合、プローブの派生させるためのその配列の有用性が示される。

#### 【0020】

一般的に、上記 (i) ファミリーをコードしている遺伝子は、癌対正常患者においてダウンレギュレーションされ、(ii) ファミリーの場合では、コードしている遺伝子はアップレギュレーションされる。

10

#### 【0021】

癌患者では、リボソーム産生及び翻訳制御に関与している遺伝子の発現が規則的に減少する場合、血液細胞は、低酸素及びエネルギー欠乏という環境に対して細胞が適合しようとしてタンパク質合成速度を減少させることによりもたらされる患者の新しい状態にตอบสนองしていると推測される。このことは、MnSODやフェリチンといった反応性酸素種 (ROS) に対する防御に関与している遺伝子は、癌試料においてアップレギュレーションされることが観察されることから支持される。エリスロポエチンが低レベルであることは、癌患者では酸素レベルが低いことを示している。また、TNFは、例えば、フェリチン、デフェンシン、MnSOD及びカルグラヌリンBの発現をアップレギュレーションすることが知られているので、TNF活性化は、上記遺伝子のファミリーにおける変化の1つの経路であると思われる。また、TNFはEPO産生を阻害し、それ自体が血液環境において低酸素状態をもたらし得る。低酸素症では、TNFレベルを誘起することが知られている。これらの変化は、血流に入って来る血管形成因子により引き起こされる。理論に縛られることを意図しないが、上記効果の基礎をなす仮説を、図1に示す。

20

#### 【0022】

したがって、本発明によれば、発現が特定の癌又は癌の病期に特徴的なパターンで影響される細胞中の遺伝子に対応する一組のオリゴヌクレオチドプローブが提供される。ここで、前記遺伝子は、前記癌又は癌の病期により系統的に影響される。好ましくは、前記遺伝子は、構成的に適度又は高度に発現される。好ましくは、上記遺伝子は、試料の細胞中で適度又は高度に発現するが、疾病細胞から得た細胞中、又はそのような疾病細胞に接触した細胞内では発現しない。このようなプローブは、特に疾病部位から離れた細胞から単離したときには、疾病が臨床的に認識できるレベルまで進展していなくてよく、癌の発病後の極めて早期、すなわち、他の自覚的又は他覚的な症状が現れる数年前であっても癌又は癌の病期を検出できる。

30

#### 【発明を実施するための最良の形態】

#### 【0023】

本明細書で使用される用語「系統的に」影響される遺伝子とは、発現が体内で疾病細胞や疾病部位と直接接触せずに影響される遺伝子を意味し、調査中の細胞は疾病細胞ではない。

40

#### 【0024】

ここで「接触」とは、一つの細胞の他細胞への直接的な影響が観察されるように互いに極めて接近する細胞、例えば、第二の細胞に影響するのに大きくはなれた所の第一の細胞から放出された二次分子により媒介されない応答、例えば免疫応答を意味する。好ましくは、接触は、物理的接触であるか、あるいは立体的にできるだけ接近している接触、都合よくは互いに接触する細胞が、例えば、 $1\text{ cm}^3$ 内の同じ単位容積に観察されることを意味する。

#### 【0025】

「異常細胞 (疾患細胞)」は、表現型の変化を発現している細胞であり、その寿命中の

50

ある時期での疾病部位に存在するもの、例えば、腫瘍部位の腫瘍細胞又は腫瘍から広まった腫瘍細胞、又は脳腫瘍の場合は脳細胞である。

【0026】

「中程度又は高度に」発現した遺伝子とは、コピー数が30～100コピー/細胞(細胞内の平均mRNA分子数が $3 \times 10^5$ と仮定)超である静止細胞に存在する遺伝子を意味する。

【0027】

上記の特性を有する特異的プローブは、本明細書に記載のようにして提供される。

【0028】

したがって、本発明による一態様によれば、一組のオリゴヌクレオチドプローブであって、前記の組が、

前記に記載したファミリー(i)又は(ii)からの遺伝子配列またはこのような配列由来の遺伝子配列に相当するオリゴヌクレオチド、又は相補的配列を有するオリゴヌクレオチドもしくは機能的に等価のオリゴヌクレオチドから選択される少なくとも10個のオリゴヌクレオチドを含む一組のオリゴヌクレオチドプローブが提供される。

【0029】

さらに、本発明によれば、本明細書に記載の方法に使用される一組のオリゴヌクレオチドの調製方法であって、ファミリー(i)からの遺伝子配列に相当する一つ以上のオリゴヌクレオチドと、ファミリー(ii)からの遺伝子配列に相当する一つ以上のオリゴヌクレオチドを選択する工程を含む方法が提供される。好ましくは、複数のオリゴヌクレオチドを、各ファミリーから(例えば、異なるサブファミリーから)選択する。この場合の選択されたオリゴヌクレオチドは、本明細書に記載の好ましい遺伝子からのものである。

【0030】

また、本発明によれば、一種以上のオリゴヌクレオチドプローブであって、各オリゴヌクレオチドプローブが、表2、表3又は表4にあげたオリゴヌクレオチドから(例えば、表2から)選択されたもの、あるいは表2、表3又は表4に記載の配列もしくはそれと相補的配列に由来するものであるオリゴヌクレオチドから選択されたものである、一種以上のオリゴヌクレオチドプローブが提供される。前記由来のオリゴヌクレオチドには、これらの表に記載の配列に相当する遺伝子、例えば、表2、表5又は表6に記載の遺伝子(受託番号参照)、又はその相補的配列に相当する遺伝子に由来のオリゴヌクレオチドなどがあげられる。このようなプローブを、本発明の産物及び方法に使用することにより、本発明のさらなる態様が提供される。

【0031】

本明細書において、「オリゴヌクレオチド」は、高分子構造において少なくとも6個のモノマー、すなわち、ヌクレオチド又はそれらの修飾形態を有する核酸分子である。核酸分子は、DNA、RNA又はPNA(ペプチド核酸)又はそのハイブリッド又はそれらの修飾型、例えば、メチル化によるか、又は合成中に修飾されたかもしくは非天然塩基から構成された、化学的に修飾された形態、例えば、LNA(固定核酸)であってもよい。但し、これらは、相補的配列に結合する能力を保持している。このようなオリゴヌクレオチドは、本明細書に準じて標的配列を調べるのに使用されるので、本明細書では、オリゴヌクレオチドプローブ又は単にプローブと称する。

【0032】

ファミリー(i)又は(ii)からの遺伝子配列に相当するオリゴヌクレオチドは、前記遺伝子配列のすべてもしくは一部分又はその転写物に相当するオリゴヌクレオチドを意味する。そうした遺伝子配列の一部分を使用するとき、この一部分は、本明細書に記載のオリゴヌクレオチドプローブの要件、例えば、長さ及び機能の要件を満足するものである。好ましくは、前記部分は、以下で説明するサイズを有する。前記オリゴヌクレオチドは、以下一次オリゴヌクレオチドと称する。誘導オリゴヌクレオチドとは、一次オリゴヌクレオチドの一部ではあるが、本明細書に記載のプローブ要件を満足するオリゴヌクレ

オチドを意味する。

【0033】

好ましくは、前記の組を構成しているオリゴヌクレオチドプローブは、標的分子が結合できるように長さが少なくとも15塩基のものである。とりわけ好ましくは、前記オリゴヌクレオチドプローブは、長さが20~200塩基、例えば、長さ30~150塩基、好ましくは長さ50~100塩基である。

【0034】

本明細書で使用される用語「相補的配列」とは、連続した相補的塩基（すなわち、T : A、G : C）を有する配列を意味する。したがって、相補的配列は、相補性を介して互いに結合できるものである。

10

【0035】

「10種のオリゴヌクレオチド」とは、10の異なるオリゴヌクレオチドを意味する。本明細書に記載の遺伝子配列ファミリーからのオリゴヌクレオチド、誘導オリゴヌクレオチド及びそれらの機能的等価物は異なるオリゴヌクレオチドとみなされるが、相補的オリゴヌクレオチドは異なるものとはみなされない。しかしながら、好ましくは、少なくとも10種のオリゴヌクレオチドは、上記した遺伝子配列ファミリー内の10の異なる遺伝子配列（又は誘導オリゴヌクレオチド又はそれらの機能的等価物）に相当する。したがって、10の異なるオリゴヌクレオチドは、好ましくは10の異なる転写物に結合できるものである。

【0036】

好ましくは、少なくとも10種のオリゴヌクレオチドは、ファミリー（i）及び（ii）からのオリゴヌクレオチドの組み合わせから構成される。例えば、各ファミリーから5種のオリゴヌクレオチドを使用してもよいし、一つのファミリーから4種と、その他のファミリーから6種を使用してもよい。これによって正常患者に対して癌患者においてアップレギュレーション又はダウンレギュレーションされる遺伝子を使用することができる。異なるサブファミリーから一種又はそれ以上のオリゴヌクレオチドを使用するのが都合がよい。例えば、サブファミリー（i）a、（i）b、（i）c、（ii）a及び（ii）bから、各々2プローブを使用してもよい。とりわけ、前記の組のオリゴヌクレオチドには、ファミリー（i）a、（ii）a及び（ii）eからのオリゴヌクレオチドが含まれることが好ましい。

20

【0037】

ファミリー（i）a遺伝子によりコードされている好ましいタンパク質はリボソームタンパク質であり、好ましくは各組にはこのようなタンパク質をコードしている遺伝子からのオリゴヌクレオチドが含まれる。

30

【0038】

ファミリー（ii）a遺伝子によりコードされている好ましい免疫応答タンパク質には、接着タンパク質、インターロイキン、それらの受容体及びスーパーファミリー、TNF、その受容体及びスーパーファミリー、免疫グロブリン成分ならびにエリスロポエチンなどがある。

【0039】

前記の組の特に好ましいものには、一種又はそれ以上のリボソームタンパク質ならびに必要に応じて一種又はそれ以上のヒストン及び必要に応じてフェリチンをコードしている遺伝子からのオリゴヌクレオチドが含まれる。

40

【0040】

好ましい前記オリゴヌクレオチドは、表2又は表3に記載されているものか、あるいは表2又は表3に記載の配列由来のもの、例えば、表2に記載のものである。前記の組は、さらに表4にあげられているか、あるいは表4に記載の配列又はその相補的配列に由来する一種又はそれ以上のオリゴヌクレオチドプローブを含んでもよい。前記誘導オリゴヌクレオチドには、これらの表に記載の配列に相当する遺伝子、例えば、表2、表5又は表6に記載の遺伝子（受託番号参照）、又はその相補的配列に相当する遺伝子由来のオリ

50

ゴヌクレオチドなどがあげられる。

【0041】

本明細書に記載の「組」とは、一群のユニークなオリゴヌクレオチドプローブ（すなわち、固有の配列）、好ましくは1000より少ないオリゴヌクレオチドプローブ、とりわけ500より少ないプローブ、例えば、好ましくは10～500プローブ、例えば、10～100、200もしくは300プローブ、とりわけ好ましくは20～100プローブ、例えば、30～100プローブからなるものである。場合によっては、10より少ないプローブ、例えば、2～9プローブ、例えば、5～9プローブを使用してもよい。

【0042】

プローブ数を増加すると、当該の特定遺伝子の発現を同様に変更できる他の疾病と比較することにより、解析がうまくいかない可能性、例えば、誤診を防止できることが分かる。また、本明細書に記載されていない他のオリゴヌクレオチドプローブも存在していてもよい。このことは、特に前記組のオリゴヌクレオチドプローブの最終用途に役立つ場合に言える。しかしながら、好ましくは、前記の組は、本明細書に記載のオリゴヌクレオチドからのみ、又はそのサブセット（上記したようなサイズのもの）からのみ構成されている。

10

【0043】

ユニークな各オリゴヌクレオチドプローブの複数のコピー、例えば、10又はそれ以上のコピーが、各組に存在していてもよいが、単一プローブのみを構成していてもよい。

【0044】

好ましくは固体担体上に固定化されているか、又はそうした固定化のための手段を有していてもよい一組のオリゴヌクレオチドプローブは、上記したものから選択された少なくとも10種のオリゴヌクレオチドプローブを含む。上記したように、これらの10プローブは、ユニークでかつ異なる配列を有していなければならない。しかしながら、そうであっても同じ遺伝子を認識するが、異なるスプライシング事象を反映する2つの別個のプローブを使用してもよい。しかしながら、特異的な遺伝子と相補的かつそれに結合するオリゴヌクレオチドプローブが好ましい。

20

【0045】

本明細書に記載の「機能的に等価な」又は誘導されたオリゴヌクレオチドとは、本明細書に記載の遺伝子配列ファミリーにおける配列からのオリゴヌクレオチドと同じ遺伝子を同定できるオリゴヌクレオチドを意味する。すなわち、機能的に等価な又は誘導されたオリゴヌクレオチドは、一次オリゴヌクレオチド又は誘導オリゴヌクレオチド（又はその相補的配列）と同じ、ある遺伝子（標的核酸分子）から転写されたmRNA分子（又はDNA）に結合できる。したがって、好ましい特徴においては、前記の誘導又は機能的に等価なオリゴヌクレオチドは、表2、表5又は表6に記載の遺伝子配列の一部分又はその相補的配列である。好ましくは、前記の機能的に等価なオリゴヌクレオチドは、一次オリゴヌクレオチド又は誘導オリゴヌクレオチドと同じスプライシング産物を認識、即ち、それに結合することができる。好ましくは、前記mRNA分子は、前記一次オリゴヌクレオチド又は誘導オリゴヌクレオチドに相当する完全長mRNA分子である。

30

【0046】

本明細書で使用される用語「結合できる」又は「結合」は、以下で説明する条件下でハイブリダイゼーションできる能力を意味する。

40

【0047】

あるいは機能的に等価な発現されるオリゴヌクレオチド（又は相補的配列）は、後記するように、一次オリゴヌクレオチド又は誘導オリゴヌクレオチド又は相補的オリゴヌクレオチドが結合する標的分子の領域と配列が同一であるか、又はそれにハイブリダイズする。好ましくは、機能的に等価なオリゴヌクレオチド（又はそれらの相補的配列）は、後記する条件下で一次オリゴヌクレオチド又は誘導オリゴヌクレオチドに相当するmRNA配列の一つにハイブリダイズするか、あるいは一次オリゴヌクレオチド又は誘導オリゴヌクレオチドに相当するmRNA配列の一つの一部分と配列が同一である。これに関連して

50

「部分」とは、一続きの少なくとも5塩基、例えば、少なくとも10又は20塩基、例えば、5～100塩基、例えば、10～50塩基又は15～30塩基を意味する。

【0048】

特に好ましい態様において、機能的に等価なオリゴヌクレオチドは、一次オリゴヌクレオチド又は誘導オリゴヌクレオチドが結合する標的核酸分子(mRNA又はcDNA)の領域のすべて又は一部分に結合する。「標的」核酸分子は、遺伝子転写物又は関連産物、例えば、mRNA又はcDNA、又はその増幅産物である。前記一次オリゴヌクレオチド又は誘導オリゴヌクレオチドが結合する前記標的分子の前記「領域」は、相補性が存在する一帯である。その最も大きな領域は、一次オリゴヌクレオチド又は誘導オリゴヌクレオチドの完全長である。しかしながら、全体の一次配列又は誘導オリゴヌクレオチドが標的配列の領域に相補的でない場合にはもっと短くてもよい。

10

【0049】

前記標的分子の前記領域の前記部分は、少なくとも5塩基の一続き、例えば、少なくとも10もしくは20塩基、例えば、5～100塩基、例えば、10～50塩基又は15～30塩基であることが好ましい。これは、例えば、一次オリゴヌクレオチド又は誘導オリゴヌクレオチドの塩基と同一である、いくつかの塩基を有する前記の機能的に等価なオリゴヌクレオチドにより得られる。これらの塩基は、例えば、機能的に等価なオリゴヌクレオチドの一部において連続する一帯にわたって同一であるか、あるいは非連続的に存在していてもよいが、この場合標的配列に結合できる程度に十分な相補性を有する。

【0050】

したがって、好ましい特徴において、前記の機能的に等価なオリゴヌクレオチドは、一次オリゴヌクレオチド又は誘導オリゴヌクレオチド又はその相補的配列に高ストリンジエンシー条件下でハイブリダイズする。別の表現では、前記の機能的に等価なオリゴヌクレオチドは、一次オリゴヌクレオチドのすべて又は一部分との配列の同一性が高い。好ましくは、前記の機能的に等価なオリゴヌクレオチドは、一次オリゴヌクレオチドのすべて又はその一部分に対する配列の同一性は、少なくとも70%、好ましくは少なくとも80%、例えば、少なくとも90%、95%、98%又は99%の同一性を有している。これに関連して使用される用語「部分」は、前記一次オリゴヌクレオチドにおける一続きの少なくとも5塩基、例えば、少なくとも10又は20塩基、例えば、5～100塩基、例えば、10～50塩基又は15～30塩基の一続きを意味する。前記一次オリゴヌクレオチドの一部に対する配列の同一性しかないときには、その配列の同一性は高い、例えば、上記したように少なくとも80%であることがとりわけ好ましい。

20

30

【0051】

上記した機能的な要件を満足する機能的に等価なオリゴヌクレオチドには、一次オリゴヌクレオチド由来のもの、ならびにさらに単一又は複数のヌクレオチド塩基(又は等価の)の置換、付加及び/又は欠失により修飾されているが、それでもまだ機能的活性を保持しているもの、例えば、一次オリゴヌクレオチド又はさらに誘導又は修飾される誘導オリゴヌクレオチドと同じ標的分子に結合するものが挙げられる。好ましくは、前記修飾は、1～50塩基、例えば、10～30塩基、好ましくは1～5塩基のものである。とりわけ、少しだけの修飾、例えば、10塩基より少ない変化、例えば、5塩基より少ない変化が存在することが好ましい。

40

【0052】

「付加」等価物の意味の中には、一次オリゴヌクレオチド又は誘導オリゴヌクレオチドが結合する標的分子上の塩基の連続した一帯に相補的である、追加配列を含有するオリゴヌクレオチドが含まれる。あるいは付加は、例えば、さらなる性質を付与され、例えば、リンカーなどの固定化手段を得てオリゴヌクレオチドプローブを固体担体に結合させることができる、異なる無関係配列を含むことができる。

【0053】

特に好ましいものは、天然の等価物、例えば、生物学的変異体、例えば、対立遺伝子、地理的又はアロタイプ変異体、例えば、異なる種に存在する遺伝的変異体に相当するオ

50

リゴヌクレオチドである。

【0054】

機能的等価物には、修飾塩基、例えば、非天然塩基を用いたオリゴヌクレオチドなどがある。このような誘導体は、合成中に調製してもよいし、又は製造後に修飾して調製してもよい。

【0055】

低ストリンジェンシーの条件下で結合する「ハイブリダイゼーション」配列は、非ストリンジェンシー条件（例えば、室温で6XSSC/50%ホルムアミド）で結合しかつ低ストリンジェンシーの条件（2XSSC、常温、より好ましくは2XSSC、42）下で洗浄したときに結合した状態を維持するものである。高ストリンジェンシー下でのハイブリダイゼーションは、洗浄を2XSSC、65（ここで、SSC=0.15M NaCl、0.015Mクエン酸ナトリウム、pH7.2）でおこなう上記条件を意味する。

10

【0056】

本明細書で使用される用語「配列の同一性」は、以下のパラメータを用いたClustal W (Thompsonら、1994、Nucleic Acids Res.、22、p4673-4680)を用いて評価して得られる値を意味する。

ペア配列パラメーター方法：正確、マトリックスIUB、ギャップオープンペナルティ：15.00、ギャップ拡張ペナルティ：6.66；

複数の配列パラメーターマトリックス：IUB、ギャップオープンペナルティ：15.00、遅延同一性%：30、ネガティブマトリックス：なし、ギャップ拡張ペナルティ：6.66、DNAトランジションウエイティング：0.5

20

【0057】

特定の塩基での配列同一性は、単純に誘導された同一塩基を含む。

【0058】

また、本発明は、表2、表3又は表4のオリゴヌクレオチド又は表2、表3又は表4の誘導オリゴヌクレオチド（例えば、表2、表5又は表6に記載の配列又はそれらに相補的な配列を有するもの）が結合するmRNA配列によりコードされているポリペプチドにも拡張される。さらに、本発明は、前記ポリペプチドのいずれかに結合する抗体にも及ぶ。

30

【0059】

上記したように、前記の組のオリゴヌクレオチドプローブは、一種又はそれ以上の固体担体上に固定化されているのが都合がよいことがある。各ユニークなプローブの単一又は好ましくは複数のコピーは、前記固体担体に結合される。例えば、各ユニークなプローブの10又はそれ以上のコピー、例えば、少なくとも100コピーが存在する。

【0060】

一つ又はそれ以上のユニークなオリゴヌクレオチドプローブは、別々の固体担体に結合させて、それらは一緒に複数の固体担体上に固定化された一組のプローブを形成してもよい。例えば、一つ又はそれ以上のユニークなプローブを、複数のビーズ、膜、フィルター、バイオチップなどに固定して、後記キットのモジュールと一緒に構成する一組のプローブと一緒に構成してもよい。各プローブに係するシグナル（後記のようにして生成）は別個に決定できなければならないけれども、異なるモジュールの固体担体は物理的に結合しているのが都合がよい。別法として、プローブを、同じ固体担体の離れた部分に固定化してもよい。例えば、各ユニークなオリゴヌクレオチドプローブ、例えば、複数のコピーにおける各ユニークなオリゴヌクレオチドプローブを、単一のフィルター又は膜の別個かつ離れた部分又は領域に固定化して、例えば、アレイを形成してもよい。

40

【0061】

このような手法を組み合わせ使用してもよい。例えば、いくつかのユニークなプローブを各々固定化するいくつかの固体担体を使用してもよい。

【0062】

50

用語「固体担体」は、疎水性、イオン性又は共有結合の架橋によりオリゴヌクレオチドを結合できる、任意の固体材料を意味する。

【0063】

本明細書で使用される用語「固定化」は、このような結合によって前記固体担体へのプローブの可逆的又は不可逆的結合を意味する。可逆的の場合には、プローブは、本発明の方法を実施するのに十分な時間、固体担体と結合したままである。

【0064】

本発明によれば固定化部分として好適な多数の固体担体は、当該技術分野において周知であり、文献に広く記載されている。一般的に言えば、固体担体は、化学的又は生化学的方法において、固定化、分離などに現在広く使用されるか又は提案されている周知の担体又はマトリックスのいずれかでよい。このような材料には、合成有機ポリマー、例えば、ポリスチレン、ポリ塩化ビニル、ポリエチレン；又はニトロセルロース及び酢酸セルロース；又はトシル活性化表面；又はガラスもしくはナイロン、又は核酸の共有結合に好適な基を担持するいずれかの表面などがあるが、これらには限定されない。該固定化部分は、例えば、高分子材料、例えば、アガロース、セルロース、アルギナート、テフロン(登録商標)、ラテックス、ポリスチレン又は磁気ビーズからできている、粒子、シート、ゲル、フィルター、膜、マイクロファイバーストリップ、チューブ又はプレート、繊維又はキャピラリーの形態をとることができる。好ましくは一次元での配列の形態を可能とする固体担体、例えば、シート、フィルター、膜、プレート又はバイオチップが好ましい。

10

【0065】

固体担体への核酸分子の結合を、直接的又は間接的におこなうことができる。例えば、フィルターを使用する場合には、結合は、UV誘発架橋によりおこなうことができる。別法として、結合を、結合部分をオリゴヌクレオチドプローブ及び/又は固体担体上に担持して使用することにより間接的におこなってもよい。したがって、例えば、一对のアフィニティ結合パートナー、例えば、アビジン、ストレプトアビジンもしくはビオチン、DNAもしくはDNA結合タンパク質(例えば、lacIレプレッサータンパク質又はそれが結合するlacオペレータ配列のいずれか)、抗体(モノクローナル又はポリクローナルでよい)、抗体断片又は抗体のエピトープ又はハプテンを使用してもよい。これらの場合、結合対の一つのパートナーを固体担体に結合するか、又はその一つのパートナーが固有的に固体担体の一部分であり、他のパートナーを核酸分子に結合するか、又は他のパートナーが固有的に核酸分子の一部である。

20

30

【0066】

本明細書で使用される用語「アフィニティ結合対」とは、特異的に互いを認識しかつ結合する(すなわち、他の分子への結合に優先して)2つの成分を意味する。このような結合対は、互いに結合したときに、複合体を形成する。

【0067】

固体担体への適切な官能基の結合は、当該技術分野において周知の方法によりおこなうことができる。このような方法には、例えば、固体担体を処理して好適な表面塗膜を得ることにより形成できる水酸基、カルボキシル基、アルデヒド基、アミノ基を介した結合が含まれる。結合パートナーの結合に適当な部分を与える固体担体は、当該技術分野において公知の通常の方法により製造できる。

40

【0068】

本発明のオリゴヌクレオチドプローブへの適当な官能基の結合は、ライゲーションによりおこなうか、又は適当な部分、例えば、ビオチン又は特定の捕捉配列を担持したプライマーを使用した合成又は増幅中に導入される。

【0069】

上記したプローブの組は、キットの形態で提供するのが都合がよい。

【0070】

したがって、本発明のさらなる態様によれば、一種又はそれ以上の固体担体上に固定化された上記のような一組のオリゴヌクレオチドプローブを含むキットが提供される。

50

## 【0071】

好ましくは、前記プローブ類を単一の固体担体上に固定化し、各ユニークなプローブを該固体担体の異なる領域に結合させる。しかしながら、複数の固体担体に結合させたとき、前記複数の固体担体は、キットを構成するモジュールを形成する。とりわけ好ましくは、前記担体は、シート、フィルター、膜、プレート又はバイオチップである。

## 【0072】

必要に応じて、キットは、正常試料又は疾病試料（該キットの使用に関して以下で詳細に説明する）、標準化材料、例えば、比較用の正常試料及び/又は疾病試料からの mRNA 又は cDNA、cDNA への取込み用標識、増幅用核酸配列導入用アダプター、増幅用プライマー及び/又は適当な酵素、バッファー及び溶液により生成されるシグナルに 10  
関係する情報を含んでいてもよい。必要に応じて、前記キットは、添付文書を含んでいてもよい。この添付文書には、本発明の方法をどのように実施するかが記載されており、本発明を実施したときに得られる標準的なグラフ、データ又は結果を解釈するためのソフトウェアが必要に応じて付けられている。

## 【0073】

後記の標準的な診断遺伝子転写物パターンを作成するこのようなキットの使用は、本発明のさらなる態様を構成する。

## 【0074】

本明細書に記載のプローブの組には、種々の用途がある。しかしながら、主にこれらは、試験用細胞の遺伝子発現状態を評価して、前記細胞を得る生物に関係した情報を得る 20  
ことに使用される。したがって、プローブは、生物における癌又は癌の病期を診断、同定又はモニタリングするのに有用である。

