

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第5913785号
(P5913785)

(45) 発行日 平成28年5月25日 (2016. 5. 25)

(24) 登録日 平成28年4月8日 (2016. 4. 8)

(51) Int. Cl.

F I

C 1 2 Q 1/68 (2006. 01)

C 1 2 Q 1/68 Z N A A

C 1 2 M 1/00 (2006. 01)

C 1 2 M 1/00 A

C 1 2 N 15/09 (2006. 01)

C 1 2 N 15/00 A

G O 1 N 33/53 (2006. 01)

C 1 2 N 15/00 F

G O 1 N 37/00 (2006. 01)

G O 1 N 33/53 M

請求項の数 36 (全 74 頁) 最終頁に続く

(21) 出願番号 特願2009-548313 (P2009-548313)
 (86) (22) 出願日 平成20年1月31日 (2008. 1. 31)
 (65) 公表番号 特表2010-516292 (P2010-516292A)
 (43) 公表日 平成22年5月20日 (2010. 5. 20)
 (86) 国際出願番号 PCT/US2008/001339
 (87) 国際公開番号 W02008/094678
 (87) 国際公開日 平成20年8月7日 (2008. 8. 7)
 審査請求日 平成23年1月21日 (2011. 1. 21)
 審判番号 不服2015-289 (P2015-289/J1)
 審判請求日 平成27年1月7日 (2015. 1. 7)
 (31) 優先権主張番号 60/898, 963
 (32) 優先日 平成19年1月31日 (2007. 1. 31)
 (33) 優先権主張国 米国 (US)

(73) 特許権者 508356065
 セレラ コーポレーション
 アメリカ合衆国 カリフォルニア 945
 02, アラメダ, ハーバー ベイ パ
 ークウェイ 1401
 (74) 代理人 100078282
 弁理士 山本 秀策
 (74) 代理人 100113413
 弁理士 森下 夏樹
 (72) 発明者 ラウ, キット
 アメリカ合衆国 カリフォルニア 945
 02, アラメダ, ハーバー ベイ パ
 ークウェイ 1401, アン アプレラ
 ビジネス ユニット, セレラ

最終頁に続く

(54) 【発明の名称】 乳癌の遠隔転移を予測するための分子予後サインおよびその使用

(57) 【特許請求の範囲】

【請求項 1】

乳癌患者の腫瘍転移のリスクの測定を補助する方法であって、

(a) 前記乳癌患者のエストロゲン受容体陽性の腫瘍細胞の試料における遺伝子 C E N P A、P K M Y T 1、M E L K、M Y B L 2、B U B 1、R A C G A P 1、T K 1、U B E 2 S、D C 1 3、R F C 4、P R R 1 1、D I A P H 3、O R C 6 L および C C N B 1 の発現レベルを測定し、前記発現レベルを転移スコア (M S) を決定するためのアルゴリズムに挿入し、それによって、前記遺伝子の発現レベルに基づく前記転移スコア (M S) を得る工程を包含し、

(b) ここで、前記転移スコアと所定の転移スコアのカットオフ閾値 (M S 閾値) との比較が、前記乳癌患者の腫瘍転移のリスクを示し、

ここで、C E N P A は配列番号 35 のヌクレオチド配列を含み、P K M Y T 1 は配列番号 36 のヌクレオチド配列を含み、M E L K は配列番号 37 のヌクレオチド配列を含み、M Y B L 2 は配列番号 38 のヌクレオチド配列を含み、B U B 1 は配列番号 39 のヌクレオチド配列を含み、R A C G A P 1 は配列番号 40 のヌクレオチド配列を含み、T K 1 は配列番号 41 のヌクレオチド配列を含み、U B E 2 S は配列番号 42 のヌクレオチド配列を含み、D C 1 3 は配列番号 43 のヌクレオチド配列を含み、R F C 4 は配列番号 44 のヌクレオチド配列を含み、P R R 1 1 は配列番号 45 のヌクレオチド配列を含み、D I A P H 3 は配列番号 46 のヌクレオチド配列を含み、O R C 6 L は配列番号 47 のヌクレオチド配列を含み、そして、C C N B 1 は配列番号 48 のヌクレオチド配列を含む、