## 【0075】

したがって、本発明のさらなる態様によれば、上記の一组のオリゴヌクレオチドプローブ又はキットの使用であって、上記オリゴヌクレオチドプローブが結合する遺伝子の遺伝子発現のレベルを反映する細胞の遺伝子発現パターンを決定するための使用が提供される。この使用は、少なくとも

a) 必要に応じて cDNA に逆転写してもよい、前記細胞から mRNA を単離する工程と；

b) 工程 (a) の mRNA 又は cDNA を、本明細書に記載の一组のオリゴヌクレオチドプローブ又はキットにハイブリダイズさせる工程と； 30

c) 前記プローブの各々にハイブリダイズしている mRNA 又は cDNA の量を評価して前記パターンを生成する工程と、  
を含む。

## 【0076】

この方法及び以下で述べる方法において言及される mRNA 及び cDNA は、前記分子の誘導体又はコピー、例えばそのような分子のコピー、具体的には相補鎖の増幅又は調製により製造されるものであるが、mRNA 配列の同一性が保持されており、すなわち、前記分子の少なくともある領域について相補性又は配列の同一性が高いので、直接に転写物（又はその相補的配列）にハイブリダイズするようなものを含む。転写物をトランケート (truncate) したり、あるいは例えば、プライマー増幅により新しい配列を導入する手法が使用された領域全体にわたっては相補性は存在しないことが認識されよう。便宜上、前記 mRNA 又は cDNA は、工程 b) の前に増幅することが好ましい。本明細書に記載のオリゴヌクレオチドと同様に、前記分子は、例えば、相補性が維持されるならば合成中に非天然塩基を使用して修飾してもよい。また、このような分子は、シグナル伝達手段又は固定化手段などのさらなる部分を担持していてもよい。

## 【0077】

このようなパターンを作成する方法に含まれる種々の工程を、以下詳細に説明する。

## 【0078】

本明細書で使用される用語「遺伝子発現」とは、特定の遺伝子の転写により特異的 m 50

R N A 産物（すなわち、特定のスプライシング産物）が生成することを意味する。遺伝子発現レベルは、転写 m R N A 分子、又は m R N A 分子から逆転写された c D N A 分子、又は例えば、増幅によりこれらの分子から得られた産物のレベルを評価することにより求めることができる。

【 0 0 7 9 】

この手法により得られた「パターン」は、例えば、表の形又はグラフ状に表すことができる情報を言い、二種又はそれ以上のオリゴヌクレオチドと関係するシグナルについての情報を伝達する。好ましくは、前記パターンは、各プローブと関連した発現レベルに関する数のアレイ（array）として表される。

【 0 0 8 0 】

好ましくは、前記パターンは、以下の線形モデルを用いて確定される：

$$y = X b + f \quad \text{式 1}$$

（式中、X は遺伝子発現データのマトリックスであり、y は反応変数であり、b は回帰係数ベクトルであり、f は推定残余ベクトルである。式 1 に表される関係を確定するのに種々の異なる方法を使用することができるが、式 1 の関係を確定するには、部分最小二乗回帰（PLSR）を使用するのがとりわけ好ましい。

【 0 0 8 1 】

したがって、プローブは、細胞が単離される時点での遺伝子発現を反映するパターンを生成するのに使用される。その発現パターンは、細胞を取り巻く周囲の状況に特徴的なものであり、細胞に及ぼす影響により異なる。したがって、特定の癌を有する個体からの細胞について特徴的な遺伝子転写パターンの標準又はフィンガープリント（標準プローブパターン）を作成して、試験細胞の転写パターンとの比較に使用されてもよい。これは、生物が特定の癌を患っているか又は癌の特定の病期にあるかどうかを診断、モニタリング又は同定することに用いることができるのは明らかである。

【 0 0 8 2 】

標準パターンは、癌を有するか、その病期にある一つ又はそれ以上の生物から得た試料の細胞についての総 m R N A （又は c D N A 又は関連産物）のプローブへの結合の程度を求めることにより作成される。これは、ユニークな各プローブに対応し、存在する転写物レベルを示す。異なるプローブに結合する核酸物質の量を評価し、この情報を一緒にして、その癌又は癌の病期に関する遺伝子転写パターン標準を形成する。このような標準パターン各々は、癌又は癌の病期に特徴的なものである。

【 0 0 8 3 】

したがって、本発明のさらなる態様によれば、生物における癌又は癌の病期に特徴的な標準遺伝子転写パターンを作成する方法であって、少なくとも

a ) 必要に応じて c D N A に逆転写してもよい、癌又は癌の病期にある一つ又はそれ以上の生物の試料細胞から m R N A を単離する工程と；

b ) 工程 ( a ) の m R N A 又は c D N A を、調査中の生物及びその試料に対応する生物及びその試料中の前記癌又は癌の病期に特異的な前記一組のオリゴヌクレオチド又はキットにハイブリダイズさせる工程と；

c ) 前記プローブの各々にハイブリダイズしている m R N A 又は c D N A の量を評価して、前記癌又は癌の病期の試料において、前記オリゴヌクレオチドが結合する遺伝子の遺伝子発現のレベルを示す特徴的なパターンを作成する工程と；

を含む方法が提供される。

【 0 0 8 4 】

前記オリゴヌクレオチドは、一種又はそれ以上の固体担体上に固定化されているのが都合がよく、好ましい。

【 0 0 8 5 】

特定のプローブを用いた、極めて多数の癌及び癌の異なる病期についての標準パターンをデータベースに蓄積して、要望のあり次第、検査室で利用できるようにしてもよい。

【 0 0 8 6 】

10

20

30

40

50

本明細書における「疾病」の試料及び生物、あるいは「癌」の試料及び生物は、例えば、腫瘍などの固形塊において異常細胞が増殖している生物（又は前記生物からの試料）を意味する。このような生物は、調査中の癌又は癌の病期を有する又は示すことが知られているものである。

【0087】

癌の「病期」は、特定の生理的又は代謝的变化を示しても、示さなくてもよいが、遺伝子発現の変化として検出できる遺伝子レベルでの変化を示す、癌の種々の病期を意味する。癌の進行中、種々の転写物の発現が異なってもよい。したがって、種々の病期での発現変化は、特定の転写物について「正常」試料と比較して示さなくてもよい。しかしながら、癌の進行を通して一つ又はそれ以上の病期で発現変化を示すいくつかの転写物から得た情報を組み合わせて使用すると、癌の特定の病期を示す特徴的なパターンを得ることができる。したがって、例えば癌の種々の病期、例えば、前段階I、病期I、病期II、病期III、又は病期IVを識別できる。

10

【0088】

本明細書で使用される用語「正常」とは、比較の目的に使用される生物又は試料を意味する。好ましくは、これらは、特にこれらが正常標準として使用される癌に関する遺伝子発現に影響を及ぼしそうな、いかなる疾病又は状態の兆候を示さないか、あるいはそのような疾病又は状態を有するとは思われない意味において「正常」のものである。しかしながら、癌の種々の病期を比較してもよく、そのような場合における「正常」試料は、当該癌の初期の病期に相当するものでもよい。

20

【0089】

本明細書で使用される用語「試料」は、生物、例えば、細胞を含有する調査中のヒト又はヒト以外の動物から得た材料を意味し、組織、体液又は体内老廃物、あるいは原核生物の場合には、その生物自体を含む。「体液」には、血液、唾液、髄液、精液、リンパ液などがある。「体内老廃物」には、尿、喀痰物（肺疾患患者）、糞などがある。「組織試料」には、バイオプシー、外科的介入又は他の手段、例えば、胎盤により得られた組織などがある。しかしながら、好ましくは、試験試料は、癌に侵されていないと思われる体の領域からのものである。このような試料における細胞は、異常細胞、すなわち、癌細胞ではなく、このような異常細胞と接触状態にあったものではなく、かつ癌部位に由来するものではない。「疾病部位」は、客観的に測定できる方法で疾病、例えば腫瘍を発現する体の領域であると考えられる。したがって、例えば、末梢血を使用して非造血性癌の診断に使用してもよく、その血液は癌からの悪性細胞又は播種性細胞が存在する必要はない。同様に、血液：脳関門のために血液中に異常細胞が見られない脳疾患でも、末梢血は、本発明の方法に使用できる。

30

【0090】

また、標準転写パターンの作成方法及び本発明の他の方法は、真核生物の生きている部分、例えば、培養細胞及び臓器培養ならびに外植片への使用にも適用できることが認識されよう。

【0091】

本明細書で使用される「相当する」試料などは、好ましくは同じ組織、体液又は体内老廃物からの細胞を意味するが、それ以外でもさらに標準又は試験パターンを作成する目的に十分に同じである組織、体液又は体内老廃物からの細胞をも含む。プローブに「対応する」遺伝子に関して使用するとき、これはプローブに対し、配列（相補的でもよい）によって関連づけられる遺伝子を意味するが、プローブは発現の異なるスプライシング産物を反映するものでよい。

40

【0092】

本明細書で使用される用語「評価」は、絶対的又は相対的な観点で測定できる定量的評価及び定性的評価の両方を意味する。

【0093】

本発明は、以下のようにして実施できる。

50

## 【0094】

特定の癌又は癌の病期のための標準転写パターンを作成するために、試料mRNAを、疾病の個体又は生物から公知の手法（例えば、Sambrookら（1989）、Molecular Cloning: A Laboratory Manual（実験マニュアル）、2nd ED.、Cold Spring Harbor Laboratory Press、Cold Spring Harbor、ニューヨーク）により、組織、体液又は体内老廃物の細胞から抽出する。

## 【0095】

RNAを用いて操作することが困難であるので、好ましくはRNAをこの段階で逆転写して第一鎖cDNAを形成する。しかしながら、cDNAのクローニング、あるいはcDNAライブラリーから又はcDNAライブラリーを用いての選択は、本発明におけるこの方法又は他の方法では必要ない。好ましくは、第一鎖cDNAの相補的鎖、すなわち、第二鎖cDNAを合成するが、これは、いずれの対応鎖がオリゴヌクレオチドプローブに存在するかによって決まる。しかしながら、別法として、RNAは、逆転写なしで直接使用することができ、必要に応じて標識することもできる。

## 【0096】

好ましくは、cDNA鎖を、適当なプライマーを使用することにより、ポリメラーゼ連鎖反応（PCR）などの公知の増幅方法により増幅させる。別法として、cDNA鎖を、大腸菌などのバクテリアを形質転換するのに使用されるベクターを用いてクローニングした後、成長させて核酸分子を増殖させてもよい。cDNAの配列が既知のものでない場合には、プライマーを、導入された核酸分子の領域に向けてもよい。したがって、例えば、アダプターを、cDNA分子、及びこれらの部分に向けられたプライマーにライゲーションして、cDNA分子を増幅することができる。別法として、真核生物試料の場合には、RNAのポリAテール及びキャップを利用して適当なプライマーを調製してもよい。

## 【0097】

特定の癌又は癌の病期の標準診断遺伝子転写パターン又はフィンガープリントを作成するために、上記オリゴヌクレオチドプローブを使用して、疾病試料のmRNA又はcDNAを探索することにより、各特定のオリゴヌクレオチドプローブ種、すなわち、ユニークな各プローブにハイブリダイズさせるためのシグナルを生成する。また、必要に応じて、正常試料からのmRNA又はcDNAを用いて、標準対照遺伝子転写パターンを作成してもよい。したがって、mRNA又はcDNAを、適当な条件下でオリゴヌクレオチドプローブと接触させてハイブリダイズさせる。

## 【0098】

複数の試料を探索するときには、例えば、一つ以上の固体担体、すなわち、プローブキットモジュール上で同じプローブを用いて連続的に実施するか、あるいは対応するプローブ、例えば、対応のプローブキットのモジュールに同時にハイブリダイズさせることにより実施できる。

## 【0099】

ハイブリダイゼーションがいつ生じたかを確認し、転写物の数ノ（オリゴヌクレオチドプローブに結合するようになったcDNA分子）という指標を得るためには、転写物（又は関連する分子）がハイブリダイズするとき（例えば、二本鎖核酸分子の検出、又は例えば、洗浄により未結合分子を除去した後に結合した分子の数を検出することにより）に生成するシグナルを確認する必要がある。

## 【0100】

シグナルを得るために、ハイブリダイズする一方又は両方の成分（すなわち、プローブ及び転写物）は、情報伝達手段又はその一部分を担持するか又は形成する。この「情報伝達手段」は、シグナルの生成又は存在により直接的又は間接的に検出できる部分である。該シグナルは、いずれの検出可能な物理的特性、例えば、放射線放出、散乱又は吸収の特性、磁気特性又は他の物理的特性、例えば、存在する分子の電荷、サイズ又は結合特性（例えば、標識）、あるいは生成することがある分子（例えば、ガス放出など）により付

10

20

30

40

50

与されるものでよい。シグナル増幅できる方法、例えば、酵素の触媒作用により単一の活性結合部位から複数のシグナル事象を生成して、複数の検出可能な産物を生成する方法が好ましい。

【0101】

情報伝達手段は、自ら検出可能なシグナルを与える標識であるのが都合がよいことがある。また、これは、cDNA産生中、相補的cDNA鎖の調製、標的mRNA/cDNAの増幅中に組み込まれることがあるか、又は標的核酸分子に直接付加した放射性又は他の標識を使用することによりおこなうのが都合がよい。

【0102】

標識は、転写物/cDNAの存在を直接的又は間接的に検出又は測定することができるものが適当である。このような標識には、例えば、放射能標識、化学標識、例えば、発色団又は蛍光体（例えば、フルオレセイン及びローダミンなどの染料）、又は高電子密度の試薬、例えば、フェリチン、ヘモシアニン又は金コロイドなどがある。代替のものとして、標識は、酵素、例えば、ペルオキシダーゼ又はアルカリ性フォスファターゼでもよい。この場合、酵素の存在は、好適な実体物、例えば、基質との相互作用により可視化される。また、標識は、情報伝達ペアの一部分を形成してもよい。この場合、このペアの他のメンバーは、転写物/cDNAが結合するオリゴヌクレオチドプローブ上に見られるか、又はオリゴヌクレオチドプローブの近くに見られるものである。例えば、蛍光性化合物及びクエンチ蛍光性基質を使用できる。また、標識を、抗体などの異なる実体物上に設けることもできる。この異なる実体物は、転写物/cDNAに結合させた、例えば、合成又は増幅中に使用された塩基に結合させたペプチド部分を認識する。

【0103】

シグナルは、ハイブリダイゼーション工程の前、間又は後の標識の導入により得ることができる。別法として、ハイブリダイズする転写物の存在は、他の物理的性質、例えば、それらの吸光度により確認できる。この場合、情報伝達手段は、複合体自体である。

【0104】

次に、各オリゴヌクレオチドプローブに関係したシグナルの量を評価する。評価は、定量的でも、定性的でもよく、各プローブへの単一の転写物種（又は関連するcDNA又は他の産生物）の結合、あるいはユニークな各プローブの複数のコピーへの複数の転写物種の結合に基づくものでよい。定量的な結果により、蓄積される癌の転写物フィンガープリントについてのさらなる情報が得られる。このデータは、絶対値（マクロアレイの場合）で表してもよいし、又は特定の標準又は基準、例えば、正常対照試料と比較して求めてもよい。

【0105】

さらに、標準診断遺伝子パターン転写物を、一種又はそれ以上の疾病試料（及び使用する場合には正常試料）を用いて調製し、ハイブリダイゼーション工程を実施することにより遺伝子発現において特定の個人のバラツキの方向に偏らないパターンを得ることができる。

【0106】

特定の生物における特定癌又は癌の病期の同定、診断又はモニタリングの目的のために作成される標準パターン及び標準診断遺伝子転写物パターンの作成における上記プローブの使用は、本発明のさらなる態様を構成する。

【0107】

選択されたオリゴヌクレオチドプローブを用いて特定癌又は癌の病期について、標準診断フィンガープリント又はパターンを一旦決定したら、この情報を、異なる試験生物又は個体におけるその癌の有無又は程度又は病期を確認するのに使用できる。

【0108】

試験試料の遺伝子発現パターンを調べるために、標準パターンの作成に使用した試料に相当する細胞を含有する組織、体液又は体内老廃物の試験試料を、調査される患者又は生物から得る。次に、試験遺伝子転写パターンを、標準パターンについて上記した方法で

作成する。

【0109】

したがって、本発明のさらなる態様によれば、試験遺伝子転写パターンの作成方法であって、少なくとも

a) 必要に応じてcDNAに逆転写してもよい、前記試験生物の試料の細胞からmRNAを単離する工程と；

b) 工程(a)のmRNA又はcDNAを、調査中の生物及びその試料に対応する生物及びその試料中の該癌又は癌の病期に特異的な一組のオリゴヌクレオチド又はキットにハイブリダイズさせる工程と；

c) 前記プローブの各々にハイブリダイズしているmRNA又はcDNAの量を評価して、前記試験試料において、前記オリゴヌクレオチドが結合する遺伝子の遺伝子発現レベルを示す前記パターンを作成する工程と；

を含んでなる方法が提供される。

【0110】

次に、この試験パターンを、一つ又はそれ以上の標準パターンと比較して、試料が癌又は癌の病期を有する細胞を含有するかどうかを評価することができる。

【0111】

したがって、本発明のさらなる態様によれば、生物における癌又は癌の病期を診断又は同定又はモニタリングする方法が提供される。この方法は、

a) 必要に応じてcDNAに逆転写してもよい、前記生物の試料の細胞からmRNAを単離する工程と；

b) 工程(a)のmRNA又はcDNAを、調査中の生物及びその試料に対応する生物及びその試料中の該癌又は癌の病期に特異的な前記一組のオリゴヌクレオチド又はキットにハイブリダイズさせる工程と；

c) 前記プローブの各々にハイブリダイズしているmRNA又はcDNAの量を評価して、前記試料において、前記オリゴヌクレオチドが結合する遺伝子の遺伝子発現のレベルを示す特徴的なパターンを作成する工程と；

d) 前記パターンを、前記調査中の生物及び試料に対応する生物からの試料を用いて本発明の方法により作成された標準的な診断パターンと比較して、前記調査中の生物において前記癌又は癌の病期の存在の有無を決定する工程と；

を含む。

【0112】

工程c)まで及びその工程c)を含む方法は、上記した試験パターンの作成である。

【0113】

本明細書で使用される用語「診断」は、生物における癌又はその病期の存在の決定を意味する。「モニタリング」は、特に個体が癌を患っていることが知られているときに癌の程度を確定することを意味し、例えば、癌の治療効果又は進行をモニタリングし、例えば、治療の適切性を決定したり、予後をおこなうことを意味する。

【0114】

癌又は癌の病期が存在するかどうかは、標準パターンと試験試料パターンとの間の相関の度合いを求めることにより決定できる。これには、正常試料及び疾病試料について得られる値の範囲を考慮する必要がある。これは、プローブに結合しているいくつかの代表的な試料についての標準偏差を得て標準を得ることにより確定できるけれども、試験試料が標準に対して密に相関している場合には、単一の試料でも癌を同定するための標準パターンを生成するのに充分であると考えてもよい。試験試料における癌又は癌の病期の有無又は程度は、試験試料における有益なプローブの発現レベルに関係するデータを、式1により確定される標準診断プローブパターンに挿入することにより予測できることが都合がよい。

【0115】

上記した方法を用いて生成したデータは、最も基本的な視覚表示に表すこと(例えば

、強度に関して)から、定量化でき、かつ数学的に表すことができ、種々のプローブが結合する各遺伝子の発現レベルの相互関係を反映した基礎パターンを同定するためのもっと複雑なデータ操作まで、種々の方法を用いて解析できる。このように生成した生データは、以下で記載するデータ処理及び統計的方法、特にデータを正規化及び標準化し、データを類別モデル(classification model)にあてはめて操作し、前記試験データが特定癌又は癌の病期のパターンを反映するかどうかを決定することが都合がよい。

【0116】

本発明の方法は、オリゴヌクレオチドプローブが有用である癌又は癌の病期又は進行の確認、モニタリング又は診断に使用することができる。本発明の「有用な」プローブは、当該癌又はその特定の病期における発現の変化を示す遺伝子を反映するものである。診断目的には本発明のプローブは、単独で使用したときにはその有用性は充分ではないが、例えば、上記の組における特徴的なパターンを得るためのいくつかのプローブのうちの一つとして使用するときには有用である。

10

【0117】

好ましくは、前記プローブは、前記癌又は癌の病期により系統的に影響される遺伝子に対応する。とりわけ好ましくは、本発明のプローブに結合し、転写物が得られる前記遺伝子が、中程度又は高度に発現する。中程度又は高度に発現した遺伝子に対するプローブを使用すると、必要な遺伝子発現データの組を生成するのに必要とする臨床試料が少なくよく、例えば、血液試料が1ml未満でよいという利点がある。

【0118】

さらに、すでに活発に転写しているこのような遺伝子は、新しい刺激によりポジティブ又はネガティブな形により影響されやすいことが分かった。さらに、該転写物は、すでに一般的に検出可能なレベルで生成されているので、これらのレベルの小さな変化は、例えば、検出可能な一定のしきい値に到達する必要がないので容易に検出できる。

20

【0119】

本発明の好ましい方法では、本発明のプローブの組は、種々の異なる癌又は癌の病期に有用である。本発明で開示されるプローブのサブセットは、特定癌又は癌の病期の診断、同定又はモニタリングに使用することができる。

【0120】

前記プローブが診断、同定及びモニタリングに使用できる癌には、胃癌、肺癌、乳癌、前立腺癌、腸癌、皮膚癌、大腸癌及び卵巣癌などがある。とりわけ好ましくは、プローブは、乳癌の解析に用いられる。

30

【0121】

この診断法は、他の診断法の代替法として単独で使用してもよいし、このような方法に加えて使用してもよい。例えば、本発明の方法は具体的に腫瘍の同定及び/又は診断において、画像形成法、例えば、磁気共鳴映像法(MRI)、超音波像形成、核イメージング又はX線イメージングを用いた診断の代替法又はそれに付加した診断方法として使用できる。

【0122】

本発明の方法は、原核生物又は真核生物からの細胞について実施できる。原核生物又は真核生物は、いずれの真核生物、例えば、ヒト、他の哺乳動物及び動物、鳥、昆虫、魚及び植物、及び原核生物、例えば、バクテリアでもよい。

40

【0123】

本発明の方法を実施できる好ましい非ヒト動物には、哺乳動物、特に霊長類、家畜、肉畜及び実験動物などがあるが、これらには限定されない。したがって、診断に好ましい動物には、マウス、ラット、モルモット、猫、犬、豚、牛、ヤギ、羊、馬などがある。特に好ましくは、ヒトの癌を、診断し、同定し、又はモニタリングする。

【0124】

上記したように、調査中の試料は、生物から得ることができるいずれかの都合のよい試料であってもよい。しかしながら、好ましくは上記のように、試料を疾病部位から離れ

50

た部位から得る。このような試料における細胞は、異常細胞ではなく、そのような細胞と接触したことはなく、そして疾病部位からのものではない。このような場合、試料は、これらの基準を満足しない細胞を含有していてもよいが、含有していない方が望ましい。しかしながら、本発明のプローブは、これらの基準を満足する細胞において発現が変化している転写物に関係しているので、たとえ他のバックグランド細胞の存在下であっても、該プローブはこれらの細胞における転写レベルの変化を特異的に検出できるようになっている。

#### 【0125】

このような試料からの細胞は、大多数の遺伝子の遺伝子発現において顕著で有用な変動を示すことが分かった。したがって、同じプローブ（又はいくつかのプローブ）は、そのプローブに結合している転写物の特定のレベル、又は他のプローブに対する当該プローブへの結合の程度という相互関係により、2つ又はそれ以上の癌又は癌の病期に関する測定に有用であることが分かる。その結果として、相対的に少数のプローブを、複数の癌のスクリーニングに使用することができる。このことは、プローブの選択に関して重要であるだけでなく、複数の診断に一組のプローブを使用することが可能となる。

10

#### 【0126】

したがって、本発明によれば、2つ又はそれ以上の癌又は癌の病期を診断、同定又はモニタリングするためのプローブ組の組であって、前記プローブの少なくとも一つは、前記癌又は癌の病期の少なくとも2つを診断し、同定し、又はモニタリングするのに好適であるプローブの組、ならびにそれを用いたキット及び方法が提供される。好ましくは、少なくとも5プローブ、例えば、5～15プローブが少なくとも2つの診断に使用される。

20

#### 【0127】

したがって、本発明のさらなる好ましい態様によれば、生物における2つ又はそれ以上の癌又は癌の病期の診断、同定又はモニタリングのための上記した診断又は同定又はモニタリングの方法であって、前記診断方法の工程c)で生成した前記試験パターンを工程d)で、上記した方法で調製した少なくとも2つの標準診断パターンと比較し、各標準診断パターンが、異なる癌又は癌の病期について生成されたパターンである方法が提供される。

#### 【0128】

好ましい態様によれば、評価方法は、試験試料からの遺伝子転写パターンの作成及びこれと標準パターンとの比較に関係するが、一定マーカーの発現の増強又は低下も、また発現産物の調査及びこれらの産物のレベルの調査により調べられることもできる。したがって、発現産物に関係する標準パターンを生成してもよい。

30

#### 【0129】

このような方法では、一次オリゴヌクレオチド又は誘導オリゴヌクレオチドが結合する遺伝子によりコードされている一組のポリペプチドの発現レベルを解析される。

#### 【0130】

種々の診断方法を使用して、ポリペプチド（又はその断片）の存在量を評価できる。ポリペプチドの存在又は濃度は、例えば、固定化されていてもよい該ポリペプチドに対する結合パートナー（例えば、抗体）を使用することにより調べて、該ポリペプチドを試料から分離した後、そのポリペプチドの量を求めることができる。

40

#### 【0131】

ポリペプチドの「断片」とは、前記ポリペプチドに由来するものであることが認識できて、特定の結合パートナーが結合することができる前記ポリペプチドのドメイン又は領域、例えば、抗原性断片を意味する。好ましくは、このような断片は、前記ポリペプチドの重要な部分を含み、正常な合成後処理の産物に相当する。

#### 【0132】

したがって、本発明のさらなる態様によれば、生物における癌又は癌の病期に特徴的な標準遺伝子転写パターンの作成方法であって、少なくとも

a) 癌又は癌の病期である一種以上の生物の試料から標的ポリペプチドを放出する工

50

程と；

b) 前記標的ポリペプチドを、一つ以上の結合パートナーであって、各結合パートナーが、一次オリゴヌクレオチド（又は誘導配列）が結合する遺伝子によりコードされているマーカポリペプチド（又はその断片）に特異的である結合パートナーと接触させて、前記結合パートナーを前記標的ポリペプチドに結合させる工程であって、前記マーカポリペプチドは、調査中の生物及びその試料に対応する生物及び試料中の該癌に特異的である工程と；

c) 前記結合パートナーに結合している前記標的ポリペプチドを評価して、前記癌又は癌の病期を有する前記試料において前記マーカポリペプチドを発現する遺伝子の遺伝子発現のレベルを反映した特徴的なパターンを作成する工程と；

10

を含む方法が提供される。

【0133】

本明細書で使用される用語「標的ポリペプチド」とは、検出される試料に存在するポリペプチドを意味し、「マーカポリペプチド」とは、一次オリゴヌクレオチド又は誘導オリゴヌクレオチドが結合する遺伝子、すなわち、遺伝子ファミリー内の遺伝子によりコードされているポリペプチドである。標的及びマーカポリペプチドは、同一であるか、又は少なくとも同一性の高い領域、例えば、エピトープ領域を有しており、結合パートナーの認識及び結合が可能である。

【0134】

標的ポリペプチドの「放出」とは、試料に適切な処理をおこなって、例えば、それらが存在する細胞の溶解により、結合パートナーの結合にアクセスできる形態のポリペプチドを得ることを意味する。標的ポリペプチドは細胞から周囲の組織又は液中に放出され得るが、この組織又は液、例えば、尿又は血液を分析できるため、この場合に使用する試料は、必ずしも細胞を含まない。しかしながら、上記した試料を使用するのが好ましい。「結合パートナー」は、上記アフィニティ結合の対と一緒に作る別個の存在物を含み、この結合の対における一つのパートナーは、標的又はマーカポリペプチドであり、他のパートナーは、そのポリペプチド、例えば、抗体に特異的に結合する。

20

【0135】

形成する結合の対の量を検出するのに、種々の配置が可能である。その最も単純な形態では、サンドイッチ型アッセイ、例えば、免疫アッセイ、例えば、ELISAを使用できる。この方法では、該ポリペプチドに特異的であって、標識（本明細書に記載のもの）を担持している抗体を、結合の対（例えば、第一抗体：ポリペプチドの対）に結合させ、標識の量を検出できる。

30

【0136】

本明細書に記載の他の方法は、遺伝子転写物及び関連核酸分子ではなく発現のタンパク質産物の分析のために同様に変更できる。

【0137】

したがって、本発明のさらなる態様によれば、試験遺伝子転写パターンの作成方法であって、少なくとも

a) 前記試験生物の試料から標的ポリペプチドを放出する工程と；

40

b) 前記標的ポリペプチドを、一つ以上の結合パートナーであって、各結合パートナーが、一次オリゴヌクレオチド（又は誘導配列）が結合する遺伝子によりコードされているマーカポリペプチド（又はその断片）に特異的である結合パートナーと接触させて、前記結合パートナーを前記標的ポリペプチドに結合させる工程であって、前記マーカポリペプチドは、調査中の生物及びその試料に対応する生物及びその試料中の該癌に特異的である工程と；

c) 前記試験試料において前記結合パートナーに結合している前記標的ポリペプチドを評価し、前記マーカポリペプチドを発現する遺伝子の遺伝子発現のレベルを反映した特徴的なパターンを作成する工程と；

を含む方法が提供される。

50

## 【0138】

本発明のさらなる態様によれば、生物における癌又は癌の病期の診断又は同定又はモニタリングの方法であって、

a) 前記生物の試料から標的ポリペプチドを放出する工程と；

b) 前記標的ポリペプチドを、一つ以上の結合パートナーであって、各結合パートナーが、一次オリゴヌクレオチド（又は誘導配列）が結合する遺伝子によりコードされているマーカポリペプチド（又はその断片）に特異的である結合パートナーと接触させて、該結合パートナーを前記標的ポリペプチドに結合させる工程であって、前記マーカポリペプチドは、調査中の生物及びその試料に対応する生物及びその試料中の該癌に特異的である工程と；

c) 前記試料において前記結合パートナーに結合している前記標的ポリペプチドを評価し、前記マーカポリペプチドを発現する遺伝子の遺伝子発現のレベルを反映した特徴的なパターンを作成する工程と；

d) 前記パターンを、調査中の生物及び試料に対応する生物からの試料を用いて上記した方法により作成した標準診断パターンと比較して、調査中の前記生物における前記癌又は癌の病期の存在を示す相関度を求める工程と；

を含む方法が提供される。

## 【0139】

標準パターン及び試験パターンの作成方法及び診断法では、有用なオリゴヌクレオチドプローブを使用して遺伝子発現データを得る。場合によっては、特定の手法、例えば、特定の癌を診断するのに有用なこれらのプローブを、利用できるプローブ、例えば、表2及び/又は表3に記載のオリゴヌクレオチド、表2及び/又は表3に記載の誘導オリゴヌクレオチド、それらの相補配列及び機能的に等価なオリゴヌクレオチド並びに必要に応じて、表4に記載のオリゴヌクレオチド、それらの誘導オリゴヌクレオチド、相補配列及び機能的に等価なオリゴヌクレオチドから選択することが必要である。前記誘導オリゴヌクレオチドには、これらの表に記載された配列に相当する遺伝子、例えば、表2、表5、表6（受託番号参照）に記載の遺伝子、又はそれらの相補的配列から誘導したオリゴヌクレオチドなどがある。以下の方法は、このような有用なプローブを同定すること、又はより詳細には本明細書に記載のプローブから好適なプローブのサブセットを選択するのに都合がよい。

## 【0140】

特定の癌又は癌の病期を分析するためのプローブは、当該技術分野において公知の多数の方法により同定できる。これらの方法には、例えば、差分的発現又はライブラリーサブトラクションなどがある（例えば、WO98/49342参照）。PCT/GB03/005102に記載され、また以下でも記載するように、ほとんどの転写物に情報量が多いことに鑑み、出発点として、本明細書に記載の配列ファミリーに相当するmRNA又はcDNA種のランダムなサブセットを単純に分析し、そのサブセットから最も有用なプローブを採取することもできる。以下の方法では、異なる試料から得たmRNA（又は関連分子）が結合した、固定化オリゴヌクレオチドプローブ（例えば、本発明のプローブ）を使用して、どのプローブが特定の種類の癌、例えば、疾病試料を同定するのに最も有用かを確認する。

## 【0141】

固定化プローブは、種々の関連性のない生物又は関連した生物から誘導することができる。ここで必要なことは、固定化プローブは、試験生物における相同的な対応物に特異的に結合するものでなければならないことだけである。また、プローブは、市販又は公開のデータベースから誘導し且つ固体担体に固定化することもできる。選択されたプローブは、本明細書に記載の遺伝子配列ファミリーにおける遺伝子の一つに必然的に相当するが、意図するプローブは、ファミリー群全体の中からランダムに選択することができる。

## 【0142】

固体担体に固定化されたプローブの長さは、標的配列に特定の結合を可能とするのに

10

20

30

40

50

十分な長さでなければならない。固定化プローブは、DNA、RNA又はそれらの修飾産物又はPNA（ペプチド核酸）の形態でよい。好ましくは、固定化されたプローブは、試験生物において高度又は中程度発現の遺伝子であるそれらの相対的対応物に特異的に結合しなければならない。使用されるプローブは、本明細書に記載のプローブであるのが都合がよい。

#### 【0143】

生物試料における細胞の遺伝子発現パターンは、以下で説明するマイクロアレイ又はマクロアレイなどの従来技術を用いるか、又は本明細書に記載の方法を用いて作成できる。現在では、高密度オリゴアレイといった、生物試料において多数の遺伝子の発現レベルを同時にモニタリングするためのいくつかの技術が開発されている（Lockhart 10  
ら、1996、Nat. Biotech., 14, p1675 - 1680）、cDNA microarrays（cDNAマイクロアレイ）（Schenarら、1995、Science、270、p467 - 470）、及びcDNA macroarrays（cDNAマクロアレイ）（Maier Eら、1994、Nucl. Acids Res., 22, p3423 - 3424；Bernardら、1996、Nucl. Acids Res., 24, p1435 - 1442）。

#### 【0144】

高密度オリゴアレイ及びcDNAマイクロアレイにおいて、数百及び数千ものプローブオリゴヌクレオチド又はcDNAが、スライドガラス又はナイロン膜上に斑点状に滴下するか、又はバイオチップ上に合成される。試験試料及び基準試料から単離したmRNA 20  
を、赤色又は緑色の蛍光染料を用いて逆転写により標識し、混合し、マイクロアレイにハイブリダイズさせる。洗浄後、結合した蛍光染料を、レーザーにより検出し、2つの像（各染料について一つ）を得る。2つの像についての得られた赤色スポット及び緑色スポットの比により、試験試料及び基準試料における遺伝子の発現レベルの変化についての情報を得る。別法として、単一のチャンネル又は複数のチャンネルマイクロアレイによる検討を行ってもよい。

#### 【0145】

cDNAマクロアレイでは、ナイロン膜などの固体担体上に、異なるcDNAを斑点状に滴下する。この場合、cDNAの量は、各スポットにハイブリダイズできる試験mRNAの量よりも多くする。試験試料から単離したmRNAを逆転写により放射性標識し、 30  
固定化プローブcDNAにハイブリダイズさせる。洗浄後、固定化プローブcDNAに特異的にハイブリダイズしている標識と関係したシグナルを検出し、定量化する。マクロアレイで得られたデータは、試験試料に存在する転写物の相対的なレベルの情報を含む。マクロアレイが好適であるのは、限定された数の遺伝子の発現を監視することだけであるのに対して、マイクロアレイは数千の遺伝子の発現を同時にモニタリングするのに使用することができ、したがって、大規模に遺伝子発現を検討するのに好ましい。

#### 【0146】

本明細書に記載のプローブ同定法を説明するのに、遺伝子発現データセットを作成するためのマクロアレイ技術を使用している。このために、mRNAを意図する試料から単離し、標識した標的分子、例えば、上記したmRNA又はcDNAを調製する。標識された標的分子を、次に固体担体に固定化したプローブにハイブリダイズさせる。これには上記の種々の固体担体を使用できる。ハイブリダイゼーション後、未結合の標的分子を除去し、固定化されたプローブにハイブリダイズしている標的分子からのシグナルを定量化する。放射性標識を実施する場合、PhosphorImagerを使用して、生のデータセットを得るのに使用できるイメージファイルを作成できる。標的分子を標識するのに選択される標識の特性に応じて、他の機器を使用することもできる。例えば、標識に蛍光を使用するとき、FluorImagerを使用してハイブリダイズされた標的分子からイメージファイルを作成することができる。

#### 【0147】

各スポットにおけるシグナルの平均強度、中央強度又は量に相当する生データは、市 50

販の画像解析用ソフトウェアを使用してイメージファイルから得ることができる。しかしながら、得られたデータは、バックグラウンドシグナルについて補正し、正規化してから解析する必要がある。これは、いくつかの因子が、ハイブリダイゼーションシグナルの質及び量に影響することがあるからである。例えば、単離されたmRNAの質及び量の試料ごとのバラツキ、各反応の間における標的分子の標識効率の微小変動、ならびに異なるマクロアレイ間での非特異的結合の量のバラツキは、すべて得られるデータセットにおけるノイズの一因であり、解析前に補正する必要がある。

**【0148】**

バックグラウンドは、いくつかの方法で補正できる。スポット内の最低ピクセル強度をバックグラウンドサブトラクションに使用することもできるし、あるいはスポット輪郭付近のピクセルのラインの平均値又は中央値をこの目的に使用できる。また、ネガティブコントロールからのシグナルに基づくバックグラウンド強度を表す面積を規定し、この面積の平均強度をバックグラウンドサブトラクションに使用してもよい。

10

**【0149】**

次にバックグラウンド補正データを変換して、データ構造の分散を安定にし、プローブ強度の差について正規化することができる。いくつかの変換法が文献には記載されており、Cui, Kerr and Churchill <http://www.jax.org/research/churchill/research/expressions/Cui-Transform.pdf>に概要が記載されている。各スポットの強度を、マクロアレイにおけるすべてのスポット又はマクロアレイにおける一群のスポットの全体強度、平均強度又は中央強度で割ることにより正規化して、マクロアレイにおける固定化プローブにハイブリダイズしているシグナルの相対強度を得ることができる。遺伝子発現データを正規化するためのいくつかの方法が報告されている(Richmond及びSomerville, 2000, Current Opin. Plant Biol., 3, p108-116; Finkelsteinら, 2001, "Methods of Microarray Data Analysis (マイクロアレイデータ解析法)、CAMDAからの論文、Lin&Johnsom編、Kluwer Academic, p57-68; Yangら, 2001", "Optical Technologies and Informatics (光技術と情報学)", Bittner, Chen, Dorsel & Dougherty編, Proceedings of SPIE, 4266, P141-152; Dudoitら, 2000, J. AM. Stat. Ass., 97, p77-87; Alterら, 2000, supra; Newtonら, 2001, J. Comp. Biol., 8, p37-52)。一般的に、倍率又はスケーリング関数をまず計算して強度効果を補正した後、強度の正規化に使用する。正規化の改善に外部コントロールの使用も示唆された。

20

30

**【0150】**

大規模遺伝子発現解析に伴うもう一つの大きな難題は、異なる時間に実施される実験から採取したデータの標準化である。同一実験で得られた試料についての遺伝子発現データは、バックグラウンド補正及び正規化後に効率的に比較できることを見いだした。しかしながら、異なる時間に実施された実験で得られた試料からのデータには、解析の前にさらなる標準化が必要である。これは、異なる実験間で実験パラメータにおける微妙な差、例えば、異なる時間に抽出したmRNAの品質及び量の差、標的分子の標識に使用される時間、ハイブリダイゼーション時間又は露光時間の差が、測定値に影響するからである。また、調査中の転写物の配列の性質といった因子(それらのGC含量)及びそれらの相対量により、これらが実験プロセスにおける微妙な変動によりどのように影響されるかが決まる。例えば、特定の転写物に対応する第一鎖cDNAが第一鎖合成中にどのように効率的に転写され、かつ標識されるか、あるいはハイブリダイゼーション中に相当する標識標的分子がそれらの相補配列にどのように効率的に結合するかが決まる。また、プリントプロセスにおけるバッチ間の差も、作成された発現データにおけるバラツキについての主要な因子である。

40

50

## 【0151】

これらの影響に適切に相当し、修正できないと、実験シリーズ間の差、すなわち、異なる実験シリーズからの組み合わせたデータ内の差が、遺伝子発現データセットに含まれる意図する主要情報が悪いものとなる。したがって、必要に応じて、発現データは、データ解析前にバッチ調整しなければならない。

## 【0152】

いくつかの試料における多数の遺伝子の発現をモニタリングすると、多量のデータが得られ、複雑すぎて容易には解釈できない。いくつかの管理されていない、及び管理されている多変量データ解析法は、これらの多量のデータセットから有用な生物学的情報を抽出するのに有用であることがすでにわかっている。クラスター分析は、遺伝子発現解析に使用されるいままでも一般的に使用された手法であり、同様の方法で調節される遺伝子を確認するのに行なわれたり、及び/又は遺伝子発現プロファイルを用いた新しい/未知腫瘍クラスを確認するのに実施されてきた (Eisenら、1998、PNAS、95、P14863-14868、上記、Alizadehら、2000、Perouら、2000、Nature、406、p747-752; Rossら、2000、Nature Genetics、24(3)、p227-235; Herwigら、1999、Genome Res.、9、p1093-1105; Tamayoら、1999、Science、PNAS、96、p2907-2912)。

10

## 【0153】

クラスタリング法では、遺伝子を、それらの発現プロファイルに基づいて、2つの基準を満足する機能的カテゴリー(クラスター)にグループ化する: “均質性” - 同じクラスターにおける遺伝子は、発現において極めて類似性が高い; 及び “分離” - 異なるクラスターにおける遺伝子は、発現の類似性は互いに低い。

20

## 【0154】

遺伝子発現解析に使用されてきた種々のクラスタリング法として、例えば、階層クラスタリング(上記、Eisenら、1998; 上記、Alizadehら、2000; 上記、Perouら、2000; 上記、Rossら、2000)、K手段クラスタリング(上記、Herwigら、1999; Tavazoieら、1999、Nature Genetics、22(3)、p.281-285)、遺伝子シェーピング(Hastieら、2000、Genome Biology、1(2)、research0003.1-0003.21)、ブロッククラスタリング(Tibshiraniら、1999、Tech report Univ Stanford.)、格子縞モデル(Lazzeroni、2002 Stat. Sinica、12、p61-86)及び自己組織化マップ(上記、Tamayoら、1999)などがある。また、多変量統計解析の関連法、例えば、特異値分析(Alterら、2000、PNAS、97(18)、p10101-10106; 上記、Rossら、2000)又は多次元スケーリングを用いたものは、調査中の対象物のディメンションを減らすのに効果的である。

30

## 【0155】

しかしながら、クラスター解析及び特異値分析などの方法は、純粋に調査のためであり、データに存在する内部構造の全体像が得られるだけである。これらは、調査中のクラスの性質に関する入手可能情報が解析に使用されない非管理法である。特定の試料がかけられた生物学的摂動(biological perturbation)の性質は、公知である。例えば、遺伝子発現パターンが解析されている試料が疾病個体由来のものであるか、あるいは健康個体由来のものであるかが分かることがある。そうした場合に、判別解析は、試料を遺伝子発現データに基づく種々のグループに類別するのに使用できる。

40

## 【0156】

このような解析では、所定クラスのメンバーと非メンバーとの間を識別できるデータをトレーニングすることにより分級器(classifier)が構成される。次に、トレーニングされた分級器を使用して、未知の試料のクラスを予測できる。文献に記載されている判別法として、例えば、Support Vector Machines(Brownら、

50

2000、PNAS、97、p262-267)、最隣接(Nearest Neighbour)(上記、Dudoitら、2000)、類別ツリー(Classification tree)、(上記、Dudoitら、2000)、ボートドクラスフィケーション(Voted Classification)(上記、Dudoitら、2000)、重み付き遺伝子ボーティング(Weighted Gene Voting)(上記、Golubら、1999)及びベイズ類別(Bayesian Classification)(Kellerら、2000、Tec report Univ of Washington)などがある。PLS(部分最小2乗(Partial Least Square))回帰分析を最初に使用して遺伝子発現データセットにおける次元を減少させた後に、ロジスティック判別分析解析及び二次判別解析(LD及びQDA)を用いる類別がなされる手法も、最近報告されている(Nguyen&Rocke、2002、Bioinformatics、18、p39-50及び1216-1226)。

#### 【0157】

遺伝子発現データが従来の判別法に及ぼす困難は、発現が解析する遺伝子の数が、解析している試料の数と比較して極めて多いことである。しかしながら、ほとんどの場合において、判別解析の問題に有用であるのは、これら遺伝子のほんの少しの部分に過ぎない。さらに、無関係の遺伝子からのノイズが有効な遺伝子からの情報をマスクし又はゆがめる恐れがある。マイクロアレイ研究に有用である遺伝子を確認及び選択するのに有効であるいくつかの方法が文献、例えば、t統計量(Dudoitら、2002、J. AM. Stat. Ass.、97、p77-87)、分散分析(Kerrら、2000、PNAS、98、p8961-8965)、近接解析(Neighbourhood Analysis)(上記、Golubら、1999)、群間：群内平方和の比(Ratio of Between Groups to Within Groups Sum of Squares)(上記、Dudoitら、2002)、ノンパラメトリックコアリング(Non Parametric Scoring)(Parkら、2002、Pacific Symposium on Biocomputing、p52-63)及び尤度選択(Likelihood Selection)(上記、Kellerら、2000)において示唆された。

#### 【0158】

本明細書に記載の方法において、正規化及び標準化された遺伝子発現データは、部分最小2乗回帰(PLSR)を用いることにより解析される。PLSRは主に連続量データの回帰分析に使用される方法(補遺A参照)であるけれども、バイナリーコードに基づくダミー応答マトリックス(dummy response matrix)を用いたモデル構築及び判別分析のための方法としても利用できる。クラスの割り当ては、単純な二分識別、例えば、乳癌(クラス1)/健康(クラス2)に基づくもの、又は複数疾病診断、例えば、乳癌(クラス1)/卵巣癌(クラス2)/健康(クラス3)に基づく複数識別に基づくものである。判別用疾病リストは、他の癌又は癌の病期に対応する、入手可能な試料に応じて増加できる。

#### 【0159】

類別法として適用されるPLSRは、PLS-DA(DAは、判別解析を意味する)と称される。PLS-DAは、Yマトリックスがn行(試料数に相当する)及びK列(クラス数に相当)を含むダミーマトリックスであるPLSRアルゴリズムの延長である。Yマトリックスは、1を第k列に挿入し、-1を他のすべての列に挿入することにより構成される(Xの相当する第i番目の対象がクラスkに属する場合)。YをX上に回帰することにより、新しい試料の分類は、適合式 $f(x) = (Y_1(x), Y_2(x), \dots, Y_k(x))$ の最大成分に相当するグループを選択することにより達成される。したがって、-1/1応答マトリックスにおいて、0未満の予測値は、試料が-1としたクラスに属することを意味し、一方0を超える予測値は、試料が1としたクラスに属することを意味する。

#### 【0160】

10

20

30

40

50

PLSR-DAには、得られる結果を、2つの異なるプロットの形態、すなわち、スコアプロット及びローディングプロットの形態で容易に表すことができるという利点がある。スコアプロットは、主要な成分上への試料の投影を表し、類別モデルにおける試料の分布及びそれらの互いの関係を示す。ローディングプロットは、データセットに存在する変数間の相関関係を示す。

【0161】

通常、PLA-DAを、共線データを取り扱えるので、類別問題のための出発点として使用し、PLSRの特質をディメンション減少法として使用することが好ましい。この目的が満足されたら、さらなる情報を抽出するのに有効であることが判明した線形判別分析LDAなどの他の方法を使用することができる (Indahlら、1999、Chem and Intell. Lab. Syst.、49、p19-31)。この手法は、まずPLS-DAを用いてデータを分解した後、スコアベクトル(最初の変数の代わりに)をLDAへの入力として用いる。LDAについての詳細は、Duda and Hart (Classification and Scene Analysis、1973、Wiley、米国)に記載されている。

10

【0162】

モデル構築に続く工程は、モデルバリデーション(model validation)である。この工程は、多変量分析の最も重要な面に含まれると考えられ、構築された校正モデルの「良好性」を試験する。この作業において、クロスバリデーション(cross validation)法が、バリデーションに使用された。この方法において、モデルが残りのデータに基づいて完全クロスバリデーションを用いて構築される間に、各セグメントにおいて1つまたは数個の試料は除外されている。次に、除外された試料は、予測/類別に使用される。単純クロスバリデーションプロセスを数回反復(各クロスバリデーションについて異なる試料を保持する)して、いわゆる二重クロスバリデーション法をおこなう。この方法は、本明細書に記載の実施例のいくつかに示すような限定されたデータ量の場合にうまくいくことが分かった。また、クロスバリデーション工程は数回反復されるため、モデルバイアス及びオーバーフィッティングの危険が減少する。

20

【0163】

校正モデルがいったん構築されかつバリデーションされると、モデルにおける所望の情報を表す最も関係のある発現パターンを示す遺伝子を、本明細書に記載の変数選択についての従来技術に記載されている方法により選択できる。変数選択は、最終的なモデルの複雑さを減少させるのに役立ち、削ぎ落としたモデルが得られ、したがって、予測に使用できる信頼のおけるモデルが実現できる。さらに、診断目的の遺伝子の数が少ないほど、診断産物のコストが減少する。かようにして関連のある遺伝子に結合するであろう有用なプローブを同定できる。

30

【0164】

本発明者らは、校正モデルを構築した後、再サンプリング法に基づくジャックナイフ(Efron、1982、The Jackknife, the Bootstrap and other resampling plans (ジャックナイフ、ブートストラップ及び他の再サンプリングプラン)。Society for Industrial and Applied Mathematics、米国フィラデルフィア)のような統計的手法が、有意な変数(有用なプローブ)を選択又は確認するのに効率的に使用できることを見いだした。PLS回帰係数Bの近似不確定分散は、下式により推定できる。

40

【0165】

【数1】

M

$$S^2B = \sum_{m=1} ((B-B_m)g)^2$$

m=1

50

(式中、

$S_2 B = B$  の推定不確定分散；

$B =$  すべての  $N$  対象を用いたクロスバリデーションランク  $A$  での回帰係数；

$B_m =$  クロスバリデーションセグメント  $m$  において除外された対象を除くすべての対象を用いたランク  $A$  での回帰係数；及び

$g =$  スケーリング係数 (但し、 $g = 1$ )。

【0166】

本発明者らの手法において、ジャックナイフを、クロスバリデーションとともに実施した。各変数について、クロスバリデーションサブモデルにおける  $B$  係数  $B_i$  と総モデルについての  $B_{t_0}$  との間の差を、まず算出する。次に、その差の平方和を、すべてのサブモデルにおいて算出して変数についての  $B_i$  推定値の分散を得る。  $B_i$  の推定値の有意性を、 $t$  検定を用いて計算する。したがって、得られた回帰係数を、2 標準偏差に対応する不確定限界を用いて示すことができ、それから、有意な変数が検出される。

【0167】

この工程の実施又は使用については、市販のソフトウェア (The Unscrambler、CAMO ASA、Norway) で実施されているので、ここではさらに詳細には説明しない。また、ジャックナイフを用いた変数選択についての詳細も、Westad & Martens (2000、J. Near Inf. Spectr.、8、p 117 - 124) において記載されている。

【0168】

以下の手法は、遺伝子発現データセットから有用なプローブを選択するのに使用できる：

a) 一クロスバリデーションセグメント当たり一つのユニークな試料 (データセットに存在する場合にはその反復を含む) を除外する；

b) PLSR - DA を用いて残りの試料について校正モデル (クロスバリデーションセグメント) を構築する；

c) ジャックナイフ基準を用いて、工程 b) におけるモデルについて有意な遺伝子を選択する；

d) 上記 3 つの工程を、データセットにおけるすべてのユニークな試料を一度は除外する (工程 a) で述べたように) まで繰り返す。例えば、データセットに 75 のユニークな試料が存在する場合には、75 の異なる校正モデルを構築して、75 組の異なる有意なプローブ群を得る；

e) 工程 d) で作成した有意なプローブの組において発生頻度基準を用いて最も有意な変数を選択する。例えば、すべての組 (100%) において現れる一組のプローブは、工程 d) において作成した組の 50% にしか現れないプローブよりも有用である。

【0169】

疾病についての有用なプローブを選択したら、最終モデルを作成し、バリデーションする。モデルをバリデーションする最も一般的に使用されている 2 つの方法は、クロスバリデーション (CV ; cross validation) 及び試験セットバリデーションである。クロスバリデーション (交差確認) では、データを、 $k$  個のサブセットに分ける。次に、モデルを  $k$  回トレーニングし、毎回、トレーニングからサブセットのうちの一つを除外する。この場合、除外されたサブセットのみを用いてエラー基準、RMSEP (予測の二乗平均平方根誤差) を算出する。 $k$  が試料サイズと等しい場合、これを、「一除外 (Leave - One - Out)」クロスバリデーションと称する。一つのバリデーションセグメント当たり一つ又は数個の試料を除外することは、種々の実験間の共分散がゼロである場合のみに有効である。したがって、「一回一試料」法は、複製物を含む場合には有効であるとは言えない。これは、複製物の一つのみを除外することは、本発明者らの解析において組織的バイアスが導入されるからである。この場合の正しい手法は、CV セグメント間の共分散がゼロであるという前提を満足することから、一度に同じ試料のすべての複製物を除外することである。

10

20

30

40

50

## 【0170】

モデルバリデーションの第二の方法は、校正モデルをバリデーションするために別個の試験組を使用するものである。これには、別個の一組の実験を試験の組として実施する必要がある。これは、実地試験データが得られることを前提として好ましい方法である。

## 【0171】

次に、最終モデルを、試験試料における癌又は癌の病期を同定するのに使用する。このために、選択された有用な遺伝子の発現データを試験試料から作成した後、最終モデルを使用して、試料が疾病クラス又は非疾病クラスであるかどうか、あるいは癌又は癌の病期を有するかどうかを判定する。