10

20

方法。

【請求項 2】

工程 (b) において、前記所定の M S 閾値より高い M S は、腫瘍転移のリスクの増加を示す、請求項 1 に記載の方法。

【請求項 3】

工程 (b) において、前記所定の M S 閾値より低い M S は、腫瘍転移のリスクの減少を示す、請求項 1 に記載の方法。

【請求項 4】

工程 (b) において、2 以上の M S 閾値に対する前記転移スコアの比較が、前記乳癌患者について腫瘍転移のリスクを示す、請求項 1 に記載の方法。

10

【請求項 5】

前記乳癌患者が、リンパ節において検出可能な腫瘍細胞を有さない、請求項 1 に記載の方法。

【請求項 6】

前記乳癌患者が、腫瘍転移を発症するリスクが増加している、請求項 1 に記載の方法。

【請求項 7】

前記エストロゲン受容体陽性の腫瘍細胞が、ホルマリン固定されたパラフィン包埋された切片に由来する、請求項 1 に記載の方法。

【請求項 8】

前記エストロゲン受容体陽性の腫瘍細胞が、腫瘍生検由来である、請求項 1 に記載の方法。

20

【請求項 9】

前記エストロゲン受容体陽性の腫瘍細胞が、凍結腫瘍組織由来である、請求項 1 に記載の方法。

【請求項 10】

前記遺伝子の m R N A が、エストロゲン受容体陽性の腫瘍細胞から得られ、c D N A に逆転写され、前記 c D N A のポリメラーゼ連鎖反応増幅によって検出される、請求項 1 に記載の方法。

【請求項 11】

前記 m R N A が、逆転写および P C R 増幅前に濃縮される、請求項 10 に記載の方法。

30

【請求項 12】

前記遺伝子の各々の m R N A が、逆転写され、表 3 の配列番号 1 ~ 28 に示される対応する遺伝子と関連した 2 つのプライマーによって増幅される、請求項 1 に記載の方法。

【請求項 13】

前記遺伝子の発現レベルが、ある内在性対照遺伝子または 2 以上の内在性対照遺伝子の平均の発現レベルに対して標準化される、請求項 1 に記載の方法。

【請求項 14】

前記内在性対照遺伝子が、N U P 2 1 4、P P I G および S L U 7 で構成される群から選択され、ここで、N U P 2 1 4 は配列番号 49 のヌクレオチド配列を含み、P P I G は配列番号 50 のヌクレオチド配列を含み、そして、S L U 7 は配列番号 51 のヌクレオチド配列を含む、請求項 13 に記載の方法。

40

【請求項 15】

前記内在性対照遺伝子の m R N A が、逆転写され、表 3 に列挙される、配列番号 29 ~ 34 の前記内在性対照遺伝子と関連した 2 つのプライマーによって増幅される、請求項 14 に記載の方法。

【請求項 16】

前記発現レベルが、マイクロアレイによって検出される、請求項 1 に記載の方法。

【請求項 17】

乳癌患者の腫瘍転移のリスクの測定を補助する方法であって、

(a) 前記乳癌患者のエストロゲン受容体陽性の腫瘍細胞の試料における遺伝子 C E N

50

P A、P K M Y T 1、M E L K、M Y B L 2、B U B 1、R A C G A P 1、T K 1、U B E 2 S、D C 1 3、R F C 4、P R R 1 1、D I A P H 3、O R C 6 LおよびC C N B 1の発現レベルを測定し、前記発現レベルを転移スコア(M S)を決定するためのアルゴリズムに挿入し、それによって、前記遺伝子の発現レベルに基づく前記転移スコア(M S)を得る工程を包含し、

(b)ここで、前記転移スコアと所定の転移スコアのカットオフ閾値(M S 閾値)との比較が、腫瘍転移のリスクを示し、

ここで、C E N P Aは配列番号35のヌクレオチド配列を含み、P K M Y T 1は配列番号36のヌクレオチド配列を含み、M E L Kは配列番号37のヌクレオチド配列を含み、M Y B L 2は配列番号38のヌクレオチド配列を含み、B U B 1は配列番号39のヌクレオチド配列を含み、R A C G A P 1は配列番号40のヌクレオチド配列を含み、T K 1は配列番号41のヌクレオチド配列を含み、U B E 2 Sは配列番号42のヌクレオチド配列を含み、D C 1 3は配列番号43のヌクレオチド配列を含み、R F C 4は配列番号44のヌクレオチド配列を含み、P R R 1 1は配列番号45のヌクレオチド配列を含み、D I A P H 3は配列番号46のヌクレオチド配列を含み、O R C 6 Lは配列番号47のヌクレオチド配列を含み、そして、C C N B 1は配列番号48のヌクレオチド配列を含み、