## 【0172】

好ましくは、類別を目的としたモデルを、上記の方法にしたがって同定したプローブに関するデータを用いて作成する。好ましくは、試料は、上記した試料である。好ましくは、工程(a)で固定化したオリゴヌクレオチドを、上記したファミリー内からランダムに選択する。別法として、例えば、異なるファミリーにおいて共通機能を有するタンパク質をコードしている遺伝子に相当するオリゴヌクレオチドの一つ以上を選択することにより、異なるファミリーを代表するように選択してもよい。とりわけ好ましくは、ファミリー(i)及び(ii)の遺伝子に由来するオリゴヌクレオチドを含むように選択される。このようなオリゴヌクレオチドは、例えば、cDNA(用語「オリゴヌクレオチド」の範囲に含まれる)を用いる場合には、かなりの長さを有している。有用なプローブとしてこのようなcDNA分子を同定することにより、cDNA分子の特異性を反映するが、製造及び取り扱いが容易であるもっと短いオリゴヌクレオチドの開発が実現できる。

## 【0173】

次に、上記モデルを使用して試験試料のデータを作成及び解析でき、したがって、本発明の診断法に使用できる。このような方法において、試験試料から作成したデータにより、遺伝子発現データセットが得られ、上記した方法で正規化及び標準化する。これを、上記した校正モデルにフィッティングして分類する。

## 【0174】

ここで説明する方法は、数種の癌について有用なプローブを同時に選択するのに使用することもできる。いずれの癌が校正又はトレーニングセットに含まれているかによって、有用なプローブを、前記癌用を選択できる。一つの癌について選択された有用なプローブは、意図する別の癌について選択された有用なプローブと同一でも、同一でなくてもよい。それは、選択された遺伝子が癌又は癌の病期中に互いの関係で発現するパターンであり、癌又は癌の病期に有効であるかどうかを決定するパターンである。

## 【0175】

換言すれば、有用な遺伝子を、これらの発現が調査中の癌又は癌の病期により生成される応答の影響下で他の選択された有用な遺伝子の発現とどのように相関しているかに基づいて選択する。

## 【0176】

有用なプローブを単離したり、いくつかの癌及び癌の病期を同時に同定するためには、遺伝子発現データセットは、被験者が調査中の特定の癌又は癌の病期を有するときどのように遺伝子が発現されるかについての情報を含んでいなければならない。データセットは、一組の健康又は疾病の試料から作成する。ここで、特定の試料は、一つの癌又は癌の病期についての情報のみを含んでいてもよいし、あるいは複数の癌又は癌の病期についての情報を含んでいてもよい。したがって、本方法は、複数の癌又は癌の病期を表す試料を選択することにより、有用なプローブを単離するのに必要とされる試料の数を減少させる効率的な実験構成も教示している。

## 【0177】

上記したように、ほとんどの転写物が高情報量を有していることから、特定の癌又は癌の病期を診断、モニタリング又は同定するのに使用される有用なプローブを同定及び選択することが大幅に簡略化されるかも知れない。したがって、有用なプローブを同定する

10

20

30

40

50

のに選択される遺伝子のプールを、大幅に減少できる。

【0178】

細胞内で発現されている数千の遺伝子の集団から有用なプローブを選択する従来技術とは異なり、マイクロアレイのように、本明細書に記載の方法においては、有用なプローブを、上記遺伝子配列ファミリーに記載されたような限られた数の遺伝子から選択する。これらのファミリー内から、意図するプローブをランダムに選択できる。

【0179】

したがって、好ましい態様によれば、上記オリゴヌクレオチドの組を、上記の一次オリゴヌクレオチドからランダムに選択する。

【0180】

本明細書で使用される用語「ランダム」は、調査中の癌又は生物と関係する転写物により担持されている情報の程度に基づくバイアスしない、すなわち、有用なプローブとしてのそれらの有用可能性の方にバイアスしない選択を意味する。ランダムな選択は、例えば、高度又は中程度発現の転写物にバイアスした転写物（又は関連産物）のプールからおこなわれるが、好ましくはランダムな選択は配列を基にした基準によりバイアス又は選択しない転写物プールからおこなう。したがって、プールが大きいと、高度及び中程度発現の遺伝子に対応するオリゴヌクレオチドを含む。別法として、高度及び中程度に発現する遺伝子に対応するものについて富化してもよい。

【0181】

高度及び中程度に発現された遺伝子からのランダムな選択は、種々の方法でおこなうことができる。例えば、有意な数のcDNAクローンを、上記した遺伝子配列ファミリーに対応するクローンを含む、調査中の生物学的検体から構成されたcDNAライブラリーからランダムに採取することによりおこなうことができる。cDNAライブラリーにおいて、多量又は中程度の量で存在する転写物に対応するcDNAクローンは、少ない量で存在するcDNAに対応する転写物よりも存在する頻度が大きいので、採取頻度がより大きい傾向がある。高度又は中程度発現の遺伝子に相当するものについて富化したcDNAのプールは、この手法により単離できる。

【0182】

本発明の方法で使用するために単離された集団のうちの多量又は中程度の量で発現する遺伝子を同定するために、意図する試料におけるそれらの転写物の相対レベルについての情報を、いくつかの従来技術を用いて作成できる。この目的には、ディファレンシャルディスプレイ又はRNAフィンガープリントといった配列に依存しない方法と、マイクロアレイ又はマクロアレイなどの配列に基づく方法の両方を使用できる。別法として、高度及び中程度発現の遺伝子についての特異的プライマー配列を構成してもよく、定量的RT-PCRなどの方法を使用して高度及び中程度発現の遺伝子のレベルを決定することもできる。したがって、当業者は、生物学的試料においてmRNAの相対レベルを決定するのに当該技術分野において公知である種々の方法を使用することができる。

【0183】

とりわけ好ましくは、上記方法におけるmRNAの単離のための試料は、上記したようなものであり、好ましくは疾病部位からのものではなく、前記試料中の細胞は疾病細胞ではなく、また疾病細胞と接触したものでなく、例えば、非造血性癌、例えば、乳癌の検出に末梢血試料を使用する。

【0184】

以下、実施例を、添付図面を参照しながら説明するが、本発明はこれらの実施例には限定されない。

【実施例】

【0185】

実施例1：乳癌の診断

方法

血液試料

10

20

30

40

50

血液試料を、ノルウェーの Regional Ethical Committee の承認の下に同意したドナーから採取した。すべてのドナーは、分析の間は匿名で処理した。血液を、乳房撮影において初期異常の疑いのある女性から採取した。これらの女性には、乳癌の女性及び乳房撮影において異常のある女性（一次スクリーニング中に観察された異常が良性か又は悪性であるのかについてわかる前）の両方が含まれていた。すべての場合において、血液試料を、午前8時～午後4時の間に採取した。各女性から、血液10mlを、熟練者により、抗凝固剤としてEDTA（Becton Dickinson、米国 Baltimore 社製）が入ったバキュテナーチューブに採取するか、又は PAXgene（登録商標）チューブ（PreAnalytiX、スイス国 Hombrechtikon 社製）に直接採取した。EDTAチューブに採取した血液を直ちに -80 で保存し、一方 PAXチューブは一晩放置した後、使用するまで -80 で保存した。

#### 【0186】

##### cDNAアレイの調製

cDNAの1435クローンを、550人の健康個人の全血から構成されたプラスミドライブラリー（米国 Clontech、米国 Palo Alto 社製）からランダムに採取した。ランダムに採取したクローンのうちの約20%は、重複するものであった。インサートの増幅のために、バクテリアクローンを、カルベニシリン50µg/mlとともにLB150µlを含有するマイクロタイタープレートで増殖させ、37 で攪拌しながら一晩インキュベーションした。細胞を溶解するために、各培養5µlをH<sub>2</sub>O 50µlで希釈し、95 で12分間インキュベーションした。この混合物のうち2µlを、1.5mM MgCl<sub>2</sub>の存在下で5' - 及び3' - 配列決定プライマー40µpmolを用いてPCR反応させた。PCR反応は、以下のサイクルプロトコルにより実施した：RoboCycler Temperature Cycler（登録商標）（米国 La Jollaにある Stratagene 社製）又はDNA Engine Dyad Peltier Thermal Cycler（米国 Walthamにある MJ Research 社製）において、95 で4分保持後、94 で1分間保持と60 で1分間保持と72 で3分間保持とからなる操作を25サイクル反復した。増幅物を、NaOH（0.2M、最終濃度）で30分間変性し、HYBOND-N+膜（英国 Little Chalfontにある Amersham Pharmacia Biotech 社製）上に、製造業者による使用説明書（英国 Cambridgeにある BioRobotics 社製）に基づいて MicroGrid IIワークステーションを用いてスポット状に滴下した。固定化cDNAを、UVクロスリンカー（米国 San Franciscoにある Hoefer Scientific Instruments 社製）を用いて固定した。

#### 【0187】

プリントアレイは、1435のcDNAの他に、アッセイのバックグラウンドレベル、コンシステンシー及び感度を評価するための対照も含んでいた。これらを複数の位置にスポット状に滴下した。これらには、PCRミックス（インサートなし）などの対照；Spot Report（商標登録）10アレイバリデーションシステム（米国 La Jollaにある Stratagene 社製）の対照、ならびに -アクチン、 -アクチン、GAPDH、HOD及びシクロフィリンといった、構成的に発現する遺伝子に相当するcDNAが含まれていた。

#### 【0188】

##### RNA抽出、プローブ合成及びハイブリダイゼーション

EDTAチューブに採取した血液を37 で解凍し、PAXチューブに移し、総RNAを、業者の取扱説明書（スイス国 Hombrechtikonにある PreAnalytiX 社製）に準じて精製した。PAXチューブに直接採取した血液から、総RNAを、上記チューブにおいて新しいチューブに移すことなく抽出した。単離したRNAから汚染DNAを、DNAフリーキット（米国 Austinにある Ambion 社製）を用いてD

RNAase I 処理により除去した。RNAの品質を、アガロースゲル電気泳動後の28S及び18Sリボソームバンドの一体性を肉眼で観察することにより決定した。良品質のRNAが抽出された試料のみを、この試験に使用した。本発明者らの経験によれば、EDTAチューブに採取した血液はRNAの品質がよくないことがしばしばあったが、一方PAXチューブに採取したものは、RNAの品質がほとんどいつも良好であった。抽出されたRNAの濃度及び純度を、260nm及び280nmでの吸光度を測定することにより求めた。総RNAから、業者の取扱説明書(ノルウェー国OsloにあるDyna1 AS社製)に準じてDynabeadsを用いてmRNAを単離した。

#### 【0189】

標識化及びハイブリダイゼーションについての実験を、16バッチで実施した。各バッチで評価した試料数は、6~9であった。プリントにおけるバッチ間のバラツキによるノイズを最小とするために、同じプリント実験中に製造されたアレイのみを、各バッチで使用した。試料を複数回(反復試験)評価したとき、同じmRNAプールからのアリコットをプローブ合成に使用した。プローブ合成では、総RNA4~5µgに相当するmRNAのアリコットを、oligo dT<sub>25</sub>NV(0.5µg/µl)及びSpot Report(商標登録)10アレイバリデーション・システムのmRNAスパイク(10pg; Spike 2、1pg)と混合し、70℃に加熱した後、氷冷した。プローブを、反応混合物35µlで、50µCi [<sup>33</sup>P] dATP、3.5µMdATP、0.6mM dCTP、0.6mM dTTP、0.6mM dGTP、SuperScript逆転写酵素(Invitrogen、Life Technologies社製)200単位及び0.1M DTT標識の存在下、42℃で1.5時間逆転写することにより調製した。合成後、酵素を、70℃で10分間失活し、mRNAを、RiboH(米国MadisonにあるPromega社製)4単位中、37℃で20分間反応混合物をインキュベーションすることにより除去した。取り込まれなかったヌクレオチドを、Probe Quant G50 Columns(米国PiscatawayにあるAmersham Biosciences社製)を用いて除去した。

#### 【0190】

膜を、4xSSC中、室温で2時間平衡化させ、プレハイブリダイゼーション溶液(4xSSC、0.1M NaH<sub>2</sub>PO<sub>4</sub>、1mM EDTA、8%デキストラン硫酸、10xデンハート液、1%SDS)10ml中、65℃で一晩、予備ハイブリダイズさせた。新しく調製したプローブを、同じプレハイブリダイゼーション溶液5mlに添加し、ハイブリダイゼーションを65℃で一晩継続した。膜を、ストリンジェンシーを増加して65℃(Cで洗浄した(2xSSC、0.1%SDS; 1xSSC、0.1%SDS; 0.1xSSC、0.1%SDSで各々2x30分間))。

#### 【0191】

##### ハイブリダイゼーションシグナルの定量化

ハイブリダイズさせた膜を、Phosphoscreen(超解像)に2日間暴露し、PhosphoImager(Cyclone、米国MeridenにあるPackard社製)を用いてイメージファイルを生成した。ハイブリダイゼーションシグナルの同定及び定量化、さらに局部バックグラウンド値のサブトラクションを、Phoretix Software(英国Non Linear Dynamics社製)を用いて実施した。バックグラウンドサブトラクションでは、各スポット輪郭周囲のピクセルのラインの中央値を、各スポットで評価されたシグナルの強度から減算した。

#### 【0192】

##### データ解析

1435のバックグラウンド減算発現データから、67の遺伝子のシグナルを各膜から除去して分散度の高い発現遺伝子を排除した。これらには、各膜から最低シグナル及び最高シグナルの1.25%を除去することが含まれていた。正規化のために、各スポットの値を、まず各アレイにおけるシグナルの平均で除した後、すべてのスポットについて立方根変換をおこなった。次に、正規化データを、一元配置分散分析(ANOVA)を用い

てバッチ調整した。

【0193】

その後、予備処理データを使用して、

a) クロスバリデーションセグメント当たり一つのユニークな試料(選択された試料のすべての反復を含む)を除外するクロスバリデーションPLSRモデルを構築すること、  
b) ジャックナイフ基準を用いて、工程a)におけるモデルについて有意の遺伝子の組を選択すること、

c) 工程b)で選択した遺伝子を用いて、工程a)のようにしてクロスバリデーションPLSR-DAモデルを構築すること、

d) 再びジャックナイフ基準を用いて、工程c)におけるモデルについて最も有意な遺伝子の組を選択すること、  
により、有用なプローブを単離した。

【0194】

工程b)で、125の遺伝子が得られた。

【0195】

工程d)で、35の有意な遺伝子が得られた。これらの遺伝子により、最終類別モデルを構築した。

【0196】

発生基準に基づいて選択された有用なプローブを使用して類別モデルを構築した。35のプローブに基づく類別モデルについての結果を、図2に示す。図2において、これらの遺伝子の発現パターンにより、乳癌を患っているほとんどの女性と乳癌のない女性を異なるグループに類別できたことが分かる。この図において、PC1及びPC2は、データに存在する系統的パラツキを最もよく定義しているデータから統計的に得た2つの主要な成分を示す。これにより、各試料、及び試料の標識化第一鎖cDNAが結合された有用プローブの各々からのデータを、主要成分-スコアプロット上の試料投影である単一点として類別モデル上に表すことができる。

【0197】

図3は、35の有意な遺伝子を用いた予測プロットを示す。図示された予測プロットにおいて、癌試料はx軸の+1であり、非癌試料は-1である。y軸は、属するクラスの予測を示す。予測の間、予測が正しい場合には、癌試料はゼロを超え、非癌試料はゼロ未満となるはずである。各例において、ほとんどすべての試料が正しく予測されている。クロスバリデーションでは、102個の実験試料を、各セグメントが一つのユニークな試料を表しており、かつその複製物(存在する場合)を含む60のクロスバリデーションセグメントに分割した。

【0198】

ほとんどの乳癌細胞を正しく予測することができた。22人の癌患者のうち19人が正しく予測され、同様に35人の正常患者のうち34人が正しく予測された。試験された個人の全詳細及び予測の精度を、表1に示す。表2に、35の有用な遺伝子の詳細を示す。これらの遺伝子は、公開データベースにおける遺伝子であり、配列の類似性及び推定生物学的機能が示されている。これらの配列を例の後に示す。

【0199】

図4は、35個の遺伝子の発現レベルを示す。この図から、正常患者における発現に対して一部が過剰発現を生じ、他のものが発現不足を生じることが分かる。

【0200】

実施例2：さらなる有用なプローブの同定と乳癌の診断における使用

方法

使用した同定と解析の方法は、試料を、cDNAアレイを調製する代わりに、大規模遺伝子発現解析用の市販のプラットフォーム(Agilent 22Kチップ)を用いて解析した以外は、実質的に実施例1と同様であった。

【0201】

合計 122 個 (対照 78 及び乳癌 44) を含む、より多数の試料を解析した。データを、上記の PLSR を用いて解析した。意図する遺伝子を、10 倍クロスバリデーション法により選択した。すなわち、122 個の試料からのデータを、各組が 12 ~ 13 試料を含むようにして 10 組に分けた。9 組について校正モデルを構築し、1 組を除外した。有意の遺伝子を、組み込みモデル (built-in model) についてジャックナイフ法により同定した。これらの工程は、10 組すべてについて、各組が少なくとも一度は除外されるようにして反復した。次に、有用遺伝子を、発生基準の頻度に基づいて同定した。すべての 10 校正モデルにおいて、109 の遺伝子が有用であることが分かった。

#### 【0202】

##### 結果

上記 109 個の遺伝子と 3 つのさらなる遺伝子を使用して、使用試料 122 個の類別を予測した。結果を、下表に示す。

#### 【0203】

##### 【表 1】

試料	試料数	正確に予測	不正確な予測	エラー率
対照	78	67	11	0.14
乳癌	44	26	18	0.41

#### 【0204】

109 個の有用な遺伝子を、3 つのカテゴリ、すなわち、本明細書に記載のファミリー (i) 及び (ii) ならびに他の遺伝子に分けてもよい。表 3 に、相当する遺伝子がファミリー (i) 及び (ii) のカテゴリに属する有用なプローブの詳細と、Agilent によりそのプローブに割り当てられた番号を示す。同様に、表 4 に、相当遺伝子がファミリー (i) 及び (ii) に属するとは思われない有用プローブの詳細を示す。表 5 及び表 6 に、表 3 及び表 4 におけるプローブがそれぞれ配列類似性を示す遺伝子の詳細、それらの推定生物学的機能 (既知の場合)、及びそれらの遺伝子についての受託番号を示す。

#### 【0205】

##### 補遺 A

部分最小 2 乗回帰 (PLSR)

多変量回帰モデルを以下のように定義する：

$$Y = XB + F$$

(式中、

X は、N 予測変数 (遺伝子) の  $N \times P$  マトリックスであり；

Y ( $N \times J$ ) は、J 予測変数である。この場合、Y は、ダミー変数を含むマトリックスを表し；

B は、回帰係数のマトリックスであり；そして

F は、残余の  $N \times J$  マトリックスである。

#### 【0206】

PLSR モデルの構造を以下のように表すことができる：

$$X = TP^T + E_A \text{ 及び}$$

$$Y = TQ^T + F_A \text{ (但し、}$$

T ( $N \times A$ ) は、x 変数の一次結合であるスコアベクトルのマトリックスであり；

P ( $P \times A$ ) は、x ローディングベクトル  $p_a$  を列としたマトリックスであり；

Q ( $J \times A$ ) は、y ローディングベクトル  $q_a$  を列としたマトリックスであり；

$E_A$  ( $N \times P$ ) は、因数 A に続く X についてのマトリックスであり；そして

$F_A$  ( $N \times J$ ) は、因数 A に続く Y についてのマトリックスである。

#### 【0207】

PLSR における基準は、[X、Y] の説明済みの共分散を最大化することである。

10

20

30

40

50

これは、 $E_a^T F_a F_a^T E_a$  ( $E_a$  及び  $F_a$  は、因数  $a$  又は PLS 成分に続く、収縮  $X$  及び  $Y$ ) の第一固有ベクトルであるローディングウエイトベクトル (loading weights vector)  $w_{a+1}$  により達成される。

【0208】

回帰係数は、以下のように表される：

$$B = W (P^T W)^{-1} Q^T$$

【0209】

最大階数、すなわち、最大成分数の PLSR モデルは、MLR 解に等しい。PLSR については、Martens & Naes、1989、Multivariate Calibration (多変量校正)、米国 John Wiley & Sons 社及び同上、Kowalski & Seasholtz、1991 にさらに詳細に説明されている。

## 【数 2】

Jackknife により選択された 34/35 遺伝子のスクレオチド配列

クローン ID とそれらの配列

## I-30

CTTTTCTCCCGCTGTCCCCACGGAGGGGACTGCTCTCCCCCGCTGCATCCTT  
 TCTGTGAGGTACCTTACCCACCTCAGCACCTGAGAGGGTGAAATAGAATTCTAA  
 CCTCGACATTGGGAAGTGTTTTTGAGAAGTCTCGGTCCGTAAGGGAAGTCTTC  
 CAAGTCCGTGCAGCACTAACGTATTGGCACCTGCCTCCTCTCGGCCACCCCC  
 AGATGAGGCAGCTGTGACTGTGTCAAGGGAAGCCACGACTCTGACCATAGTCTT  
 CTCTCAGCTTCCACTGCCGTCTCCACAGGAAACCCAGAAGTTCTGTGAACAAGT  
 CCATGCTGCCATCAAGGCATTTATTGCAGTGTACTATTTGCTTCCAAAGGATCA  
 GGCCCTGAGAACAATGACCTTATTTCTACAACAGTGTCTGGGTTGCGTGCCAG  
 CAGATGCCTCAGATACCAAGAGATAACAAAGCTGCAGCTCTTTGATGCTGACC  
 AAGAATGTGGATTTTGTGAAGGATGCACATGAAGAAATGGAGCAGGCTGTGGAA  
 GAATGTGACCCTTACTCTGGCCTCTTGAATGATACTGAGGAGAACAACCTCTGAC  
 AACCACAATCATGAGGATGATGTGTTGGGGTTTCCCAGCAATCAGGACTTGTAT  
 TGGTCAGAGGACGATCAAGAGCTCATAATCCCATGCCTTGGCTGGTGAGAGCA  
 TCCAAAGCCTGCCTGAAGAAAATTCGGATGTTAGTGGCAGAGAATGGGAAGAAG  
 GATCAGGTGGCACAGCTGGATGACATTGTGGATATTTCTGATGAAATCAGCCCT  
 AGTGTGGATGATTTGGCTCTGAGCATATATCCACCTATGTGTACCTGACCGTG  
 CGAATCAATTCTGCGAAACTTGTATCTGTTTTAAAGAAGGCACCTTGAAATTACA  
 AAAGCAAGTCATGTGACCCCTCAGCCAGAAGATAGTTGGATCCCTTTACTTATT  
 AATGCCATTGATCATTGCATGAATAGAATCAAGGAGCTCACTCAGAGTGAACCT  
 GAATTATGACTTTTCAGGCTCATTGTACTCTCTTCCCCTCTCATCGTCATGGT  
 CAGGCTCTGATACCTGCTTTTTAAAATGGAGCTAGAAATGCTTGCTGGATTGAAAG  
 GGAGTGCCTATCTATATTTAGCAAGAGACACTATTACCAAAGATTGTTGGTTAG  
 GCCAGATTGACACCTATTTATAAACCATATGCGTATATTTTCTGTGCTATATA  
 TGAAAAATAATTGCATGATTTCTCATTCCCTGAGTCAATTTCTCAGAGATTCTAG  
 GAAAGCTGCCTTATTCTTTTTTGACGTAAGTATGTTGTTTTTATTGTAAAGA  
 TGTTGATGGTCTCAATAAAATGCTAACTTGCCAGTGAIAAAAAAAAAAAAAA

10

20

30

## III-02

AGGATCTAAGACCAGCCTGGCAGCCACCAGATGGTGATTCTAGTCCTGGCTCAG  
 TCAGTAATAGGTCACTGACCCAGAGAAATCAATTCAGCCTCCCCAGGTCCTTG  
 GATTTCTTTCTGTGAAAATGAAAGCATAGGTAGGAATTTCCCATGGAACAGCTA  
 GCAGAGGAGAAATATTAAGTCAAGGAGACTCATGCTATAGTTTTTACTACTTCA

40

【数 3】

TTACAACAATGTTGTTTAGGACAAGTGAGTTAACCTGTTAGCTTCCTCTATATA  
 AAATGGAAAGTCATTAACAACTACATAGCAGGGTTCTTGTGAAGATCAAGTGA  
 TAATGTAGGAAGCATGTACAAATGTCACATTCTGCCGTCACGTAATGGTCCTCA  
 CAGCTTGAGGTAGCATTTAGCATGTGTCATGATTTAGTACAAGGGTTGGCAAAC  
 TGTTGCTCTTGGATTAAGTCTGGCTCATTGCCTGTTTTTCAAAGAAAAAATTG  
 TATATGTGTGTATATATGTTATATATAGGTACACACACATATGTGCTATATATA  
 GCATATACACACATAATATATAAACATGTACATATATAGCATTATATATATA  
 CGTGTATAATATCTCCAGTCTCATGACCAGCCATGCTTGTTCAATTACATTG  
 CATACTCTATGATTGCTTTCATGCAACAATGGCAGAGTTGAGTGATTGTTTTGC  
 AACAGAGACTGTATGGCCCACTAAACCTAAAATATTTAGTCTCTGACCCGAAA  
 TGTAAGATTGATAGCCAGGACCAGGCGTGGTGGCTCACACTTGTAATCCTAGC  
 ACTTTGGCAGGCCAAGGAGGGTGGATCACCTGAGGTCAGGAGTTAAAGACCAGC  
 CTGGCCAACATGGTGAAACCCTGACTCTACTAAAAATACAGAAATTAGCTGGGC  
 GTGGTAATGGGTGCCTGCAATCCAAGCTACTCTGGAGGCTGAGGCAGGAGAATC  
 ACTTGAACCCAGGAGGCAGAAGTTACAGTGAGCTGAGATGGTGCCACTGCACTC  
 CAGCCTGGACGACAGAGTGAGACTCCATCTCAAAA

10

20

III-27

CCATTCTCCTGCCTCAGCCTCTCAAGTAGCTGGGACTACAGGCGCCACAACCA  
 CGCCCGGCTAATGTTTTTGGTATTTTTCGTAGAGACGGGGTTTACCTTGTTAG  
 CCAGGATGGTCTTGATCTCCTGACCTCGTGATCTGCCTGCCTCGGCCTCCCAA  
 GTGTTGGGATTACAGGCACATTTTCACAATTTTAACTTAAGAATGACTT  
 AACTGAATCATGCCTTTAGAAGAACTTTCTGTTAAAAA

III-60

CTGCCGCCGCCAGCTCCCCGCCTCGGGAGGGCACCAGGTCCTGCAGCC  
 AGAGGGGTCCAGAAGAGAGAGGAGGCACTGCCTCCACTACAGCAACTGCACCCA  
 CGATGCAGAGCATCAAGTGCGTGGTGGTGGGTGATGGGGCTGTGGGCAAGACGT  
 GCCTGCTCATCTGCTACACAACCTAACGCTTTCCCAAAGAGTACATCCCAACCG  
 TGTTGACAATTACAGCGCGCAGAGCGCAGTTGACGGGCGCACAGTGAACCTGA  
 ACCTGTGGGACACTGCGGGCCAGGAGGAGTATGACCGCCTCCGTACACTCTCCT  
 ACCCTCAGACCAACGTTTTTCGTCATCTGTTTCTCCATTGCCAGTCOGCCGTCCT  
 ATGAGAACGTGCGGCACAAGTGGCATCCAGAGGTGTGCCACCACTGCCCTGATG  
 TGCCCATCCTGCTGGTGGGCACCAAGAAGGACCTGAGAGCCCAGCCTGACACCC  
 TACGGCGCCTCAAGGAGCAGGGCCAGGCGCCCATCACACCGCAGCAGGGCCAGG  
 CACTGGCCAAGCAGATCCACGCTGTGCGCTACCTCGAATGCTCAGCCCTGCAAC  
 AGGATGGTGTCAAGGAAGTGTTCGCCGAGGCTGTCCGGGCTGTGCTCAACCCCA

30

40

【数4】

CGCCGATCAAGCGTGGGCGGTCTGCATCCTCTTGTGACCCTGGCACTTGGCTT  
 GGAGGCTGCCCCTGCCCTCCCCCACCAGTTGTGCCTTGGTGCCTTGTCCGCCT  
 CAGCTGTGCCTTAAGGACTAATTCTGGCACCCTTCCAGGGGGTTCCCTGAAT  
 GCCTTTTTCTCTGAGTGCCTTTTTCTCCTTAAGGAGGCCTGCAGAGAAAGGGGC  
 TTTGGGCTCTGCCCCCTCTGCTTGGGAACACTGGGTATTCTCATGAGCTCATC  
 CAAGCCAAGGTTGGACCCCTCCCCAAGAGGCCAACCCAGTGCCCCCTCCATTT  
 TCCGTA CTGACCAGTTCATCCAGCTTCCACACAGTTGTTGCTGCCTATTGTGG  
 TGCCGCCTCAGGTTAGGGGCTCTCAGCCATCTCTAACCTCTGCCCTCGCTGCTC  
 TTGGAATTGCGCCCCAAGATGCTCTCTCCCTTCTCCAATGAGGGAGCCACAGA  
 ATCCTGAGAAGGTGAATGTGCCCTAACCTGCTCCTCTGTGCCTAGGCCTTACGC  
 ATTTGCTGACTGACTCAGCCCCATGCTTCTGGGGACCTTTCCTACCCCCATCA  
 GCATCAATAAAACCTCCTGTCTCCAGTGA

10

IV-26

CAGCCCTCCGTCACCTCTTACCGCACCCCTCGGACTGCCCAAGGCCCCCGCCG  
 CCGCTCCAGCGCCGCGCAGCCACCGCCGCGCCGCGCCCTCTCCTTAGTCGCCG  
 CCATGACGACCGCGTCCACCTCGCAGGTGCGCCAGAACTAOCACCAGGACTCAG  
 AGGCCGCCATCAACCGCCAGATCAACCTGGAGCTCTACGCCTCCTACGTTTACC  
 TGTCCATGTCTTACTACTTTGACCGCGATGATGTGGCTTTGAAGAACTTTGCCA  
 AATACTTTCTTACCAATCTCATGAGGAGAGGGAACATGCTGAGAACTGATGA  
 AGCTGCAGAACCAACGAGGTGGCCGAATCTTCTTCCAGGATATCAAGAAACCAG  
 ACTGTGATGACTGGGAGAGCGGGCTGAATGCAATGGAGTGTGCATTACATTTGG  
 AAAAAAATGTGAATCAGTCACTACTGGAAGTGCACAACTGGCCACTGACAAAA  
 ATGACCCCCATTTGTGTGACTTCATTGAGACACATTACCTGAATGAGCAGGTGA  
 AAGCCATCAAAGAAATTGGGTGACCACGTGACCAACTTGCGCAAGATGGGAGGCG  
 CCGAATCTGGCTTGGCGGAATATCTCTTTGACAAGCACACCCTGGGAGACAGTG  
 ATAATGAAAGCTAAGCCTCGGGCTAATTTCCCATAGCCGTGGGGTGACTTCCC  
 TGGTACCAAGGCAGTGCATGCATGTGGGGTTTCTTTACCTTTTCTATAAGT  
 TGTACCAAAACATCCACTTAAGTTCTTTGATTTGTACCATTCTTCAAATAAAG  
 AAATTTGGTACCCAAAAAAA

20

30

IV-41

GCCATTTCTAAGACCTACAGCTACCTGACCCCGACCTCTGGAAGGAGACTGTA  
 TTCACCAAGTCTCCCTATCAGGAGTTCACTGACCACCTCGTCAAGACCCACACC  
 AGAGTCTCOGTGCAGCGGACTCAGGCTCCAGCTGTGGCTACAACATAGGGTTTT  
 TATACAAGAAAAATAAAGTGAATTAAGCGTGAAAA

40

IV-51

## 【数 5】

ATTTCTGTGGATACAGTGCCACCGCCCTCCTCCACTTGGAAACGGTATCCTCC  
 CTGCCCATCCGTCTGTCTGTGCGCCCTTCTCCCGGCCCTCACTAAGCCCCGGCAC  
 TTCTAGTGGTCTCACCTGGAGGCAAGAGGGAGGGGACAGAGGCCCTGCCACGTC  
 CCGCTGCCTCCTGCTCTCTGGAGGTACTGAGACAGGGTGCTGATGGGAAGGAGG  
 GGAGCCTTTGGGGGGCCACCCGGGGCCTGGACCTATGCAGGGAGGCCACGTCCC  
 ACCCCACCTCTTGTCTTCTGGGTCCCTGCTCCCTTTGGGGGTGTGTGTGTGTGT  
 TTTAATTTCTTTATGGAAAAATTGACAAAAAAAATAGAGAGAGAGGTATTTA  
 ACTGCAATAAACTGGCCCCATGTGGCCCCCGCCTTGTCAAAAAAAAAA

10

## V-09

TGGATCCCGTCGTAACCTTAAAGGGAAACTTTCACAATGTCCGGAGCCCTTGAT  
 GTCCTGCAAATGAAGGAGGAGGATGTCCCTAAGTTCCCTGCAGCAGGAACCCAC  
 TTAGGTGGCACC AATCTTGACTTCCAGATGGAACAGTACATCTATAAAAGGAAA  
 AGTGATGGCATCTATATCATAAATCTCAAGAGGACCTGGGAGAAGCTTCTGCTG  
 GCAGCTCGTGCAATTGTTGCCATTGAAAACCCTGCTGATGTCAGTGTATATCC  
 TCCAGGAATACTGGCCAGAGGGCTGTGCTGAAGTTTGCTGCTGCCACTGGAGCC  
 ACTCCAATTGCTGGCCGCTTCACTCCTGGAOCCTTCACTAACCAGATCCAGGCA  
 GCCTTCCGGGAGCCACGGCTTCTTGTGGTTACTGACCCAGGGCTGACCACCAG  
 CCTCTCACGGAGGCATCTTATGTTAACCTACCTACCATTGCGCTGTGTAACACA  
 GATTCTCCTCTGCGCTATGTGGACATTGCCATCCCATGCAACAACAAGGGAGCT  
 CACTCAGTGGGTTTAAATGTGGTGGATGCTGGCTCGGGAAGTTCTGCGCATGCGT  
 GGCACCATTTCCCGTGAACACCCATGGGAGGTCATGCCTGATCTGTACTTCTAC  
 AGAGATCCTGAAGAGATTGAAAAAGAAGAGCAGGCTGCTGCTGAGAAGGCAGTG  
 ACCAAGGAGGAATTTCAAGGTGAATGGACTGCTCCCGCTCCTGAGTTCACTGCT  
 ACTCAGCCTGAGGTTGCAGACTGGTCTGAAGGTGTACAGGTGCCCTCTGTGCT  
 ATTCAGCAATTCCTACTGAAGACTGGAGCGCTCAGCCTGCCACGGAAGACTGG  
 TCTGCAGCTCCCACTGCTCAGGCCACTGAATGGGTAGGAGCAACCACTGACTGG  
 TCTTAAGCTGTTCTTGCATAGGCTCTTAAGCAGCATGGAAAAATGGTTGATGGA  
 AAATAAACATCAGTTTCT

20

30

## V-38

GTTAAATTTGACAAACTAAAGCTAATTAAGGCTATAAGAGTAATAACTGCTCA  
 TTTTCCATAACTCATTCTTAAAGTTTTAGTAATGTAAAAGTTATTTTTTGCAG  
 TAAGTTATAATGATAGAAGCTTACATGTTTTTTTCATGCCTCATCTGTTTCCOCT  
 TAAAACTATAATTATCAGTAAAGTCCCTGTGGTATTTTTCAATTTGTAAGAACT  
 AGGCTATATATACATTGGGAAAAACAGCCTTCATTTGTCAATGCACTAGTGTTT  
 CAAAGGTTTCTGGTAATTGTGTGCTATTGCTTTTTGTTGACTTGCAAAAAAAAAA

40

【数 6】

AAAAAAAAAATTACTATGACTTGTGGTAGCCCTGCAACCTTCGGAAGTGCTTAG  
 CCCAGTCTGACCATACATTTATATTTAGAAATGCTTAGGTAATAAATAATATGC  
 CTAAACCCAATGCTATAAGATACTATATAATATCTCATAATTTTAAAAATCACT  
 GTTTTGTATAATAATAAAACAAGGCAGGCAAGCTGTTCTACAATGACTGTTGGT  
 AAGGGTGCTGAGGAAGAAAAACAACAATCTTGATTGAGGGATAGTGAATAGAC  
 AAAAAATGTCTAATCAATGAAGCTGTGTGATGATTCTGATTGACAGAGAGTGC  
 TGCCACAAGATTCTTAGGCTACACTCAAATCAGCAGAAAAAGTGCTACAATAAA  
 TTAGAAGTGACTATTACAGGTGCAGATGAGGGTTGGTAGTACCTGTTTGCCATT  
 TCTCTTCTAATCTTATATTTTCTGACCCTCCTACTGTAAGTCGCGCGGAGGCGG  
 AGGCTTGGGTGCGTTCAAGATTCAACTTCACCCGTAACCCACCGCCATGGCCGA  
 GGAAGGCATTGCTGCTGGAGGTGTAATGGACGTTAATACTGCTTTACAAGAGGT  
 TCTGAAGACTGCCCTCATCCAGATGGCCTAGCACGTGGAATTCGCGAAGCTGC  
 CAAAGCCTTAGACAAGCGCAAGCCCATCTTTGTGTGCTTGCATCCAAGTGTGA  
 TGAGCCTATGTATGTCAAGTTGGTGGAGGCCCTTTGTGCTGAACACCAAATCAA  
 CCTAATTAAGGTTGATGACAACAAGAACTAGGAGAATGGGTAGGCCTTTGTAA  
 AATTGACAGAGAGGGGAAACCCCGTAAAGTGGTTGGTTGCAGTTGTGTAGTAGT  
 TAAGGACTATGGCAAGGAGTCTCAGGCCAAGGATGTCATTGAAGAGTATTTCAA  
 ATGCAAGAAATGAAGAAATAAATCTTTGGCTCACAAA

10

20

VI-44

GAGAATGGCTTGAACCCAGTAGGCAGAGGTTGTAGTGAGCCGAGATTGGGCCAC  
 TGCACCTTAGCCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAAAAAAAAAAAAAA  
 AATTTAAATAAAATAAAAAACCTTTACTTATTTTTAAATTGGGTTGTCTTTTTG  
 GTATTGAGTTGTTAAAGTTCTTTATATATTTTAGGTACAAATCCCTTATGAGAT  
 ACGTGATTTGAAAATATTTCTCCATTCTGTGGTTGCTTTTTCACTTTCTTG  
 GTTGTATCCTTGAAGCACAGAAGTTTTAAATTTTGATGAAGTCCAGTTTATTT  
 ATTTTTTTGCTGTTGTTCTGCTCATACTTTGAGGTCATGTCTGAGAAACCAT  
 TGTCAAATCCAAGGTCGTGATGACTTACCCCTGTGTTTTCTTCTAAGAGTTTTA  
 AAGGCATCTGAAGCTTAATGTGCACTAGATGGATTCTAAATATCATCTCATCCA  
 AAACCTGCTATATACTACCTTCCTCATCTCAGTTGAAGGCAAGTCCATTGTT  
 TCAATTGCCTGGGCAAAAAATATTCTAAATAATTCATAATTTTCTCAACTCC  
 ACATCTATTGGTAAATCCTGTGGGTTCTCCTTTTAAACATATCCAAAATAGAA  
 TCATTTCTCACTATCATTCCACTGCAGGCACCAAGTCTCAATAGTCTCCTAGCA  
 GATAATCATGTCTACATTTATTCTCAATGTAGCAGCTAGAGAGCTTTTTTG

30

40

VI-49

GCGGTCGTAAGGGCTGAGGATTTTGGTCCGCACGCTCCTGCTCCTGACTCACC

【数7】

GCTGTTTCGCTCTCGCCGAGGAACAAGTCGGTCAGGAAGCCCGCGCGCAACAGCC  
 ATGGCTTTTAAGGATACCGGAAAAACCCCGTGGAGCCGGAGGTGGCAATTCAC  
 CGAATTCGAATCACCTAACCAAGCCGCAACGTAATAATCCTTGAAAAAGGTGTGT  
 GCTGACTTGATAAGAGGCGCAAAAGAAAAGAATCTCAAAGTGAAAGGACCAGTT  
 CGAATGCCTACCAAGACTTTGAGAATCACTACAAGAAAACTCCTTGTTGGTGAA  
 GGTTCTAAGACGTGGGATCGTTTCCAGATGAGAATTCACAAGCGACTCATTGAC  
 TTGCACAGTCCTTCTGAGATTGTTAAGCAGATTACTTCCATCAGTATTGAGCCA  
 GGAGTTGAGGTGGAAGTCACCATTGCAGATGCTTAAGTCAACTATTTTAATAAA  
 TTGATGACCAGTTGTTAAAAAAAAAAAAAAAAA

10

VI-52

GAAAAGGGNTNGCNCCCAANGGGCAGAGGTTGGGCTGATGCCGATATTGGGCCN  
 CTGCNCTNCANACCTGGGTGACATGAATGAAACTCTGTCTCACATAAAAAACCA  
 AAAAAANCTAAATGAAATAAAAGACCTTTGCTTATTNCTAANTTGGGTACGC

VII-15

CCCATCCCCTCGACCGCTCGCGTCGCATTTGGCCGCTCCCTACCGCTCCAAGC  
 CCAGCCCTCAGCCATGGCATGCCCCCTGGATCAGGCCATTGGCCTCCTCGTGGC  
 CATCTTCCACAAGTACTCCGGCAGGGAGGGTGACAAGCACACCCTGAGCAAGAA  
 GGAGCTGAAGGAGCTGATCCAGAAGGAGCTCACCATTGGCTCGAAGCTGCAGGA  
 TGCTGAAATTGCAAGGCTGATGGAAGACTTGGACCGGAACAAGGACCAGGAGGT  
 GAACTTCCAGGAGTATGTCACCTTCTGGGGGCCTTGGCTTTGAT

20

VII-32

AATTAGAGAGGTGAGGATCTGGTATTTCTGGACTAAATCCCCTTGGGGAAGA  
 CGAAGGGATGCTGCAGTTCCAAAAGAGAAGGACTCTCCAGAGTCATCTACCTG  
 AGTCCCAAAGCTCCCTGTCTGAAAGCCACAGACAATATGGTCCCAAATGACTG  
 ACTGCACCTTCTGTGCCCTCAGCCGTTYTTGACATCAAGAATCTTCTGTTCCACA  
 TCCACACAGCCAATACAATTAGTCAAACCACTGTTATTAACAGATGTAGCAACA  
 TGAGAAAACGCTTATGTTACAGGTTACATGAGAGCAATCATGTAAGTCTATATGA  
 CTCAGAAATGTTAAAATAGACTAACCTCTAACAACAAATTTAAAAGTGATTGTT  
 TCAAGGTGATGCAATTATTGATGACCTATTTTATTTTTCTATAATGATCATATA  
 TTACCTTTGTAATAAAACATTATAACCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
 AAAA

30

40

【数 8】  
VII-48

CTTAAGTATGCCCTGACAGGAGATGAAGTAAAGAAGATTTGCATGCAGCGGTTTC  
 ATTAAAATCGATGGCAAGGTCCGAACCTGATATAACCTACCCTGCTGGATTCATG  
 GATGTCATCAGCATTGACAAGACGGGAGAGAATTTCCGTCTGATCTATGACACC  
 AAGGGTCGCTTTGCTGTACATCGTATTACACCTGAGGAGGCCAAGTACAAGTTG  
 TGCAAAGTGAGAAAAGATCTTTGTGGGCACAAAAGGAATCCCTCATCTGGTGACT  
 CATGATGCCCGCACCATCCGCTACCCCGATCCCTCATCAAGGTGAATGATACC  
 ATTCAGATTGATTTAGAGACTGGCAAGATTACTGATTTTCATCAAGTTCGACACT  
 GGTAACCTGTGTATGGTGACTGGAGGTGCTAACCTAGGAAGAATTGGTGTGATC  
 ACCAACAGAGAGAGGCACCCTGGATCTTTGACGTGGTTCACGTGAAAGATGCC  
 AATGGCAACAGCTTTGCCACTCGACTTTCCAACATTTTGTATTGGCAAGGGC  
 AACAAACCATGGATTTCTCTCCCGAGGAAAAGGGTATCCGCCTCACCATTGCT  
 GAAGAGAGAGACAAAAGACTGGCGGCCAAAACAGAGCAGTGGGTGAAATGGGTCC  
 CTGGGTGACATGTCAGATCTTTGTACGTAATTAATAAATATTGTGGCAGGATTAA  
 TAGCC

10

## VII-76

AGACACACGAGCATATTTACCTCCGCTACCATAATCATCGCTATCCCCACGGG  
 CGTCAAAGTATTTAGCTGACTCGCCACACTCCACGGAAGCAATATGAAATGATC  
 TGCTGCAGTGCTCTGAGCCCTAGGATTCATCTTTCTTTTACCCTAGGTGGCCT  
 GACTGGCATTGTATTAGCAAATCATCACTAGACATCGTACTACACGACACGTA  
 CTACGTTGTAGCCCACTTCCACTATGTCCTATCAATAGGAGCTGTATTTGCCAT  
 CATAGGAGGCTTCATTCACTGATTTCCCCTATTCTCAGGCTACACCCTAGACCA  
 AACCTACGCCAAAATCCATTTCACTATCATATTCATCGGCGTAAATCTAACTTT  
 CTTCCACAAACACTTTCTCGGCCTATCCGGAATGCCCGACGTTACTCGGACTA  
 CCCCAGTGCATACACCACATGAAACATCCTATCATCTGTAGGCTCATTATTTC  
 TCTAACAGCAGTAATATTAATAATTTTCATGATTTGAGAAGCCTTCGCTTCGAA  
 GCGAAAAGTCCTAATAGTAGAAGAACCCTCCATAAACCTGGAGTGACTATATGG  
 ATGCCCCCACCCTACCAACATTCGAAGAACCCTGATACAT

20

30

## IX-24

AGAGTGCAAGACGATGACTTGCAAAATGTCGCAGCTGGAACGCAACATAGAGAC  
 CATCATCAACACCTTCCACCAATACTCTGTGAAGCTGGGGCACCCAGACACCCT  
 GAACCAGGGGGAATTCAAAGAGCTGGTGCGAAAAGATCTGCAAAATTTTCTCAA  
 GAAGGAGAATAAGAATGAAAAGGTCATAGAACACATCATGGAGGACCTGGACAC  
 AAATGCAGACAAGCAGCTGAGCTTCGAGGAGTTCATCATGCTGATGGCGAGGCT  
 AACCTGGGCCTCCACGAGAAGATGCACGAGGGTGACGAGGGCCCTGGCCACCA

40

## 【数 9】

CCATAAGCCAGGCCTCGGGGAGGGCACCCCCTAAGACCACAGTGGCCAAGATCA  
 CAGTGGCCACGGCCACGGCCACAGTCATGGTGGCCACGGCCACAGCCACTAATC  
 AGGAGGCCAGGCCACCCTGCCTCTACCCAACCAGGGCCCCGGGGCCTGTTATGT  
 CAAACTGTCTTGGCTGTGGGGCTAGGGGCTGGGGCCAAATAAAGTCTCTTCCTC  
 CAAAAAAAA

## IX-39

CTTGGCTCCTGTGGAGGCCTGCTGGGAACGGGACTTCTAAAAGGAACTATGTCT  
 GGAAGGCTGTGGTCCAAGGCCATTTTTGCTGGCTATAAGCGGGGTCTCCGGAAC  
 CAAAGGGAGCACACAGCTCTTCTAAAATTGAAGGTGTTTACGCCGAGATGAA  
 ACAGAATTCTATTTGGGCAAGAGATGCGCTTATGTATATAAAGCAAAGAACAAC  
 ACAGTCACTCCTGGCGCAAACCAAACAAAACCAGAGTCATCTGGGGAAAAGTA  
 ACTCGGGCCCATGGAAACAGTGGCATGGTTCGTGCCAAATCCGAAGCAATCTT  
 CCTGCTAAGGCCATTGGACACAGAATCCGAGTGATGCTGTACCCCTCAAGGATT  
 TAAACTAACGAAAAATCAATAAATAAATGTGGATTTGTGCTCTTGTA

10

## IX-46

ACGCGAGATGGCAGTGCAAATATCCAAGAAGAGGAAGTTTGTGCTGATGGCAT  
 CTTCAAAGCTGAAGTGAATGAGTTTCTTACTCGGGAGCTGGCTGAAGATGGCTA  
 CTCTGGAGTTGAGGTGCGAGTTACACCAACCAGGACAGAAATCATTATCTTAGC  
 CACCAGAACACAGAATGTTCTTGGTGAGAAGGGCCGGCGGATTCGGGAACTGAC  
 TGCTGTAGTTTCAAGAGGTTTGGCTTTCCAGAGGGCAGTGTAGAGCTTTATGC  
 TGAAAAGGTGGCCACTAGAGGTCTGTGTGCCATTGCCAGGCAGAGTCTCTGCG  
 TTACAAACTCCTAGGAGGGCTGTGTGCGGAGGGCCTGCTATGGTGTGCTGCG  
 GTTCATCATGGAGAGTGGGGCCAAAGGCTGCGAGGTTGTGGTGTCTGGGAACT  
 CCGAGGACAGAGGGCTAAATCCATGAAGTTTGTGGATGGCCTGATGATCCACAG  
 CGGAGACCTGTAACTACTAAGTTGACACTGCTGTGCGCCACGTGTTGCTCAG  
 ACAGGGTGTGCTGGGCATCAAGGTGAAGATCATGCTGCCCTGGGACCCAACTGG  
 TAAGATTGGCCCTAAGAAGCCCCTGCCTGACCAGTGAGCATTGTGGAACCCAA  
 AGATGAGATACTGCCACCACCCCATCTCAGAACAGAAGGGTGGGAAGCCAGA  
 GCCGCCTGCCATGCCCCAGCCAGTCCCCACAGCATAACAGGGTCTCCTTGCCAG  
 CTGTATTCTGGAGTCTGGATGTTGCTCTCTAAAGACCTTTAATAAAATTTGT

20

30

## IX-50

GTCCATCCTGCAGGCCACAAGCTCTGGATGAGGAACTTGAGGCAAGTCACCAGC  
 CCCTGATCATTTGCGCTAAAAGAGCAAGGACTAGAGTTCTGACCTCCAGGCCA  
 GTCCCTGATCCCTGACCTAATGTTATCGCGGAATGATGATATATGTATCTACGG

40

【数 1 0】

GGGCCTGGGGCTGGGCGGGCTCCTGCTTCTGGCAGTGGTCCTTCTGTCCGCCTG  
 CCTGTGTTGGCTGCATCGAAGAGTAAAGAGGCTGGAGAGGAGCTGGGCCCAGGG  
 CTCCTCAGAGCAGGAACTCCACTATGCATCTCTGCAGAGGCTGCCAGTGCCAG  
 CAGTGAGGGACCTGACCTCAGGGGCAGAGACAAGAGAGGCACCAAGGAGGATCC  
 AAGAGCTGACTATGCCTGCATTGCTGAGAACAAACCCACCTGAGCACCCAGAC  
 ACCTTCTCAACCCAGGCGGGTGGACAGGGTCCCCCTGTGGTCCAGCCAGTAAA  
 AACCATGGTCCCCCACTTCTGTGTCTCAGTCCTCTCAGTCCATCTCGAGCCTC  
 CGTTCAAATGATCATCATCAAACTTATGTGGCTTTTTGACCTTTGAATAGGG  
 AATTTTTTAAATTTTTTAAAAATTAATAAAAAAACAACATGGCTCACCCCTC  
 CACCCAAAAA

10

X-77

CCTCCCGGGCTCTTAAGCCCTCTCTTTCTCTAACAGAAAAAGCGGATGGTGGT  
 TCCTGCTGCCCTCAAGGTCGTGCGTCTGAAGCCTACAAGAAAGTTTGCCTATCT  
 GGGGCGCCTGGCTCACGAGGTTGGCTGGAAGTACCAGGCAGTGACAGCCACCT  
 GGAGGAGAAGAGGAAAGAGAAAGCCAAGATCCACTACCGGAAGAAGAAACAGCT  
 CATGAGGCTACGGAAACAGGCOGAGAAGAACGTTGGAGAAGAAAATTGACAAATA  
 CACAGAGGTCCTCAAGACCCACGGACTCCTGGTCTGAGCCCAATAAAGACTGTT  
 AATTCCTCATGCGTTGCCCTGCCCTTCCATTGTTGCCCTGGAATGTACGGGA  
 CCCAGGGGCAGCAGCAGTCCAGGTGCCACAGGCAGCCCTGGGACATAGGAAGCT  
 GGGAGCAAGGAAAGGGTCTTAGTCACTGOCTCCCGAAGTTGCTTGAAAGCACTC  
 GGAGAATTGTGACAGGTGTCATTTATCTATGACCAATAGGAAGAGCAACCAGTTA  
 CTATGAGTGAAAGGGAGCCAGAAGACTGATTGGAGGGCCCTATCTTGTGAGTGG  
 GGCATCTGTTGGACTTCCACCTGGTCAATACTCTGCAGCTGTTAGAATGTGC  
 AAGCACTGGGGACAGCATGAGCTTGCTGTTGTACACAGGGTATT

20

30

XI-13

CTGCCAACATGGTGTTCAGGCGCTTCGTGGAGGTTGGCCGGGTGGCCTATGTCT  
 CCTTTGGACCTCATGCOGAAAATTGGTTCGCGATTGTAGATGTTATTGATCAGA  
 ACAGGGCTTTGGTCGATGGACCTTGCACTCAAGTGAGGAGACAGGCCATGCCTT  
 TCAAGTGCATGCAGCTCACTGATTCATCCTCAAGTTTCCGCACAGTGCCCACC  
 AGAAGTATGTCCGACAAGCCTGGCAGAAGGCAGACATCAATACAAAATGGGCAG  
 CCACACGATGGGCCAAGAAGATTGAAGCCAGAGAAAGGAAAGCCAAGATGACAG  
 ATTTTGATCGTTTTAAAGTTATGAAGGCAAAGAAAATGAGGAACAGAATAATCA  
 AGAATGAAGTTAAGAAGCTTCAAAGGCAGCTCTCCTGAAAGCTTCTCCAAAA  
 AAGCACCTGGTACTAAGGGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG  
 CTGCTGCTGCTGCTGCTAAAGTTCCAGCAAAAAGATCACCGCCGCGAGTAAAA