ここで、前記転移スコア(M S)が、下記：

【数8】

$$MS = a0 + \sum_{i=1}^M ai * Gi$$

(式中、

Mは14であり、G iは14個の前記遺伝子の各遺伝子(i)の発現レベルであり、a 0は0でありかつa iは1であるか、または、a 0は0.022であり、a iは前記遺伝子の各々について表2にある値に対応する)

によって計算される、方法。

【請求項18】

前記各々の遺伝子(i)の発現レベルG iが、下記：

$$(Ct) = (Ct_{GOI} - Ct_{EC})_{testRNA} - (Ct_{GOI} - Ct_{EC})_{refRNA}$$

(式中、

C tは指数関数的な標的増幅のP C R閾値サイクルであり、G O Iは対象とする遺伝子であり、E Cは内在性対照であり、t e s t R N Aは患者の試料RNAであり、r e f R N Aは参照RNAである)

によって(C t)値として算定される、請求項17に記載の方法。

【請求項19】

乳癌患者の腫瘍転移のリスクの測定を補助する方法であって、

(a)前記乳癌患者のエストロゲン受容体陽性の腫瘍細胞の試料における遺伝子C E N P A、P K M Y T 1、M E L K、M Y B L 2、B U B 1、R A C G A P 1、T K 1、U B E 2 S、D C 1 3、R F C 4、P R R 1 1、D I A P H 3、O R C 6 LおよびC C N B 1の発現レベルを測定し、前記発現レベルを転移スコア(M S)を決定するためのアルゴリズムに挿入し、それによって、前記遺伝子の発現レベルに基づく前記転移スコア(M S)を得る工程を包含し、

(b)ここで、前記転移スコアと所定の転移スコアのカットオフ閾値(M S 閾値)との比較が、腫瘍転移のリスクを示し、

ここで、C E N P Aは配列番号35のヌクレオチド配列を含み、P K M Y T 1は配列番号36のヌクレオチド配列を含み、M E L Kは配列番号37のヌクレオチド配列を含み、M Y B L 2は配列番号38のヌクレオチド配列を含み、B U B 1は配列番号39のヌクレオチド配列を含み、R A C G A P 1は配列番号40のヌクレオチド配列を含み、T K 1は配列番号41のヌクレオチド配列を含み、U B E 2 Sは配列番号42のヌクレオチド配列

を含み、D C 1 3 は配列番号 4 3 のヌクレオチド配列を含み、R F C 4 は配列番号 4 4 のヌクレオチド配列を含み、P R R 1 1 は配列番号 4 5 のヌクレオチド配列を含み、D I A P H 3 は配列番号 4 6 のヌクレオチド配列を含み、O R C 6 L は配列番号 4 7 のヌクレオチド配列を含み、そして、C C N B 1 は配列番号 4 8 のヌクレオチド配列を含み、

ここで、前記転移スコア (M S) が、下記：

【数 8 A】

$$MS = a0 + \sum_{i=1}^M ai * Gi$$

(式中、

10

M は 1 4 であり、G i は 1 4 個の前記遺伝子の各遺伝子 (i) の発現レベルであり、a 0 が 0 であり、a i が 1 である)

によって計算される、方法。

【請求項 2 0】

前記各々の遺伝子 (i) の発現レベル G i が、下記：

$$(Ct) = (Ct_{GOI} - Ct_{EC})_{testRNA} - (Ct_{GOI} - Ct_{EC})_{refRNA}$$

(式中、

C t は指数関数的な標的増幅の P C R 閾値サイクルであり、G O I は対象とする遺伝子であり、E C は内在性対照であり、t e s t R N A は患者の試料 R N A であり、r e f R N A は参照 R N A である)

20

によって (Ct) 値として算定される、請求項 1 9 に記載の方法。

【請求項 2 1】

前記転移スコア (M S) が、下記：

【数 9】

$$MS = a0 + b * \left[\sum_{i=1}^M ai * Gi \right]$$

(式中、

30

M が 1 4 であり、G i が 1 4 個の前記遺伝子の各遺伝子 (i) の基準化された発現レベルであり、a 0 が 0 . 0 2 2 であり、b が - 0 . 2 5 1 であり、a i が前記遺伝子の各々について表 2 に示される値に対応する か、または、a 0 が 0 であり、b が - 1 であり、a i が 1 である)

によって計算される、請求項 1 に記載の方法。

【請求項 2 2】

前記各々の遺伝子 (i) の発現レベル G i が、下記：

$$(Ct) = (Ct_{GOI} - Ct_{EC})_{testRNA} - (Ct_{GOI} - Ct_{EC})_{refRNA}$$

(式中、

40

C t は指数関数的な標的増幅の P C R 閾値サイクルであり、G O I は対象とする遺伝子であり、E C は内在性対照であり、t e s t R N A は患者の試料 R N A であり、r e f R N A は参照 R N A である)

によって (Ct) 値として算定される、請求項 2 1 に記載の方法。

【請求項 2 3】

前記転移スコア (M S) が、下記：

【数 1 0】

$$MS = a0 + b * \left[\sum_{i=1}^M ai * Gi \right]$$

50

(式中、

Mが14であり、 G_i が14個の前記遺伝子の各遺伝子(i)の発現レベルであり、 a_0 が0.8657であり、 b が-0.04778であり、 a_i が1であるかまたは、 a_0 が0であり、 b が-1であり、 a_i が1である)

によって計算される、請求項1に記載の方法。

【請求項24】

前記各々の遺伝子(i)の発現レベル G_i が、下記：

$$(Ct) = (Ct_{GOI} - Ct_{EC})_{testRNA} - (Ct_{GOI} - Ct_{EC})_{refRNA}$$

(式中、

Ct は指数関数的な標的増幅のPCR閾値サイクルであり、GOIは対象とする遺伝子であり、ECは内在性対照であり、 $testRNA$ は患者の試料RNAであり、 $refRNA$ は参照RNAである)

によって (Ct) 値として算定される、請求項23に記載の方法。

【請求項25】

前記転移スコア(MS)が、下記：

【数11】

$$MS(新規) = -1/14 * \left[\sum_{i=1}^{14} G_i \right]$$

(式中、 G_i が14個の前記遺伝子の各遺伝子(i)の発現レベルである)

によって計算される、請求項1に記載の方法。

【請求項26】

前記各々の遺伝子(i)の発現レベル G_i が、下記：

$$(Ct) = (Ct_{GOI} - Ct_{EC})_{testRNA} - (Ct_{GOI} - Ct_{EC})_{refRNA}$$

(式中、

Ct は指数関数的な標的増幅のPCR閾値サイクルであり、GOIは対象とする遺伝子であり、ECは内在性対照であり、 $testRNA$ は患者の試料RNAであり、 $refRNA$ は参照RNAである)

によって (Ct) 値として算定される、請求項25に記載の方法。

【請求項27】

前記転移スコア(MS)が、14個の前記遺伝子の各遺伝子の発現レベルを一緒に加えることによって計算される、請求項1に記載の方法。

【請求項28】

前記各遺伝子の発現レベルが (Ct) 値によって表される、請求項27に記載の方法。

【請求項29】

前記 (Ct) 値の総計が負の係数によって乗じられるか、または、前記 (Ct) 値の各々が負の係数によって乗じられる、請求項28に記載の方法。

【請求項30】

前記負の係数が-1である、請求項29に記載の方法。

【請求項31】

遺伝子CENPA、PKMYT1、MELK、MYBL2、BUB1、RACGAP1、TK1、UBE2S、DC13、RFC4、PRR11、DIAPH3、ORC6LおよびCCNB1の発現レベルを検出するための試薬、および酵素；並びにバッファーを含む、請求項1～30に記載の方法において使用するためのキット。