40

## 【数 1 1】

AGGCTCCAGCCCAGAAAGGTTCTGCCCAGAAAGCCACAGGCCAGAAAGCAGCGC  
 CTGCTCCAAAAGCTCAGAAGGGTCAAAAAGCTCCAGCCCAGAAAGCACCTGCTC  
 CAAAGGCATCTGGCAAGAAAGCATAAGTGGCAATCATAAAAAGTAATAAAGGTT  
 CTTTTGACCTGTTAAAAAA

## XI-49

GATCAACCTGGAGCTCTACGCCTCTACGTTTACCTGTCCATGTCTTACTACTT  
 TGACCCGCGATGATGTGGCTTTGAAGAACTTTGCCAAATACTTTCTTCACCAATC  
 TCATGAGGAGAGGGAACATGCTGAGAACTGATGAAGCTGCAGAACCAACGAGG  
 TGGCCGAATCTTCCTTCAGGATATCAAGAAACCAGACTGTGATGACTGGGAGAG  
 CGGGCTGAATGCAATGGAGTGTGCATTACATTTGGAAAAAATGTGAATCAGTC  
 ACTACTGGAAGCTGCACAACTGGCCACTGACAAAAATGACCCCCATTTGTGTGA  
 CTTCAATTGAGACACATTACCTGAATGAGCAGGTGAAAGCCATCAAAGAATTGGG  
 TGACCACGTGACCAACTTGCGCAAGATGGGAGCGCCGAATCTGGCTTGGCGGA  
 ATATCTCTTTGACAAGCACACCCTGGGAGACAGTGATAATGAAAGCTAAGCCTC  
 GGGCTAATTTCCCATAGCCGTGGGGTGACTIONCCTGGTCACCAAGGCAGTGCA  
 TGCATGTTGGGGTTTCCTTTACCTTTCTATAAGTTGTACCAAAAACATCCACTT  
 AAGTTCTTTGATTTGTACCATTCTTCAAATAAAGAAATTTGGTACCC

10

20

## XI-81

AGAGCAGCAGCCATGGCCCTACGCTACCCTATGGCCGTGGGCCTCAACAAGGGC  
 CACAAAGTGACCAAGAACGTGAGCAAGCCCAGGCACAGCCGACGCCGCGGGCGT  
 CTGACCAAACACACCAAGTTCGTGCGGGACATGATTCGGGAGGTGTGTGGCTTT  
 GCCCGTACGAGCGGCGGCCATGGAGTTACTGAAGGTCTOCAAGGACAAACGG  
 GCCCTCAAATTTATCAAGAAAAGGGTGGGGACGCACATCCGCGCCAAGAGGAAG  
 CGGGAGGAGCTGAGCAACGTACTGGCCGCCATGAGGAAAGCTGCTGCCAAGAAA  
 GACTGAGCCCCTCCCCTGCCCTCTCCCTGAAATAAA

30

## XII-35

CTCTCCTGTCAACAGCGGCCAGCCTCCCAACTACGAGATGCTCAAGGAGGAGCA  
 GGAAGTGGCTATGCTGGGGGCGCCCCACAACCTGCTCCCCGACGTCCACCGT  
 GATCCACATCCGCAGCGAGACCTCOGTGCCCGACCATGTGCTCTGGTCCCTGTT  
 CAACACCCTCTTCATGAACACCTGCTGCCTGGGCTTCATAGCATTGCGCTACTC  
 CGTGAAGTCTAGGGACAGGAAGATGGTTGGCGACGTGACCGGGGCCAGGCCTA  
 TGCTCCACCGCCAAGTGCTGAACATCTGGGCCCTGATTTTGGGCATCTTCAT  
 GACCAATTCTGCTCGTCATCATCCAGTGTGGTTCGTCCAGGCCAGCGATAGAT  
 CAGGAGGCATCATTGAGGCCAGGAGCTCTGCCCGTGACCTGTATCCCACGTACTION

40

【数 1 2】

CTATCTTCCATTCTCGCCCTGCCCCAGAGGCCAGGAGCTCTGOCCTTGACCT  
 GTATTCCACTTACTCCACCTTCCATTCTCGCCCTGTCCCCACAGCCGAGTCCT  
 GCATCAGCCCTTTATCCTCACACGCTTTTCTACAATGGCATTCAATAAAGTGTA  
 TATGTTTCTGGTGCTGCTGTGACTTCAA

XII-77

GTAAGAAAGCCCTTAAATAAAGAAGGTAAGAAACCTAGGACCAAAGCACCCAAG  
 ATTCAGCGTCTTGTTACTCCACGTGTCTGCAGCACAAACGGCGGGCGTATTGCT  
 CTGAAGAAGCAGCGTACCAAGAAAAATAAAGAAGAGGCTGCAGAATATGCTAAA  
 CTTTTGGCCAAGAGAATGAAGGAGGCTAAGGAGAAGCGCCAGGAACAAATTGCG  
 AAGAGACGCAGACTTTCTCTGCGAGCTTCTACTTCTAAGTCTGAATCCAGT  
 CAGAAATAAGATTTTTTGAGTAACAAATAAATAAGATCAGACTCTGAAAAAAA  
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

10

XIII-29

CTCGCTCACGCAGCACTCGTGGCAGTCCCTGAAGGACCGCTACCTCAAGCACCT  
 GCGGGGCCAGGAGCATAAGTACCTGCTGGGGACGCGCCGGTGAGCCCCCTCCTC  
 CCAGAAGCTCAAGCGGAAGGCGGAGGAGGCCCGGAGGCCGCGGATAGCGGGGA  
 ACCACAGAATAAGAGAACTCCAGATTTGCCCTGAAGAAGAGTATGTGAAGGAAGA  
 AATCCAGGAGAATGAAGAAGCAGTCAAAAAGATGCTTGTGGAAGCCACCCGGGA  
 GTTTGAGGAGGTTGTGGTGGATGAGAGCCCTCCTGATTTTGAATAACATATAAC  
 TATGTGTGATGATGATCCACCCACACCTGAGGAAGACTCAGAAACACAGCCTGA  
 TGAGGAGGAAGAAGAAGAAGAAGAAAAAGTTTCTCAACCAGAGGTGGGAGCTGC  
 CATTAAGATCATTTCGGCAGTTAATGGAGAAGTTTAACTTGGATCTATCAACAGT  
 TACACAGGCCCTTCTAAAAAATAGTGGTGGAGCTGGAGGCTACTTCGCTTCTT  
 AGCGTCTGGTCAGAGAGCTGATGGATATCCCATTTGGTCCCGACAAGATGACAT  
 AGATTTGCAAAAAGATGATGAGGATACCAGAGAGGCATTGGTCAAAAATTTGG  
 TGCTCAGAATGTAGCTCGGAGGATTGAATTTGAAAGAAATAATTGGCAAGATA  
 ATGAGAAAAGAAAAAGTCATGGTAGGTGAGGTGGTTAAAAAAATTGTGACCA  
 ATGAACTTTAGAGAGTTCTTGCAATTGGAAGTGGCACTTATTTTCTGACCATCGC  
 TGCTGTTGCTCTGTGAGTCCTAGATT

20

30

XIII-84

ATTATCCTCAGTTCCCAAGAGCAATCATACTTTTCCACACATACCGTGTGTCTC  
 ATGTTAGGTAATGTATTTTTACAATGAGCACCCTTCTGTGGAAAAAGTTCCC  
 TGCACGGGGAGGTCCAGCTTCCAGACTGCTCCATCGCATAAGGACTTCCCCATT  
 CCCCATAATGCTGCTCTGTGAGAACCTGCCAGGTAATGGTAATGACCCTAGAG

40

## 【数 1 3】

AGATGATTTCTGAACCGCAATTTTGAGCCCATTAGAAGGTGTGTGGTGGGCATT  
 TATTTTCATCCTGATGCTCTGGTGAGAACTTTGCAGACGCACTAGATCCAGAAG  
 CTGTTAATCTTGGTGCATTTATTTTCTACCTAAAAGAACCAAGCAGCTCAGAG  
 GCAGTGACTGTACAGGATGCAGTGTATAATAATGCTGAGCTTGCTGGTCTGG  
 AACCCACACTTCAGCAATCCAGCATTGTTTCTGTTTATGAAGTTGACAAAGT  
 GACCAGGGCAAGGGGGTATTATCATTAAATACACTCTAGGAGAGGCAGAACACA  
 TGAGGGCAATGTTTTTCAGAGGTCTTTAGGCCACCGCATCAGATTCTCCTGGAG  
 CATAAAGCAAATGCTTTATGAGTCCAGGGCCCCTGCAGACCTACTGTATACTAG  
 TATACAGCTCCCTCTTAGTGGATCTCAAGCTTGTTCAAAAAGTCATTACACT  
 CCTTACCAAAGCCCATGACACATTCATACAGATTCATCCAGACATAACCCACTG  
 CATGGTCCAGTGCATGCTTGTGTGCTTAACTTATTATAGATCAAGTGTATTTA  
 AGTCCAACATATTAACGTGACTGAATATT

10

## XV-49

AAGTCTGCCAGAAAGCTCAGAAGGCTAAATGAATATTATCCCTAATACCTGCC  
 ACCCCACTCTTAATCAGTGGTGGAAGAACGGTCTCAGAAGTGTGTTTCAATT  
 GGCCATTTAAGTTTAGTAGTAAAAGACTGGTTAATGATAACAATGCATCGTAAA  
 ACCTTCAGAAGGAAAGGAGAATGTTTTGTGGACCACTTTGGTTTTCTTTTTGC  
 GTGTGGCAGTTTAAAGTTATTAGTTTTTAAAATCAGTACTTTTAAATGGAAACA  
 ACTTGACCAAAAATTTGTACAGAATTTTGAGACCCATTA AAAAAGTTAAATGAG

20

## XV-54

AAGAGCAGGTCTCTGGAGGCTGAGTTGCATGGGGCCTAGTAACACCAAGCCAGT  
 GAGCCTCTAATGCTACTGCGCCCTGGGGCTCCAGGGCCTGGGCAACTTAGCT  
 GCAACTGGCAAAGGAGAAGGGTAGTTTGAGGTGTGACACCAGTTTGCTCCAGAA  
 AGTTTAAAGGGTCTGTTTCTCATCTCCATGGACATCTTCAACAGCTTCACCTGA  
 CAACGACTGTTTCTATGAAGAAGCCACTTGTGTTTTAAGCAGAGGCAACCTCTC  
 TCTTCTCCTCTGTTTCGTGAAGGCAGGGGACACAGATGGGAGAGATTGAGCCAA  
 GTCAGCCTTCTGTTGGTTAATATGGTATAATGCATGGCTTTGTGCACAGCCCAG  
 TGTGGGATTACAGCTTTGGGATGACCGCTTACAAAGTTCTGTTTGGTTAGTATT  
 GGCATAGTTTTTCTATATAGCCATAAATGCGTATATATACCCATAGGGCTAGAT  
 CTGTATCTTAGTGTAGCGATGTATACATATACACATCCACCTACATGTTGAAGG  
 GCCTAACCAGCCTTGGGAGTATTGACTGGTCCCTTACCTCTTATGGCTAAGTCT  
 TTGACTGTGTTTCAATTAACCAAGTTGACCCAGTTTGTCTTTTAGGTTAAGTAAGA  
 CTCGAGAGTAAAGGCAAGGAGGGGGCCAGCCTCTGAATGCGGCCACGGATGCC  
 TTGCTGCTGCAACCCTTTCCAGCTGTCCACTGAAACGTGAAGTCCTGTTTTG  
 AATGCCAAACCCACCATTCAGTGGTGTGACTACATAGAATGGGGTTGAGAGAA  
 GATCAGTTTGGGCTTCACAGTGTCAATTTGAAAACGTTTTTTGTTTGTGTTTGTGTA

30

40

【数 1 4】

ATTATTGTGGAAAACCTTCAAGTGAACAGAAGGATGGTGTCTACTGTGGATGA  
 GGGATGAACAAGGGGATGGCTTTGATCCAATGGAGCCTGGGAGGTGTGCCAGGA  
 AAGCTTGTCTGTAGCGGGTTTTGTGAGAGTGAACACTTTCCACTTTTGTACACC  
 TTATCCTGATGTATGGTTCCAGGATTTGGATTTTGGATTTTCCAAATGTAGCTTG  
 AAATTTCAATAAACTTTGCTCTGTTTTTCTAAAAATAAAAAAAAAAAAAAAAAAAAA  
 AAAAAAAA

XV-75

10

AGCAGATGACCOCTTCGTGGCACCCCTCAAGGGCCACAACGGCTGGGTAACCCAGA  
 TCGCTACTACCCCGCAGTTCCCGGACATGATCCTCTCCGCCTCTCGAGATAAGA  
 CCATCATCATGTGAAAACCTGACCAGGGATGAGACCAACTATGGAATCCACAGC  
 GTGCTCTGCGGGGTCCTCCACTTTGTTAGTGATGTGGTTATCTCCTCAGATG  
 GCCAGTTTGCCTCTCAGGCTCCTGGGATGGAACCCTGCGCTCTGGGATCTCA  
 CAACGGGCACCACCACGAGGCGATTTGTGGGCCATACCAAGGATGTGCTGAGTG  
 TGGCCTTCTCCTCTGACAACCGGCAGATTGTCTCTGGATCTCGAGATAAAACCA  
 TCAAGCTATGGAATACCCTGGGTGTGTGCAAATACACTGTCCAGGATGAGAGCC  
 ACTCAGAGTGGGTGTCTTGTGTCCGCTTCTCGCCCAACAGCAGCAACCCTATCA  
 TCGTCTCCTGTGGCTGGGACAAGCTGGTCAAGGTATGGAACCTGGCTAACTGCA  
 AGCTGAAGACCAACCACATTGGCCACACAGGCTATCTGAACACGGTGAAGTGTCT  
 CTCCAGATGGATCCCTCTGTGCTTCTGGAGGCAAGGATGGCCAGGCCATGTTAT  
 GGGATCTCAACGAAGGCAAACACCTTTACACGCTAGATGGTGGGGACATCATCA  
 ACGCCCTGTGCTTCAGCCCTAACCGCTACTGGCTGTGTGCTGCCACAGGCCCA  
 GCATCAAGATCTGGGATTTAGAGGGAAAGATCATTGTAGATGAACTGAAGCAAG  
 AAGTTATCAGTACCAGCAGCAAGGCAGAACCACCCAGTGCACCTCCCTGGCCT  
 GGTCTGCTGATGGCCAGACTCTGTTTGTCTGGCTACACGGACAACCTGGTGCAG  
 TGTGGCAGGTGACCATTGGCACACGCTAGAAGTTTATGGCAGAGCTTTACAAAT  
 AAAAAAAAAACTGGCTTTTCTGACAAAAAAAAAAA

20

30

XV-86

GCAAAATGTCGCAGCTGGAACGCAACATAGAGACCATCATCAACACCTTCCACC  
 AATACTCTGTGAAGCTGGGGCACCCAGACACCCTGAACCAGGGGGAAITCAAAG  
 AGCTGGTGCAGAAAAGATCTGCAAAATTTTCTCAAGAAGGAGAATAAGAATGAAA  
 AGGTCATAGAACACATCATGGAGGACCTGGACACAAATGCAGACAAGCAGCTGA  
 GCTTCGAGGAGTTCATCATGCTGATGGCGAGGCTAACCTGGGCCTCCACGAGA  
 AGATGCACGAGGGTGCAGAGGGCCCTGGCCACCACATAAGCCAGGCCTCGGGG  
 AGGGCACCCCTAAGACCACAGTGGCCAAGATCACAGTGGCCACGGCCACGGCC

40

【数 1 5】

ACAGTCATGGTGGCCACGGCCACAGCCACTAATCAGGAGGCCAGGCCACCCTGCCT  
 CTACCCAACCAGGGCCCCGGGGCCTGTTATGTCAAACCTGTCTTGGCTGTGGG  
 GCTAGGGGCTGGGGCCAAATAAAGTCTCTCCTCCAAAAAAAAAAAAAAAAAAAA  
 AAAAAAAAAAAAAAAAAAAAAA

XVI-74

CGCCGCCGCGCCGCGCTCTCCAACGCCAGCGCCGCTCTCGCTCGCCGAG  
 CTCCAGCCGAAGGAGAAGGGGGTAAGTAAGGAGGTCTCTGTACCATGGCTCGT  
 ACAAAGCAGACTGCCCGCAAATCGACCGGTGGTAAAGCACCCAGGAAGCAACTG  
 GCTACAAAAGCCGCTCGCAAGAGTGCGCCCTCTACTGGAGGGGTGAAGAACT  
 CATCGTTACAGGCCTGGTACTGTGGCGCTCOGTGAAATTAGACGTTATCAGAAG  
 TCCACTGAACTTCTGATTCGCAAACCTCCCTTCCAGCGTCTGGTGGGAGAAAT  
 GCTCAGGACTTTAAAACAGATCTGCGCTTCCAGAGCGCAGCTATCGGTGCTTTG  
 CAGGAGGCAAGTGAGGCCTATCTGGTTGGCCTTTTGAAGACACCAACCTGTGT  
 GCTATCCATGCCAAACGTGTAACAATTATGCCAAAAGACATCCAGCTAGCACGG  
 CGCATACTGGAGAACGTGCTTAAGAATCCACTATGATGGGAAACATTTCAATC  
 TCAAAAAAAAAAAAAAAAAAAATTTCTCTTCTTCTGTTATTGGTAGTTCTGAACG  
 TTAGATATTTTTTTTCCATGGGGTCAAAGGTACCTAAGTATATGATTGCGAGT  
 GGAAAAATAGGGGACAGAAATCAGGTATTGGCAGTTTTTCCATTTTCATTTGTG  
 TGTGAATTTTAAATATAAATGCGGAGACGTAAAGCATTAAATGCAAGTTAAAATG  
 TTTCACTGAACAAGTTTCAGCGGTTCAACTTTATAATAATTATAAATAAACCTG  
 TTAATTTTTCTGGACAATGCCAGCATTTGGATTTTTTAAAACAAGTAAATTT  
 CTTATTGATGGCAACTAAATGGTGTGTTGTAGCATTTTTATCATAACAGTAGATTC  
 CATCCATTAATACTTTTCTAACTGAGTTGTCTACATGCAAGTACATGTTT  
 TTAATGTTGTCTGTCTTCTGTGCTGTTCTGTAAGTTTGCTATTAATAACATT  
 AACTATAAAAAAAAAAAAAAAAAA

10

20

30

XVII-77

CAGACACCCTGAACCAGGGGAATTCAAAGAGCTGGTGCGAAAAGATCTGCAAA  
 ATTTTCTCAAGAAGGAGAATAAGAATGAAAAGGTCATAGAACACATCATGGAGG  
 ACCTGGACACAAATGCAGACAAGCAGCTGAGCTTCGAGGAGTTCATCATGCTGA  
 TGGCGAGGCTAACCTGGGCCTCCCAAGAGAAGATGCACGAGGGTGACGAGGGCC  
 CTGGCCACCACATAAGCCAGGCCTCGGGGAGGGCACCCCTAAGACCACAGTG  
 GCCAAGATCACAGTGGCCACGGCCACGGCCACAGTCATGGTGGCCACGGCCACA  
 GCCACTAATCAGGAGGCCAGGCCACCCTGCCTCTACCCAACCAGGGCCCCGGGG  
 CCTGTTATGTCAAACCTGTCTTGGCTGTGGGGCTAGGGGCTGGGGCCAAATAAAG  
 TCTCTCCTCCAAAAA

40

XII-78

入手できず

【表 2】

表1. 試料の詳細。ステージ0、インサイチュ癌;ステージ I、腫瘍サイズ<20mmの浸潤性癌;ステージ I I、腫瘍サイズ>20~50mmの浸潤性癌;ステージ I I I、腫瘍サイズ>50mmの浸潤性癌;ステージ I V、離れた部分までの癌の広がり、I DC、浸潤性膵管癌;DC I S、インサイチュ膵管癌;ILC、浸潤性小葉癌、n. a.、該当なし;ND、非決定。\* 同じ女性からの連続した5週間で得た血液試料

女性 ID	年齢	ステージ	組織	グレード	サイズ (mm)	ノード	他の疾病(存在する場合/注釈)	時間評価	最終予測
1	51	II	IDC	3	20	1/7	-	2	+
2	84	II	IDC	1	22	2/2	-	2	+
3	50	I	IDC,多病巣性	1	5x14	0	-	1	+
4	66	I	IDC	2	15	0	リウマチ性疾患	3	+
5	66	II	IDC	1	26	0	てんかん	1	+
6	47	I	IDC	2	15	0	-	2	ND
7	69	III	ILC+管状腺癌	2+1	50+3	2/19	-	2	ND
8	50	II	IDC	2	24	0	-	2	+
9	65	I	IDC	1	15	0	-	1	-
10	63	II	IDC	3	23	0	-	1	+
11	65	IV				スプレ及び鎖骨下ノードにおける転移	-	1	-
12	52	I	IDC	1	3	0	-	2	+
13	60	II	IDC	2	23	0	-	2	+
14	54	I	IDC	1	11	0	-	2	+
15	67	0	DCIS	2	20	0	-	3	+
16	na	0	DCIS	2	9	0	-	1	-
17	48	I	IDC	2	4	0	-	2	+
18	na	I	IDC	2	14	0	乾癬	1	+
19	68	I	IDC	1	7	0	-	1	+
20	63	I	IDC	1	10	0	-	2	+
21	65	I	IDC	1	11	0	I I型糖尿病	3	+
22	44	II	IDC	2	25	0	-	1	+
23	55	III	IDC	1	35	0	-	1	+
24	71	I	IDC	1	8	0	-	1	+

10

20

【表 3】  
サブグループ A2:異常第一マンモグラフィの女性

女性ID	年齢	胸部異常	他の疾病(存在する場合)/注釈	時間評価	最終予測
25	44	良性密集状態	-	2	+
26	46	良性密集状態	-	2	+
27	53	良性微小石灰化	左臑における被包性 嚢胞	2	+
28	52	良性微小石灰化	癌、大腸、1992年	1	ND
29	45	良性密集状態	-	2	+
30	59	良性腫瘍、線維腺腫	-	2	+
31	46	良性密集状態	-	2	+
32	46	良性密集状態	潰瘍性大腸炎(1983 年以來)	2	ND
33	50	良性密集状態	1型糖尿病	2	+
34	47	良性微小石灰化	-	2	+
35	46	良性密集状態、嚢胞	クローン病	2	+
36	na	良性密集状態	リウマチ性疾患	1	+
37	44	良性微小石灰化	-	2	+
38	47	良性密集状態	-	2	+
39	50	繊維症、良性	組織サイズ 60mm	1	+
40	45	良性密集状態	I I 型糖尿病	2	+
41	63	良性密集状態、嚢胞	線維筋痛	2	+
42	44	良性密集状態	-	2	+
43	51	放射状瘢痕	組織サイズ 10mm	1	+

10

20

【表 4】  
サブグループ A3: 胸部に異常のない女性

女性ID	年齢	注釈	時間評価	予測
44	22	-	2	+
45	34	妊娠8カ月	3	+
46	27	妊娠6カ月	1	+
47*	18	1週目	2	+
		2週目	1	+
		3週目	1	+
		4週目	2	+
		1週目	1	+
48	29	妊娠9カ月	1	-
49	30	授乳	2	+
50	26	-	1	+
51	43	-	1	+
52	42	-	3	+
53	43	-	2	+
54	34	授乳	3	+
55	-	-	1	+
56	51	慢性 EBV 感染の他 に急性細菌感染	1	+

30

40

【表 5】

表2. ジャックナイフにより選択された35の有意な遺伝子の詳細。アレイにおける位置、クローンID並びにそれに一致する公開データベースにおける配列の受託番号及びそれらの公知又は推定細胞機能を示す。

## アップレギュレーション遺伝子

クローンID	位置ID	受託番号	遺伝子類似性	推定生物学的機能
III-2	6A	ヒットせず	-	-
III-27	10M	AC096970	染色体 3 クローン RP11-321A23,配列番号 135183-135446 から	-
III-60	14AC	NM_001665	Ras 同族体遺伝子ファミリー、メンバーG (rho G)	シグナル伝達、キナーゼインヒビター? (RhoH は、キナーゼインヒビターとして記載されていた)
IV-26	6N	BC016857	フェリチン、重ポリペプチド1	鉄蓄積; ROS に対する防御
IV-51	10Z	BC042655	上流転写因子2、USF2	転写制御因子
VI-44	14X	AC087441	染色体 11、配列番号 116068-116692 から	-
VI-52	14AB	ヒットせず	-	-
VI-15	27E	BC001431	S100 カルシウム結合タンパク質A6(カルサイクリン)	防御; カゼインキナーゼ II の阻害
VII-32	31M	M28697	ヒト低アフィニティ IgGFc 受容体(アルファ-Fc-ガンマーRII)	免疫反応
IX-24	31J	BC047681	S100 カルシウム結合タンパク質A9(カルグラヌリンB)	防御; カゼインキナーゼ II の阻害
IX-50	7Z	NM_007161	白血球特異的転写物1	防御関連
XI-49	3AB	BC016857	フェリチン、重ポリペプチド1、mRNA	鉄蓄積; ROS に対する防御
XII-35	12Q	BC009696	インターフェロン誘発膜貫通タンパク質2	免疫反応
XII-78	24K	-	-	-
XIII-84	16AP	AL391903	配列番号 75875-76710 から	-
XV-54	24AA	BC018148	デルタ睡眠誘発ペプチド、イムノリアクター	免疫反応?
XV-86	24AQ	BC047681	S100 カルシウム結合タンパク質A9(カルグラヌリンB)	防御; カゼインキナーゼ II の阻害
XVI-74	5AK	BC066901	H3 ヒストン、ファミリー-3B (H3.3B)	クロマチンリモデリング
XVII-77	20AN	BC047681	S100 カルシウム結合タンパク質A9(カルグラヌリンB)	防御; カゼインキナーゼ II の阻害

10

20

30

【表6】  
表2(続き)

## ダウンレギュレーション遺伝子

クローンID	位置ID	受託番号	遺伝子類似性	推定生物学的機能
I-30	21O	BC009689	サイクリンD型結合タンパク質	E2F 介在転写
IV-41	2V	BC010165	リボソームタンパク質 S2	リボソーム産生
V-09	2G	BC053370	リボソームタンパク質 SA	リボソーム産生
V-38	22S	NM_001016	リボソームタンパク質 S12	リボソーム産生
VI-49	2AB	NM_001023	リボソームタンパク質 S20 (RPS 20)	
VII-48	31U	M22146	リボソームタンパク質 S4	リボソーム産生
VII-76	15AK	AY495316	チトクロームcオキシダーゼサブユニット、COX 1	ミトコンドリア電子伝達鎖
IX-39	27R	BC001037	リボソームタンパク質 L35a	リボソーム産生
IX-46	23V	BC034149	リボソームタンパク質 S3	リボソーム産生
X-77	19AM	BC000514	リボソームタンパク質 L13a	リボソーム産生
XI-13	19H	D87735	リボソームタンパク質 L14	リボソーム産生
XI-81	3AR	AF077043	60S リボソームタンパク質 L36	リボソーム産生
XII-77	20AK	BC035447	リボソームタンパク質 S6	リボソーム産生
XIII-29	20N	BC004465	テロメア反復結合因子2、相互作用タンパク質	テロメア長調節
XV-49	4AA	BC018641	真核翻訳伸長因子1アルファ1、(EEF1A)	タンパク質翻訳
XV-75	12AM	BC019093	グアニンヌクレオチド結合タンパク質、ベータポリペプチド2様、RACK(活性化Cキナーゼ用受容体用)	タンパク質翻訳

10

20

【表 7】

Table 3: 乳癌に有用なプローブファミリー (i) および (ii) 遺伝子

Probe No.	Agilent ID	Oligonucleotide sequence
2	A_23_P164011	ACTCCAGACTGGGAAGACCTTCCATTTCCAGGATCGACGCTTACGTTGAGGGGAGGGC
3	A_23_P94111	TTACCAAACTCAAAGCTTATTTAGTAGAATGGGCTCATGGCAATGTGATGTTCCCTGT
5	A_23_P155009	TGTTGGTTGGAGGACAAGTGGCACTGAGACCCCTGGTGAACCCATGGAAGGGTGGCCTG
6	A_23_P84323	TGGAGAAAGGACCCCTGGACCTGTGGGTCATCGTCCGTTCCAGGAGCAGCAGGCTGGGG
8	A_23_P121716	TGGACATTCGAACAGAGTCAAGAAGCATTATGGCTATTCCTATATTCAGCAATTAAT
10	A_23_P111037	ATCAGAAGTCCACTGAACIGCTTATCGTAACTACCTTCCAGCGCCTGGTGGCGGAGA
13	A_23_P75830	TTTGTGAAACTGTGTATACTTTGTGGTATAGACTGCCTGTTAGTATGAAGGGCGG
16	A_23_P149836	CCTCCAGCAGTTAAGTAACTTGTGAAGATGGGACCCCTGTTCCCTAATGGTTCTAGAA
17	A_23_P134805	CTGAATCTGTTTGTCTTCCCTAATCTATCACAATGGCCACCCATCGGGTTTGGGTGTGT
18	A_23_P154235	CCATGTTTCTGAATCTTCTTTGTTCAATGGTGTGCATGTTTCAACTACAATAAGTG
19	A_23_P2616	ATCATTCAAGATCTGAAAGAAATCTTCTTATTTTCTGGGCTGTGAGATCCAGGGGGT
20	A_23_P333484	CCACCGAGCTGCTGATCAGAAAGCTGCCCTTTCAGCGTCTGGTGGTGAAGATCGGCAGG
24	A_23_P259874	TCTCAGAAGAATGTTGGCCATGAGACTATCATTGAGAGGAGGGGATTTCTCTTCA
26	A_23_P206568	AATCCTGTGATTTCTGTGTGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT
27	A_23_P115091	GATGGCTGAAGGAGCTATGACCATGCTGAAGCCACGATCGTCAATGATGATGATGATGATG
28	A_23_P46718	TGCATGGGGAGTACATTCATCTGGAGGCTGGCTGCTGATGAATGCTGCTGCTGGGGT
29	A_23_P218456	GTTTTGAGTTTTTGCAGTTTCAAGTATCCCTCTGCTATTCACACTTCGTTAGTGGTAA
31	A_23_P76610	CAGTTTATGGATGCTGGGCAATCATAGCACTTGCATTAAACAAACATGCTACAGGGGCA
32	A_23_P206396	ATTATCAACTCACTGGTAACAACAGTATTCATGCTCATCTGTGTGGCACTGATA
34	A_23_P56091	GAAACGGATCGCAAGCTTCCAGGATCCCTTTCGTTGCTGGGGTGGGAAGCATGG
35	A_23_P55184	TCAATTCAGGCTCCCTGCCTCTACTAGGCGCTTAGCTCACTATGGGAAACCCTTG
37	A_23_P150974	CAAAATAGCTACATCCCTGAACACAGTCCGGAATATACGGCCGACCCAGGGAATCGGGA
40	A_23_P111689	TTAATTCATTGGCTCTTAGTCACTTGGAACTGATTAATTTGACTTTCTGTCACTAAGC
41	A_23_P58937	GTCTCAAAACAGCCGAAACCTGCTTGCATGGGGGAGGGGGTTCGCTTTCCTTCTT
42	A_23_P74828	TTGGCTTTAGACATTAATATATATATATATATATATATATATATATATATATATATATAT
45	A_23_P42168	GGAACTGTGAAAGTTACTTGGGAGGGTGGGCGGTTGGGCGGTTAGCTCTCTACCTCT
47	A_23_P81278	TCAGACAGGCTTGGTAAGTGAACCTTGAACCTTGAACCTTGAACCTTGAACCTTGAACCTT

10

20

30

40

【表 8】

49	A_23_P251695	AGGTTGAACTCTTTTGTGCTCAAGTTCTAGGAGTCCCTTCTGAAATATATACTTGT
52	A_23_P393645	CTACTTAGAGTCTTCCAAATGCCAAAAGGCTAGGGGTTGGAGGTGGGACTCTGGA
53	A_23_P208683	ATAGTCATGGGTGTCTAGAAAATACCAAAATGTAAGAGAACTCCAAAGTCAGGGCCAG
55	A_23_P72016	GACATTGAGAAAGAAAACCGGGAGTGGAGACTGGCCCAAGAAATCGATGCACCTAAGT
58	A_23_P16915	CATATCCATTTTAAAGAGAGGTGCCAGTTCTGCATCTGATACCGTCTCCTTTCCCT
60	A_23_P37076	GAAAATCCCCTTCTATGCTTTCCCTACTAGAAATGTTCTAGAAATCGCTGGACGGTGGGT
61	A_23_P166408	TGGTGGTGGATCCCTGGAATTTCTACGCCAGGACCCATGCTCTCCTAGAGGGGCTCA
63	A_23_P94501	GGCTCTTTCTGGAGGAACTAAACATCCCTTGATGGTCTCAAGCTATGATCAGAAGACT
67	A_23_P94230	TGAAGCTATTTCTGGAGGCCAGAAATGCTCTTTGCTTGGAGTTTGTCCATCCCTACA
68	A_23_P154037	TTGGTTCCCTTAGGGTGATATCGTCAATTAATGCTCTGCTCTTCAATCCATCCAGCTAAAT
69	A_23_P47938	AAGAAAACACACCTCGGGCAGCAATGCTTTGCTGCTCGGATTAGGTGGGGATGGGCGACA
70	A_23_P90743	CTGAGTTTGCCTGTTAATCTTCAATAGTTTACCTACCCAGTCTTTGGAAACCCTAAAT
71	A_23_P171249	AAAGTGTCAAGTGATTAAGTGTGATTTGTACCCCTAGATGATGAACCAGCAGCTTTGT
72	A_23_P142675	TGAGCTGTCCCTTCTTAAGCCATAATCTCTTAGTGGATTGAGCCCTTTGGAAAGACT
73	A_23_P153637	TGTTATTGGCCTAGAGCTACAGTATATGGGTTTGTCTGAGTCCGTTTCAAATGACCT
76	A_23_P76749	CCTGTTCTGTTTTGCTTTTCCCTCTTGGACCAAGCAATGGCCACTAGCTGTCCCTTGA
80	A_23_P169061	TTGAAGGCAAAGATCATCAATCTGCACTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT
91	A_23_P206253	CGAATTGGGAGGCTTATTTTTCAGCAAGAAATTTGGGGGGTTTTGTTGTTGGGGC
95	A_23_P151995	AATAACAACCTTGTATGATGTAACCTTGAACCTCCAGAGTTATGGAAATTTGTCCCCATG
97	A_23_P138011	AGAGACCTGCAGGGGCCCTCGGCCCTCACATCGTGTATGCTCTCCTTGGATTTGTGTTGT
100	A_23_P35912	GCCAAAAGCTCAAAATGCCACCATAGAAGGACTGCCATGACAAGATATTTTACCTCTTT
101	A_23_P99424	TGGCTCCCATCATGTATCCTCCCGATTTGCGTATTTCTAAAATAGGAAACAAGACTTT
105	A_23_P418986	GATGACACTGCCACCCTCTGACTTCTGCCTCTGGCCCTCCACTCTCAGTAAGAAGAGCCAG

10

20

30

40

【表 9】

Table 4: 乳糖に有用なプローブファミリー (i) および (ii) 遺伝子

Probe No.	Agilent ID	
1	A_23_P368812	TTACTTCTACCTGCTCTCCCAACTCCCTGAGCCTGAGTGAGCGTGTGGCCATCATCA
4	A_23_P388391	TGGCCCTCAAAATGAGATGGATCCAGGTCTTGGGACCCCTGGGATGTTGGGACTT
7	A_23_P4096	TAATATCCCAAACTGAGATGAGCACTACGATGGCAGAGAGCAGCCTGTTGGACCTGCT
9	A_23_P15450	GACTGAAAATCAGCTTCTAATTAATGAAACACTTGGGGTCAATGGGAGTGCACAGC
11	A_23_P379596	AGGATAAATCAAACTGACAACCTGTCAGTCCCGTGGAGGTAGGGAGTGGGTGAT
12	A_23_P391275	TAAATATGATTTACTCTGTGCTGTTCCAAATGGGACCCAGGAGAAATATGAACCTC
14	A_23_P124661	TCATATTTAATAACTTCAGACTGGGCCCTGTTCTTCCCATTAACCTGAGTG
15	A_23_P44257	AACATTTACTTCTGCGCTTCTATGTTGGAAACATGCTCTGATAAAAAATAGCTGTC
21	A_23_P128183	CTGAGAGTTTTGCAGAAATGGGCAGAGGGACACCCCTTGGCGTGGCTTCCCTGGTGAT
22	A_23_P331211	CGAGTGGCTCACTCAGAAATCTCATTTGATGGCTAGGACCCCTACTCGTGGGTCATGC
23	A_23_P94932	TCTGTTGATGACCTTGGATGCTGTAAGTATCGTCATAGTCTCTGGGGTACCCATGTA
25	A_23_P102122	AAGCGGCTGGCAACTGAAGCTGGAACACTTGCTACTGGATAATCGTAGCTTTTAATGTT
30	A_23_P407654	GAGGAGCTCTTTCTAGAGAGCCGGAGTGGGGAGGGGATTTATTTGTTATTTAAT
33	A_23_P392457	CCTCTGACTGCCCTCAACGTAAAAATGTAATATAAATTTGGTTGAGATCTGGAGGGGGG
36	A_23_P406376	GCCACACTGGCTTAGGACCTGTGACACGGAGGGGGTTTTAATTTGGTTTTTAACAA
38	A_23_P22723	AACAACTACAGTTTTACCGTGTTTGCCATTTGAGCTGTGTGGTGGGCAGGGGGCTGG
39	A_23_P70258	AGAGAGGATGGCTGTATCCCTATCCAGCTCAAGCTGCCAGCAGCAATGTTGGCTGCCCA
43	A_23_P104005	AATTTTCAAGACTCTTTTCACTCTTTGATTTGGATCTGGCAAATGGGAGGGGATGCT
44	A_23_P119652	TTGCCAACTGACCGTGGCTGAACACACGTTCTGCTTGACTCATTTAGGGGGAGGGAA
46	A_23_P22957	ATGAGGTGATCACTGTGTTGAGTGTGTTGGAATGGATTCAGACTGGCTAAATGGGGGAAA
48	A_23_P8072	GGGGAGATCAGAATCGTCCAGCTGGCTCGACTTGGATGGATTCAGACTGGCTAAATGGGGGAAA
50	A_23_P23346	AATCTTCTGAACGGCATAAGTCTTATTTAGCCTTACCTCCTGCAATGGCAATACGTAAT
51	A_23_P92342	CGAACAAACAATAACTTGGCGGGCCCGAGAGGGCTCGTTGGCCATTCGTTGGGGAT
54	A_23_P153183	ACAGAAAAACAGACTTGTAAAAAGCTTAGATCATCAAGTGTTTGGATGGGGCCCTCCCA
56	A_23_P157231	TGCAGAAATGCATAAGATGAACATTCATGACCCGGATCATTTAGTGTCTTTGCCGTTAAA
57	A_23_P103282	TGAAGATCATGAAGAAGCAGGGCTCTACCTACAAAAGTGAATCTTGCTCATCTGAAAT
59	A_23_P109462	CTGGATGTTTACCTGGAGACCGGAGCCCATGACGACAGTGTGGAGGGCCCAAGGAATTT

10

20

30

40



【表 1 1】

109	A_23_P112251	AGAAATCTTAAC TTCACAAGTGTTTACTTCGACGATGTGCCTTTGATTTAATTGGGAC
110	A_23_P313330	TCATTAGACATCGGGGATTCACCTGCAGAGTAATCCTGGAACCTACATTAAAGTGGGGG
111	A_23_P27414	TGCGGGAAGCCCTTCAGCCACCGTTGCAACCTCAACGAGCACCAGAAAGCCGCGGGGGC
112	A_23_P210981	TTGTAGGACCTTAATGGCTAAGAATTAGAACATAGCAAGGGGGCTCCTCTGTTGGAGTAAT

10

20

30

40

【表 1 2】

表 5 乳癌に有用な遺伝子-ファミリー①および②の遺伝子

グループ番号	受託番号1	受託番号2	遺伝子ファミリー①	遺伝子ファミリー②
			転写因子 (ファミリー①)	遺伝子類似性及び推定生物学機能
2	NM_006942	AB006867	SRY (性決定領域) ーボックス20, 転写因子のファミリーを含む SRY 関連 HMG ボックスのメンバー	
3	NM_002095	X63469	一般転写因子 IE 2 (34 HDa サブユニット), RNA ポリメラーゼ I I 転写因子 TFE のペータサブユニット, 転写開始に必要, 活性化因子及び DNA 修復タンパク質と相互作用し, 転写運動修復の役割を果たすことができる	
6	NM_016942	M95987	H6 ホモオボックス1, DNA 結合タンパク質のホモドメイン含有ファミリーのメンバー, マウス Nr2f5 介在転写活性を担うことができる転写レプレッサー	
26		BC026031.1	T-box 6, 転写因子の T-box DNA 結合ドメインファミリーのメンバー, 胚の軸傍中胚葉形成及び体節形成に関与することがある	
45	NM_005596	U78313	MyoD ファミリーインヒビター, 筋形成を抑制調節する推定転写レプレッサー	
69	NM_014212	AJ000041	ホメオボックス C11, ホメオドメイン含有転写因子, HNF 1 アルファ (HNF1) 依存転写を活性化することができ, 膵の早期の発達及び分化において機能することができ, NUP98-HOX11 融合タンパク質は, 骨髄性悪性腫瘍に関与	
105	NM_004348		ホモサビエンズラント (hnt) 阻害転写因子 2 (RUNX2), mRNA	
			防癌転写因子 (ファミリー②)	
63	NM_000700	BC035993	アネキシン1, ホスホリパーゼA2を阻害し, 抗炎症活性を有するカルシウム依存性リン脂質結合タンパク質, ストレス応答に関与, 食道癌及び前立腺癌における腫瘍の早期診断に関連	
8	NM_005139	M63310	アネキシンA3 (リポコルチン I I I), カルシウム依存性リン脂質結合タンパク質のアネキシニンファミリーのメンバー, コリンを結合, 膜融合及び透過性及びフラゴサイトーシスの調節に役立つ	
13		BC007022.1	血清アミロイドA1, 白血球遊走能に作用し, マトリックス金属プロテアーゼを誘発する急性期アポリポプロテイン, リウマチ様関節炎, アテローム性動脈硬化症, 反応性全身性AAアミロイドーシス, アルツハイマー病及び多発性硬化症に関与することがある	
18	NM_004688	BC001268	Nmyc (及び STAT) 相互作用因子, N-myс (MYCN) 及び STAT タンパク質と相互作用するタンパク質, CBP/p300 と STAT タンパク質との会合を促進することにより IL2 及び IFN ガンマ応答性転写を増加, BCR A1 が乳癌の発生を抑制しやすくする	

10

20

30

40

【 表 1 3 】

19	NM_208503	AF325460	C型(カルシウム依存性水化物認識)メインレクチンチンズスバーファミリーメンバー11、インターフェロオンアルファ及びベータの誘発を阻害し、抗原補綴に介在してT細胞依存性免疫反応を開始できるテンドライト細胞糖タンパク質
27	NM_020387	AF274025	アゴサイトーシス及び自己免疫に関与する推定GTPaseである、Ras21 様小型 GTP 結合タンパク質11a(ヒトRAB11A)と類似性の高いタンパク質、GTP 結合タンパク質の Ras スーパーファミリーのメンバー
29	NM_012218	AJ271747	インターロイキンエンハンサー結合因子3、NF-AT のサブユニット、正又負の転写制御因子として作用、T細胞の IL2 発現に必要、mRNA 処理、翻訳の阻害、宿主防御及び自己免疫に関与することがある
32	NM_181640	BC004380	ケモカイン様因子1、白血球、好中球、単球及びリンパ球のための分泌化誘引物質、炎症反応及び血管細胞増殖及び増産を刺激、筋形成の調節に関与
37	NM_153633	X07495	ホメオボックスC4、DNA結合タンパク質のホメオボックスファミリーのメンバー、リンパ球活性化の調節及び造血中の細胞系決定に関与することがある
61	NM_020530	BC011589	オンコスチチンM、インターロイキン-6 サイトカインファミリーのメンバー、活性化単球及びTリンパ球により産生、JAK-STAT 及び MAPK 経路の活性化により細胞増殖及び分化を調節、カポジ肉腫細胞増殖の調節
67	NM_015364	AB018549	MD-2 タンパク質、リポ多糖受容体複合体の一部、リポ多糖及びトール様受容体4 (TL4) 情報伝達に寄与、細胞防御反応に関与することがある
80	NM_000912	L37382	カンパセチド受容体1、阻害的Gタンパク質を介して情報伝達するGタンパク質結合受容体、痛み等の知覚を変化させることがある;発現変化は、アルツハイマー病に関連している;アゴニスト刺激は、HSV 感染を抑制することがある
95	NM_004049	U29680	BCL2 関連タンパク質 A1、アポトーシス制御因子の Bcl-2 ファミリーのメンバー、アポトーシスを阻害、腫瘍発生を促進、そして炎症中の防御役割に関与することがある
17	NM_003580	BC041124	中性スフィンゴミエリナーゼ(N-Smase)活性関連因子、N-Smase の腫瘍転移因子受容体 CD40(TNFRSF5)誘発を介在、アポトーシスのTNFアルファ介在誘導に関与、TNF受容体 TNF-R55(TNFRSF1A)に結合
34	NM_144615	BC015655	タンパク質-タンパク質及びタンパク質-リガンド相互作用に関与することがある、免疫グロブリン4(IgD)ドメインを含むタンパク質
10	NM_003529	BC087491	クロマチンリモデリングファミリー(n) H3ヒストンファミリーメンバー-A、コアヒストン H2A、H2B、H4 および DNA とともにヌクレオソームの成分

10

20

30

40

【表 1 4】

20	NM_003536	BC082305	ホモサピエンシステトン1、H94HSTH94、miRNA
31	NM_018282	AK090873	バラスベックルタンパク質1、2つのRNA結合(RRM)ドメインを含む、推定RNA結合タンパク質、バラスベックルインタークロマチンスペースコンパートメントと核小体との間を移動、転写保存方式で核小体と相互作用
16		AK024156	リボソーム生成(ファミリー①)
			未知の機能のタンパク質、ScerevisiaeBms1pの領域との類似性は中程度、RNA処理及び40Sリボソームサブユニット生成に関与タンパク質代謝(ファミリー①)
52	NM_139028	AY358118	ホモサピエンシスジスインテグリン様及びメタロプロテアーゼ(レプロリジン型) (トロンボスポンジンタイプ1)モチーフを有する、13(ADAMTS13)、転写変異体1、miRNA
72	NM_012100	AK001777	サイトソルアスミルチルアミニペプチダーゼ、メタロプロテアーゼのM18ファミリーのメンバー、N末端アスミルチル及びグルタミル残基に対する基質の優先性がある、細胞内ペプチド代謝に関与することがある
100	NM_001225	U28979	カスパーゼ4、アポトーシスの誘発に関与するICEシステインプロテアーゼファミリーのメンバー、半胱氨酸セルピニン0mMによる阻害により、アポトーシスの阻害による感染が容易になることがある
101	NM_003291	AK097678	トリペプチジルペプチダーゼI1、非プロテアソームタンパク質タンパク質タンパク質に作用することがあるセルノエクソペプチダーゼ、神経ペプチド及びMHCクラスI抗原が基質、シグナル伝達にアポトーシスにおいて作用、MYCを過剰発現するバネキリン腫瘍抑制剤においてアップレギュレーション
5	NM_012265	BC002705	内在性膜タンパク質の偏離形ファミリーのメンバー、UBA(ユビキチン関連)又はTS-NITドメインを含む
55		M64247.1	トロポニン阻害サブユニットである心臓トロポニンI (マウス Troi)と非特異的に高い類似性を有するタンパク質、カルシウム誘発筋収縮を調節するトロポニンファミリーのメンバー
76		BC038812	UDP-N-アセチルアルブアーD-ガラクトサミン:トリペプチドN-アセチルガラクトサミントランスフェラーゼ(ヒト GALANT2)に対する類似性が高いタンパク質、グリコシルトランスフェラーゼファミリー2のメンバー、20XWUリンBLクエン酸阻害剤ドメインを含む
			酸化ストレス(ファミリー①)
58	NM_012413	X71125	グルタミルペプチドシクロトランスフェラーゼ(グルタミルシクラゼ)、下垂体において発現、酸化ストレスの阻害タンパク質レギュレーション
68	NM_001159	L11005	アルテロトオキシダーゼ、酸素ラジカル、外因性物質及び薬物代謝に関与するモノテレン含有アラチン素、常染色体性染色体型型の筋萎縮性側索硬化症の原因と考えられる遺伝子

10

20

30

【表 15】

60	NM004873	AK023145	BCL2 関連アミノ酸残基6、BAGドメイン含有、BAGドメインを介してATPaseドメインに結合してHsc70/Hsp70タンパク質を調節すると予測される。
24	NM012430	AF100749	タンパク質分泌(タンパク質合成-ファミリー④) Sec22 同族体、小胞輸送タンパク質の SEC22 ファミリーのメンバー、小胞体からゴルジ体へのタンパク質輸送に関与することがある
28		M65189	エンドセリン 2、血管作動性ペプチドホルモンのファミリーのメンバー、血圧調節に関与、プロラクチンの分泌を阻害、心臓発生に關係する細胞増殖に作用することがある。遺 伝子座標は、高血圧症に関連がある
35	NM001661	L38490	ADPリボシル化因子4様、GTPase 及び ADPリボシル化因子ファミリーのメンバー、小胞の細胞内輸送及びタンパク質分泌に関与することがある
41		BC028121	分泌タンパク質の転位を刺激する推定小胞体受容体である転位調節タンパク質(ト-TRAM)との類似性が高いタンパク質、長寿調節タンパク質(LAG1)ファミリーのメン バー
42	NM030772	AF271281	分子が細胞質を通ずるの可能性があるキヤップ結合タンパク質のコンネキシンファミリーのメンバー、臓器心房逆位置と関連するキヤップ結合タンパク質/アルファ 1(コネキシン43、ヒト-GJA1)との類似性が中等度
49	NM013248	AK026360	NIF2 様輸送タンパク質1、PAN と結合、ORMIXPO1 依存核輸送経路において機能
53	NM012946		核孔アセタンパク質 p62、核孔の成分、核細胞質輸送に関与することがある、ポリオウイルス感染中の宿主の標的
73	NM032139	AL138784	空胞運送タンパク質(VPS9)ドメイン及び8つのアンキリン/Ank)反復を含むタンパク質、軸索芽イダンス及び適切な軸索形式に必要とされるC、エレガンスUNC-44の領域C の領域との類似性が低い領域を有する
91		AB010418.1	コア結合因子タンパク質ドメインアルファサブユニット-転位3、MTG8(ETO/ODR)タンパク質ファミリーのメンバー、推定転写因子、RUNX1 への対応遺伝子の融合が、急性骨髄性白血 血病にみられる
97		AL137537	家族性肝細胞肝癌症と関連するアミリン-脂質ATPase 輸送体(家族性肝細胞肝癌症1、ヒトATP8B1)との類似性が高いタンパク質、ハロ酸デハロゲナーゼ又はエポキ シドヒドロラーゼファミリーのメンバー

10

20

30

【表 16】

70	NM1002909	M27190	日細胞形態形成免疫応答、ファミリー(ii) 島由来1アルファ(糖タンパク質)を再生、降ベータ細胞再生を誘発、動物において糖尿病を改善、異常発現は慢性石灰化肺炎及び結腸癌発生と関連している
71	NM1001551	BC004137	免疫グロブリン結合タンパク質1、6G 疫病原体在B細胞情報伝達に関与することがある
40		AC007032	プレB細胞コロニー増強因子、早期B系細胞において、幹細胞因子(KL6)とインターロイキン7(L7)のコロニー形成能を相乗的に向上させるサイトカイン、感染誘発 早産及び一対腸癌に関与することがある
47	NM152547	AK057097	免疫応答ファミリー(ii) 増加及びインターフェロングamma合成を正に調節するT細胞の同時刺激分子であり、炎症サイトカインにより誘発されるB7同族体3(ヒトB7-H3)との類似性が低いタンパク質

受託番号1及び2は、遺伝子の代替受託番号である。関連配列は、NCBIデータベース([www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov))で確認できる。