【請求項32】

1以上の内在性対照遺伝子の発現レベルを検出するための試薬またはその組み合わせをさ

10

20

30

40

50

らに含む、請求項 3 1 に記載のキット。

【請求項 3 3】

前記内在性対照遺伝子が、N U P 2 1 4、P P I G および S L U 7 で構成される群から選択される、請求項 3 2 に記載のキット。

【請求項 3 4】

遺伝子 C E N P A の配列番号 3 5 のヌクレオチド配列からなる領域、P K M Y T 1 の配列番号 3 6 のヌクレオチド配列からなる領域、M E L K の配列番号 3 7 のヌクレオチド配列からなる領域、M Y B L 2 の配列番号 3 8 のヌクレオチド配列からなる領域、B U B 1 の配列番号 3 9 のヌクレオチド配列からなる領域、R A C G A P 1 の配列番号 4 0 のヌクレオチド配列からなる領域、T K 1 の配列番号 4 1 のヌクレオチド配列からなる領域、U B E 2 S の配列番号 4 2 のヌクレオチド配列からなる領域、D C 1 3 の配列番号 4 3 のヌクレオチド配列からなる領域、R F C 4 の配列番号 4 4 のヌクレオチド配列からなる領域、P R R 1 1 の配列番号 4 5 のヌクレオチド配列からなる領域、D I A P H 3 の配列番号 4 6 のヌクレオチド配列からなる領域、O R C 6 L の配列番号 4 7 のヌクレオチド配列からなる領域および C C N B 1 の配列番号 4 8 のヌクレオチド配列からなる領域に配列特異的な様式でハイブリダイズする、複数のポリヌクレオチドを含む、請求項 1 ~ 3 0 に記載の方法において使用するためのマイクロアレイ。

10

【請求項 3 5】

内在性対照遺伝子の領域に配列特異的な様式でハイブリダイズする複数のポリヌクレオチドをさらに含む、請求項 3 4 に記載のマイクロアレイ。

20

【請求項 3 6】

前記内在性対照遺伝子が、N U P 2 1 4、P P I G および S L U 7 からなる群から選択される、請求項 3 5 に記載のマイクロアレイ。

【発明の詳細な説明】

【技術分野】

【0 0 0 1】

関連する出願への相互参照

本願は、2007年1月31日に出願された米国仮特許出願第60/898,963号の利益を主張する。米国仮特許出願第60/898,963号の内容は、本願中にその全体が参考として本明細書により援用される。

30

【0 0 0 2】

本発明は、乳癌転移の予後に関する。特に、本発明は、乳癌患者のリンパ節転移陰性のエストロゲン受容体(E R)陽性腫瘍の転移のリスクの予測に有用である多重遺伝子の予後サインに関する。多重遺伝子の予後サインは14個の遺伝子を含み、乳癌患者のE R陽性腫瘍に含まれるそれらのm R N Aは、ホルマリン固定のパラフィン包埋(F F P E)の組織切片から得ることができ、それらの発現レベルは当該技術分野において知られている方法によって測定される。このようにして、本発明は、リンパ節転移陰性のE R陽性乳癌の遠隔転移のリスクを評価するための日常的な臨床検査室試験に適している。

【背景技術】

【0 0 0 3】

40

乳癌は、複雑な異種疾患である。乳癌の早期の検出により、成功的な治療および回復が改善される。日常的に検査されているマンモグラフィーは、ステージIの乳癌の検出を増やし、付随して、より多くの女性がリンパ節転移陰性腫瘍であると診断される。(非特許文献1)。米国において、毎年、乳癌であると診断されたおよそ240,000人の女性のうち約43%がリンパ節転移陰性である。

【0 0 0 4】

現在のガイドラインに基づくと、85~90%のリンパ節転移陰性の患者が、外科処置後、それによる全身アジュバント療法の候補者となる。このような全身アジュバント療法には、化学療法およびホルモン療法が含まれてもよい。しかしながら、局所的処置(乳腺切除または腫瘍摘出手術および放射線療法)を受けているリンパ節転移陰性乳癌の女性の

50

うち約60～70%が遠隔再発を経験することはない。乳腺患者の処置決定は、転移に関する各々の患者のリスクの評価、複数の臨床的であり、病理組織学的なパラメータを用いた処置への応答から恩恵を受ける。

【0005】

最近のいくつかの研究では、マイクロアレイが使用され、患者の遺伝子発現プロフィールも有用な予後情報を提供することができることを示している。これらの研究の一部は、それらの大きさ、それらの検定の程度により、集中した注目を浴びている。(非特許文献2；非特許文献3；非特許文献4；非特許文献5；非特許文献6)。