10

20

30

40

【表 17】

表6. 乳癌に有用な遺伝子-非ファミリー0および1の遺伝子

グループ番号	登録番号1	登録番号2	遺伝子類似性及び推定生物学的功能
			チャンネル及びポンプ
1	NM_001661	BC034366	ホモサビエンズ アクアポリン5(AQP5), mRNA
4	NM_005072	AF054506	ホモサビエンズ 溶質キャリアファミリー12(カリウム/塩化物輸送体), メンバー4(SLC12A4), mRNA
11	NM_004983	U52152	ホモサビエンズ カリウム内流流チャンネル, サブファミリーJ, メンバー9(KCNJ9), mRNA
38	BC047580		形質膜 Ca <sup>2+</sup> -輸送 ATPase3, カルシウム輸送に関与すると予測される, 主に脳で発現
107	NM_174873	AF260427	プリン性受容体 P2X2, カルシウムイオン輸送及び情報伝達に関与している, 細胞外 ATP により開閉されるカチオンチャンネル
65	NM_130840	AK055789	ATPase (H <sup>+</sup> 輸送) シノーム V0 サブユニット A1 型4, 液胞プロトンポンプの非触媒ドメイン
44		AC004659	興奮性アミノ酸輸送体4(溶質キャリアファミリー6), リガンド開閉塩化物チャンネル活性を有する, 高アフィニティグルタメート及びアスパラチレート輸送体, 小脳内で興奮性神経伝達を調節すると思われる
74	NM_017836	AK000480	Mag <sup>2</sup> 又は他の二価カチオンを細胞内に輸送できる二価カチオン輸送体ファミリーのメンバー, 糸状菌中のヒト DKFZP434K0427 との類似性が高い
48	NM_032464	L26280	推定キナーゼ又はキナーゼ相互作用タンパク質
54		AK056549	セリン/スレオニンキナーゼ 19, ほとんど核内局在化するマンガン依存タンパク質キナーゼ
			腸動運タンパク質キナーゼ1, SAM, PDZ および PHドメインを含み, シナプス足場キナーゼ S-SCAM および PSD-95/SAP90 と相互作用する, ラット Maguin1 と高い類似性のあるタンパク質
66	NM_003137	BC038292	RNA スプライシング因子のセリン及びアルギニンリッチ(SCR)ファミリーのためのタンパク質キナーゼ, 核内のスプライシング因子の局在化を制御するように作用すると思われる; 広く使用されている抗腫瘍剤であるシスプラチンに対する感受性を決定に関与していると思われる
78	NM_015518	BC056423	タンパク質キナーゼドメインを含むタンパク質, 小脳神経節細胞軸突突起伸張の初期段階に関与しているタンパク質キナーゼであり, 情報伝達カスケードにおいて作用することができる, unc-51 様キナーゼ1(マウス Ukn1)の領域との類似性が中程度である

10

20

30

40

【 表 1 8 】

81	NM_007061	BC009356	骨髄腫内皮血清構成タンパク質、非キナーゼ CRIB (Coc-42/Rac 相互作用結合ドメインを含む、GTP 依存的に CD42 と結合する、細胞骨格再構成及び Rac タンパク質情報伝達において機能すると思われる
108	NM_004125	BC016319	グアニンシクレオチド結合タンパク質ガンマサブユニット10、情報伝達に關与しているヘテロトリマーGタンパク質複合体の推定成分、Gタンパク質ベータ1 (GNB1)及びベータ2(GNB2)並びに Ras のネズミキナーゼサブユニットと相互作用する
7	NM_000717	MB3670	代謝 カルボニックアンヒドラーゼ IV、重炭酸塩およびプロトン形成する二酸化炭素の可逆水和を触媒する、pH調節に關与している、腎臓での重炭酸塩吸収に關与して作用することがある、欠乏は、真正の近位尿細管性アシトシスに關連することがある
22	NM_153446	AJ517771	ホモサビエンズベータ1.4N-アセチルガラクトサミントランスフェラーゼ(GALGT2)、mRNA
102	NM_001303	U09466	ヘムA:フェルレドキシトランスフェラーゼ、ヘムAの生合成に必要とするフェルレドキシトランスフェラーゼ、遺伝子の欠損又は破壊は、圧迫性麻痺及び1型シヤルコーマリートウース病の傾向がある遺伝性神経障害と關連することがある
108	NM_130468	BC023653	デルマトン-4スルホトランスフェラーゼ-1、デルマトン-4糖鎖生合成に關与してデルマトンのN-アセチルガラクトサミンの O-4 水酸基へのサルフェートの取替を触媒する
			癌関連
21	NM_145897	D89667	プレフォルディン5、折り畳みしていないタンパク質の細胞質内チャペロンの送達に關与しているプレフォルディン5チャペロン複合体の成分、MYC と相互作用し且つ MYC の活性化を抑制する、癌細胞において一般的に置き換えられる腫瘍抑制因子
56	NM_006136	BC005338	キヤピリンタンパク質 2 線(アルファ2)、細胞運動に關与することがあるアクチン結合タンパク質のサブユニット、対応遺伝子は、悪性神経膠腫に關与して増幅し、腫瘍発生に關与することがある
79	NM_004728.1		DEAD-H (Asp-Glu-Ala-Asp/His)ボックスポリペプチド 21、抗癌剤トリアマシンにより阻害される RNA ヘリカーゼ、ssRNA に分子内二次構造を導入するRNA フォルダゼ、スライム疾患における自己抗原
83		AF010315	ホモサビエンズ腫瘍タンパク質p53誘発性タンパク質 11 (TP53I11)、mRNA
84	NM_033137	X65779	線維芽細胞増殖因子1 (酸性)、細胞移動、胎形成、臓器発生、血管新生に關与するマイトジェン及びアポトーシスインヒビター
85	NM_025216	AK024363	ウイングレス型 MMTV 組込み部位ファミリーメンバー-10a、wnt、ファミリーのメンバー、情報伝達及び癌形成に關与することがある、ある種の食道癌、胃癌及び結直腸癌において過剰発現

10

20

30

40

93	NM1021070	AF318354	8つの表皮細胞成長因子(EGF)様ドメイン及び2つの TGF 結合タンパク質ドメインを含むタンパク質、潜在性転写因子結合タンパク質3(マウス Ltbp3)の領域と高い類似性がある
103	NM1001550	BC001272	神経成長因子(NGF)により誘発され、筋分岐に関与し、ならびに組織の増殖及び分岐において発現する、ラット RN13723 との類似性が高いタンパク質
104	NM1198407	U60179	成長ホルモン分泌促進物質受容体、グレルリン(GHRH)及び合成成長ホルモン分泌促進物質と結合するGタンパク質結合受容体、成長ホルモン分泌を調節することができ、発現高値は内分泌腫瘍と関連することがある
88	NM1016041	BC010890	F-LAN-1、肝細胞癌においてアップレギュレーションされるタンパク質、細胞増殖の正の調節に関与する
			アクチン関連
25	NM1006731	U50523	アクチン関連タンパク質 2/3 複合体サブユニット2、アクチン細胞骨格の構築に関与している Arp2/3 複合体の成分、おそらくArp2/3 複合体形成における初期の中間体としてARPC4と直接相互作用する
98	NM1006135	BX646738	キャッピングタンパク質筋 2 線アルファ1、アクチン重合を調節し且つ反矢状リネアクチンキヤンペイング、細胞移動、サルコミア組織化及び筋機能改善とすることがあるアクチンキヤンペイングタンパク質
94	NM1006136	U03269	ホモサビエンスキヤンペイングタンパク質(アクチンフィラメント)筋 2 線、アルファ2(CAPZA2)、mRNA
			細胞分化
39	NM1001858	U08279	XX 型コラーゲンのアルファ1サブユニット、コラーゲンの FACIT ファミリーのメンバー、細胞分化に関与することがある、あるいは機能的に肉腫細胞に結合される
50	NM1006818	AK056089	染色体1qからのALL1融合遺伝子、胸腺、造血及び白血病細胞系で発現されるタンパク質、対称遺伝子は、MLLを含み且つ急性骨髄単球性白血病を引き起こす染色体転位の部位である
51	NM1006168		NK ホモオボクスタファミリー6A、遺伝子発現を調節し且つ細胞分化の制御に関与し、高度保存NKデカヘブチド及びホモオボドメイン領域を含む、DNA結合タンパク質のホモオボドメインファミリーのメンバー、
92	NM1001496	AY358037	GDNF ファミリー受容体アルファ3、GDNF/ニューロリン/バーセフィン受容体ファミリーのグリコシルホスファチジルイノシトール(GPI)連結オーファンメンバー、発達の末梢神経系及び成人知覚神経節及び交感神経節において高度発現する
			他の機能
46	NM1016009	AK001954	SH3ドメイン GFB2 様エンドフィリン B1、C末端における Src ホモロジ-3 (SH3)ドメインを含み、BAX アポトーシス情報伝達経路のレギュレーターとしての役割を果たすことがある
89	NM1021724	M24898	ホモサビエンスキヤンペイングタンパク質ファミリー1、グルーブD、メンバー1(NPR1D1)、mRNA

【表 20】

96	NM_006152	U10485	リンパ系制御遺伝子タンパク質、小胞体の細胞質側の膜タンパク質
90	NM_018304	BC009778	ホスファチジルセリン特異的ホスホリラーゼA1、ホスファチジルセリン及び1-アシナル-L-リノホスファチジルセリンの sn-1 位で脂肪修飾添加分解する、ホスファチジルセリン又はリンホスファチジルセリン介在機能の調節に関与している
87	NM_015900	BC047703	両特異性ホスファターゼ 13、ホスホプロリン、ホスホセリン及びホスホレオニン残基を脱リンすることがある
9	NM_018286	AK095175	機能不明
12	NM_013441	AF176117	機能不明のタンパク質
14	NM_148415		ホモサビエンズダウン症候群発症領域遺伝子1様2(DSCR1L2)、mRNA
15	NM_017845	BC015145	腎臓/脳神経運動制御タイプ2と関連している SCA2(アタキシン 2)との類似性が中程度であるタンパク質
23	NM_015702	BC022559	機能が不明のタンパク質、特徴付けされていないマウス DBPuc28e との類似性が高い
30	NM_173504	AK124773	機能が不明のタンパク質、特徴付けされていないマウス 2010311D03Rk との類似性が高い
33	NM_002336	AK074543	ホモサビエンズ、仮定タンパク質 FLJ37538(FLJ37538)、mRNA
36	NM_152383	BC036113	ホモサビエンズ、低密度リポタンパク質受容体関連タンパク質 6 (LPP6)、mRNA
43	NM_020141	AF164793	ホモサビエンズ、仮定タンパク質 MGC42174(MGC42174)、mRNA
57	NM_004872	BC016374	機能が不明のタンパク質、特徴付けされていないマウス K07F515 との類似性が高い
59	NM_003678	AK025385	機能が不明のタンパク質、特徴付けされていないマウス ORF18 との類似性が高い
62	NM_004321	EX537556	機能が不明のタンパク質、特徴付けされていないマウス Frip との類似性が高い
64	NM_152587	BC029536	ホモサビエンズ、仮定タンパク質 BC009491(LOC151568)、mRNA
75		BC030200.1	ホモサビエンズ、仮定タンパク質 MGC33948(MGC33948)、mRNA
77	NM_174918	BC035847	機能が不明のタンパク質、特徴付けされていないマウス D430039ND5Rk との類似性が低い
82	NM_174899	BC033935	ホモサビエンズ、仮定タンパク質 LOC130888(LOC130888)、mRNA
86	NM_178525	AY248901	ホモサビエンズ、仮定タンパク質 MGC33407(MGC33407)、mRNA
99	NM_025109	AL133017	ホモサビエンズ、仮定タンパク質 FLJ22865(FLJ22865)、mRNA
106	NM_152362	AK024161	ホモサビエンズ、仮定タンパク質 MGC17791(MGC17791)、mRNA
110		XM_088567	不明
111	NM_198458	AK126727	不明
112		BC054888	不明

受託番号は、表5に記載の通りである

10

20

30

40

【図面の簡単な説明】

【0210】

【図1】図1は乳癌を患う個体における発現の変化に関与する種々の因子の可能性のある相互作用を示す。

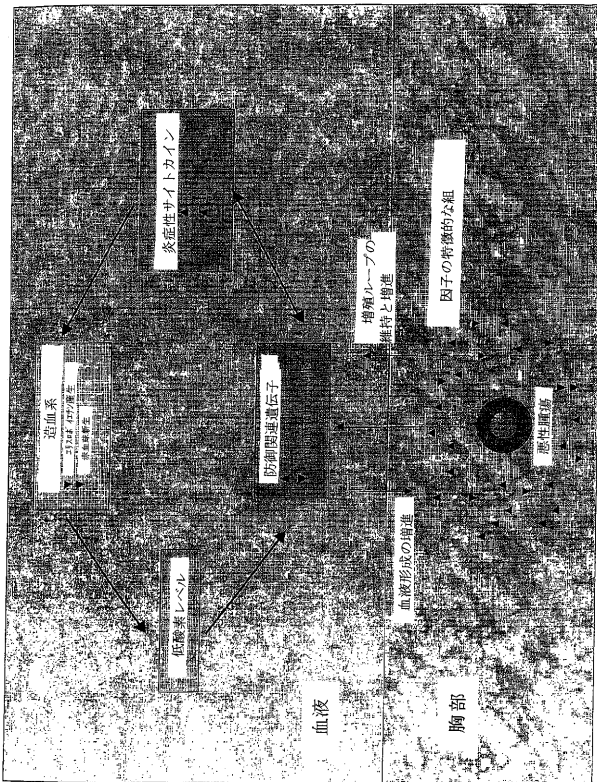
【図2】図2は、PLSRを使用する、35遺伝子に基づくスコアプロットを示す。102個の正常(良性を含む)試料及び乳癌試料を、35の有用遺伝子のデータを用いたPLSR-DAにより作成された類別モデルへのプロジェクションである。PCは主要成分を示し、N及びCはそれぞれ正常試料及び乳癌試料を示す。

50

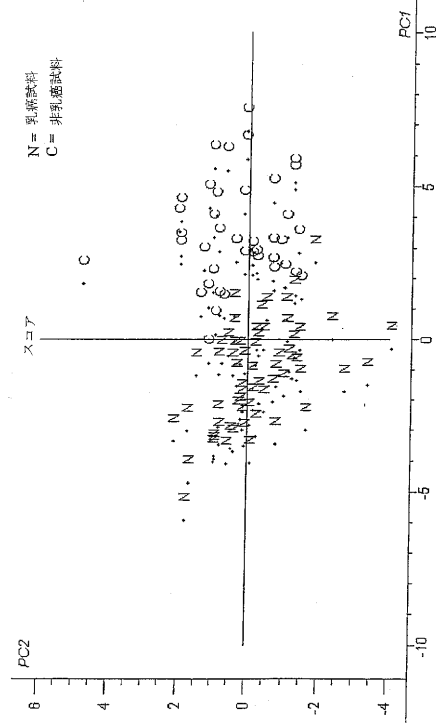
【図3】図3は、PLSRを使用する、35遺伝子に基づく予測プロットを示す。35個のcDNAデータを用いた、3つの主要成分に基づく予測プロットである。

【図4】図4は、乳癌試料及び非乳癌試料における35遺伝子の平均発現レベルを示す。すなわち乳癌の予測に使用される35個の遺伝子発現の平均レベルである。遺伝子は、配置(array)上の位置ID(position ID)で示されている。

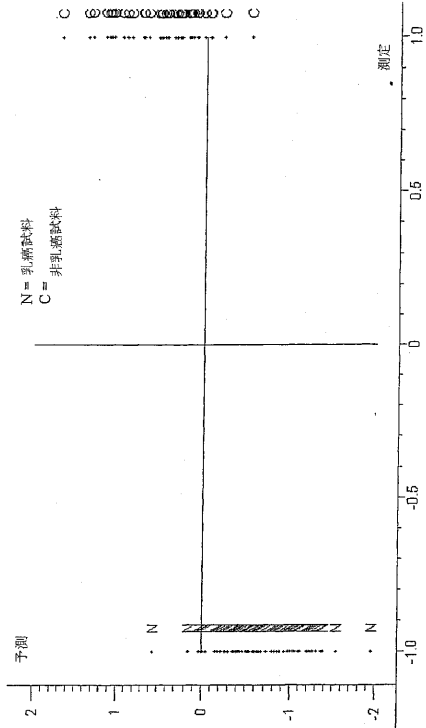
【図1】



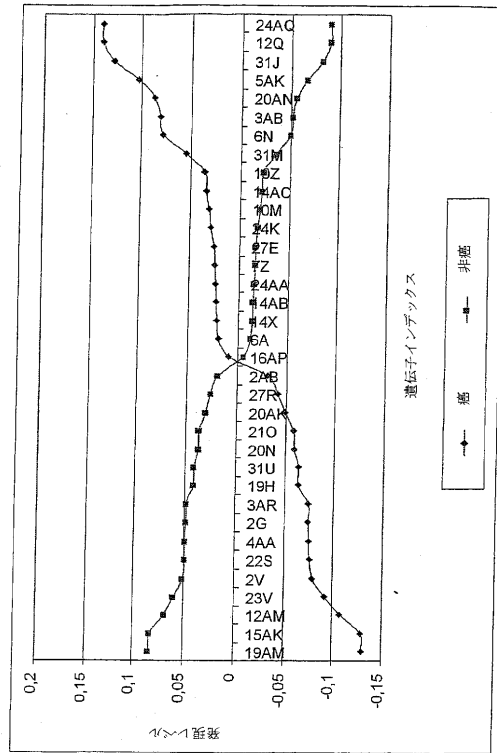
【図2】



【 図 3 】



【 図 4 】



## 【 国際調査報告 】

INTERNATIONAL SEARCH REPORT		Internal Application No PCT/GB2005/002180
<b>A. CLASSIFICATION OF SUBJECT MATTER</b> C12Q1/68		
According to International Patent Classification (IPC) or to both national classification and IPC		
<b>B. FIELDS SEARCHED</b>		
Minimum documentation searched (classification system followed by classification symbols) C12Q		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practical, search terms used) EPO-Internal, WPI Data, PAJ, BIOSIS, EMBASE		
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE GEO PROFILES 'Online! NCBI; 2 January 2004 (2004-01-02), "http://www.ncbi.nih.gov/projects/geo/quer y/acc.cgi?acc=GPL887" XP002353005 retrieved from NCBI Database accession no. GPL887	29-31
Y	abstract	1-28
Y	EP 0 979 308 A (DIAGENIC AS) 16 February 2000 (2000-02-16) the whole document	1-28
X	WO 02/057787 A (YALE UNIVERSITY; ALTIERI, DARIO, C; WEISS, ROBERT, M; SMITH, SHANNON,) 25 July 2002 (2002-07-25) claims 1,2,7,18,26	33-35
-/-		
<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		*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
*E* earlier document but published on or after the international filing date		*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
*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)		*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.
*O* document referring to an oral disclosure, use, exhibition or other means		*S* document member of the same patent family
*P* document published prior to the international filing date but later than the priority date claimed		
Date of the actual completion of the international search <b>9 November 2005</b>	Date of mailing of the international search report <b>23/11/2005</b>	
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax (+31-70) 340-3016	Authorized officer <b>Bellmann, A</b>	

## INTERNATIONAL SEARCH REPORT

Internal	Application No
	PCT/GB2005/002180

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No
A	US 2003/068642 A1 (URNOVITZ HOWARD B) 10 April 2003 (2003-04-10) the whole document	
X	BROOKS ET AL: "Transcript Profiles of Microdissected Tissue in a Mouse Model of Idiopathic Parkinson's disease Using the Ribo-SPIATM Amplification Process" Online! November 2003 (2003-11), page 1, XP002353003 Retrieved from the Internet: URL: <a href="http://www.nugentechnologies.com/pdfs/ASofNEURO-Nov03_AA_Poster.pdf">http://www.nugentechnologies.com/pdfs/ ASofNEURO-Nov03_AA_Poster.pdf</a> 'retrieved on 2005-11-08! the whole document	29-32
X	US 2003/219760 A1 (GORDON GAVIN J ET AL) 27 November 2003 (2003-11-27) claims 1,12,15,77; table 6	33-35
X	US 2003/232350 A1 (AFAR DANIEL ET AL) 18 December 2003 (2003-12-18) paragraph '0068!; claims 1,5,7,17; table 44	33-35
P, X	WO 2004/046382 A (DIAGENIC AS; SHARMA, PRAVEEN; SAHNI, NARINDER, SINGH; LOENNEBORG, ANDE) 3 June 2004 (2004-06-03) the whole document	1-35
A	WO 02/059271 A (GENE LOGIC, INC; ORR, MICHAEL, S; NATION, MICHELE; DIGGANS, JAMES, C;) 1 August 2002 (2002-08-01) the whole document	
A	ALIZADEH A A ET AL: "DISTINCT TYPES OF DIFFUSE LARGE B-CELL LYMPHOMA IDENTIFIED BY GENE EXPRESSION PROFILING" NATURE, NATURE PUBLISHING GROUP, LONDON, GB, vol. 403, 3 February 2000 (2000-02-03), pages 503-512, XP002943414 ISSN: 0028-0836 the whole document	
A	US 2003/175736 A1 (CHINNAIYAN ARUL M ET AL) 18 September 2003 (2003-09-18) the whole document	

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/GB2005/002180

**Box II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)**

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

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

**Box III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)**

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

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

Remark on Protest

The additional search fees were accompanied by the applicant's protest.

No protest accompanied the payment of additional search fees.

International Application No. PCT/GB2005/002180

## FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box II.2

Claims Nos.: -

The subject-matter of independent claim 29 lacks clarity and conciseness in the sense of Article 6 PCT. Claim 29 refers to a set of at least 10 oligonucleotide probes as defined in any one of claims 1 to 28. Claims 1 to 28 however relate to an extremely large number of possible probe sets, as they are defining the probes only via gene families from which the probes are supposed to be chosen. Hence, the subject-matter of said independent claim 29 includes an extremely large number of possible set of oligonucleotides. Even without taking into account the fact that "oligonucleotides which are derived from such a sequence" or "functionally equivalent oligonucleotides" are also encompassed in its scope, this claim contains so many options, variables and possible permutations that its scope could not be meaningfully searched.

Support and disclosure in the sense of Article 6 and 5 PCT is to be found for only a very small proportion of the probes claimed, see examples 1 and 2, pages 46 to 54. The search of claim 29 was restricted to those probe sets which appear to be supported by the examples, namely the 109 probes identified in example 2 to be informative for breast cancer. (As in example 1 the probes itself are not disclosed but only the gene names and sequences of said genes without any indication where the probes are localized the probe set is not described in a sufficiently clear and concise way to be meaningfully searched.) Table 3 and 4 provide the probe sequences of informative probes identified in example 2. The search has therefore be limited to the probe sets disclosed in Table 3 and 4 respectively.

Independent claims 1, 2, 3, 30 and 32 refer to methods using said probe sets, kits comprising said probe sets and uses of said probe sets as defined in claim 29. As the probe sets are not described in a sufficiently clear and concise way, these claims also lack conciseness in the sense of Article 6 PCT. Therefore the search has been limited as described above to kits comprising the probe sets as defined in Table 3 and 4, uses of said probe sets and methods using said probes sets.

The applicant's attention is drawn to the fact that claims relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure. If the application proceeds into the regional phase before the EPO, the applicant is reminded that a search may be carried out during examination before the EPO (see EPO Guideline C-VI, 8.5), should the problems which led to the Article 17(2) declaration be overcome.

## INTERNATIONAL SEARCH REPORT

Information on patent family members

Internal	Application No
	PCT/GB2005/002180

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
EP 0979308	A	16-02-2000	AT 286143 T	15-01-2005
			AU 7222698 A	24-11-1998
			DE 69828392 D1	03-02-2005
			ES 2230688 T3	01-05-2005
			WO 9849342 A1	05-11-1998
			JP 2001526534 T	18-12-2001
			PT 979308 T	31-03-2005
			US 2002022222 A1	21-02-2002
WO 02057787	A	25-07-2002	CA 2437305 A1	25-07-2003
			CN 1636139 A	06-07-2005
			EP 1350114 A2	08-10-2003
			JP 2004533602 T	04-11-2004
US 2003068642	A1	10-04-2003	US 2003036077 A1	20-02-2003
US 2003219760	A1	27-11-2003	NONE	
US 2003232350	A1	18-12-2003	NONE	
WO 2004046382	A	03-06-2004	AU 2003286262 A1	15-06-2004
			CA 2506887 A1	03-06-2004
			EP 1565574 A2	24-08-2005
WO 02059271	A	01-08-2002	NONE	
US 2003175736	A1	18-09-2003	CA 2456184 A1	13-02-2003
			EP 1511855 A2	09-03-2005
			JP 2005518522 T	23-06-2005
			WO 03012067 A2	13-02-2003
			US 2005136493 A1	23-06-2005

## フロントページの続き

(51) Int.Cl.		F I		テーマコード(参考)
<b>G 0 1 N 37/00</b>	<b>(2006.01)</b>	G 0 1 N 33/53		M
		G 0 1 N 37/00	1 0 2	

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

(72) 発明者 ロンネボルグ アンデルス

ノルウェー エヌ - 1 4 3 0 アース フェルツパトヴェーン 1 4

Fターム(参考) 4B024 AA11 AA19 CA04 CA05 CA06 CA09 CA10 CA12 GA18 HA08  
 HA09 HA12 HA14 HA19  
 4B063 QA01 QA18 QA19 QQ03 QQ08 QQ42 QQ44 QQ47 QQ48 QQ53  
 QR07 QR08 QR14 QR32 QR36 QR42 QR55 QR62 QR66 QR77  
 QR82 QS03 QS12 QS16 QS25 QS34 QS36 QS39 QX02

专利名称(译)	用于癌症诊断的寡核苷酸		
公开(公告)号	<a href="#">JP2008501332A</a>	公开(公告)日	2008-01-24
申请号	JP2007514130	申请日	2005-06-02
[标]申请(专利权)人(译)	迪亚增益尼克等耶稣		
申请(专利权)人(译)	Daiagenikku Eiesu		
[标]发明人	シャルマ プラヴェーン ロンネボルグ アンデルス		
发明人	シャルマ プラヴェーン ロンネボルグ アンデルス		
IPC分类号	C12Q1/68 C12Q1/02 C12N15/09 G01N33/574 G01N33/53 G01N37/00		
CPC分类号	C12Q1/6886 C12Q2600/112 Y10T436/10		
FI分类号	C12Q1/68.ZNA.A C12Q1/02 C12N15/00.F C12N15/00.A G01N33/574.A G01N33/53.M G01N37/00.102		
F-TERM分类号	4B024/AA11 4B024/AA19 4B024/CA04 4B024/CA05 4B024/CA06 4B024/CA09 4B024/CA10 4B024/CA12 4B024/GA18 4B024/HA08 4B024/HA09 4B024/HA12 4B024/HA14 4B024/HA19 4B063/QA01 4B063/QA18 4B063/QA19 4B063/QQ03 4B063/QQ08 4B063/QQ42 4B063/QQ44 4B063/QQ47 4B063/QQ48 4B063/QQ53 4B063/QR07 4B063/QR08 4B063/QR14 4B063/QR32 4B063/QR36 4B063/QR42 4B063/QR55 4B063/QR62 4B063/QR66 4B063/QR77 4B063/QR82 4B063/QS03 4B063/QS12 4B063/QS16 4B063/QS25 4B063/QS34 4B063/QS36 4B063/QS39 4B063/QX02		
代理人(译)	鈴木 俊一郎		
优先权	2004012301 2004-06-02 GB		
其他公开文献	JP5060945B2		
外部链接	<a href="#">Espacenet</a>		

#### 摘要(译)

本发明涉及用于评估细胞中基因转录水平的寡核苷酸探针，其可用于分析方法，特别是诊断方法。涉及编码参与蛋白质合成和/或稳定性，或特定用途的蛋白质，以创建用于癌症诊断防御的转印图案的调节和/或染色体重塑基因本发明编码的蛋白质我们提供对应于该基因的寡核苷酸组提供。这种探针以试剂盒的形式方便地提供。不同类型的探针，创造基因表达模式，和不同的癌症或癌阶段的鉴定，可用于诊断或监测技术。此外，本发明提供了这样的使用，以获得的试剂盒通过所述基因编码的标记多肽的分析一组和相关的方法，以及含有这样的一组的特性的表达谱。

試料	試料数	正確に予測	不正に予測	正率
対照	78	67	11	0.86
乳癌	44	26	18	0.59