【0006】

van 't Veer, Daiら(非特許文献2)によって同定された70個の遺伝子の予後サインに関して集められた、結果として生じた信頼は、欧州試験であるリンパ節転移陰性疾患のためのマイクロアレイによって化学療法を避けられ得る(the Microarray for Node-Negative Disease May Avoid Chemotherapy: MINDACT)にその取り込みをもたらしている。同様に、非特許文献7によって記載されているPCRに基づく21個の遺伝子の予測サインは、北アメリカの乳癌インターグループ(The Breast Cancer Intergroup of North America)による第III相試験(臨床癌試験の評価のためのプログラム(Program for the Assessment of Clinical Cancer Tests: PACCT))に含まれている。(非特許文献8)。

【0007】

非特許文献9による21個の遺伝子の予測サイン(5個の標準化遺伝子を含む)は、タモキシフェン処置患者に由来された。そのサインの独立性は、従来の免疫組織化学(IHC)試験にすでに用いられていた遺伝子および/またはタンパク質とのその実質的な重複により関心が寄せられてきた。(非特許文献10)。ER陽性の乳癌患者に対する標準的なホルモン療法は、タモキシフェン単独から、タモキシフェン+アロマターゼ阻害剤の連続使用へ、またはアロマターゼ阻害剤単独へと変更されつつある。(非特許文献11；非特許文献12)。タモキシフェン処置に独立している予後ツールは、アロマターゼ阻害剤の摂取を計画している患者に対する基準リスクの測定の提供に重要であり得る。

【0008】

このようにして、乳癌患者の遠隔転移のリスクを予測する日常的な臨床検査室試験に用いることができる遺伝子に基づく予後アッセイは必要である。理想的には、このアッセイは、相対的に少数の遺伝子の発現レベルの測定を必要とし、このような遺伝子によってコードされるmRNAは、FFPE腫瘍切片などの日常的な回収法によって保存された腫瘍組織から容易に得ることができる。遠隔転移のリスクに関する情報は、乳癌患者、特に初期段階のリンパ節転移陰性患者の処置戦略の指針に用いることができ、その結果、遠隔転移のリスクが高い患者は適切に処置され、遠隔転移のリスクが低い患者はある種の処置の副作用から免れる場合がある。

【先行技術文献】

【非特許文献】

【0009】

【非特許文献1】B. Cady, 1997, Surg Oncol Clin N Am 6: 195 - 202

【非特許文献2】L J van 't Veer, H. Dai, 2002, Nature 415: 530 - 536

【非特許文献3】M J van de Vijver, Y D He, 2002, N Engl J Med 347: 1999 - 2009

【非特許文献4】Y. Wang, J G Klijn, 2005, Lancet 365: 671 - 679

【非特許文献5】H. Dai, U van 't Veer, 2005, Cancer

10

20

30

40

50

Res 15 : 4059 - 4066

【非特許文献6】HY Chang, DS Nuytenら, 2005, Proc Natl Acad Sci USA 102 : 3738 - 3743

【非特許文献7】SP Paik, S. Shakら, 2004, N Engl J Med 351 : 2817 - 2826

【非特許文献8】VG Kaklamani and WJ Gradishar, 2006, Curr Treat Options Oncol 7 : 123 - 8

【非特許文献9】SP Paik, 2004, N Engl J Med 351 : 2817 - 2826

【非特許文献10】DR Carrizosa and LA Carey, 2005, The American Journal of Oncology Review 4 : 7 - 10 10

【非特許文献11】EP Winer, C. Hudisら, 2005, J Clin Oncol 23 : 619 - 629

【非特許文献12】SM Swain, 2005, N Engl J Med 353 : 2807 - 9

【発明の概要】

【課題を解決するための手段】

【0010】

本発明は、乳癌患者のER陽性腫瘍の転移のリスクを予測するための14個の遺伝子サインに関する。本発明は、部分的には、初期段階のリンパ節転移陰性のER陽性患者の試験に基づき、この患者の大部分は、一次診断後の治療決定を導くための更なる情報を必要とする。本発明の分子サインに含まれる14個の遺伝子を表2に開示する。当業者は、多数の可能な供給源から得られるRNAを用いて、本明細書に記載されている14個の遺伝子に関して発現プロファイリングを行い、次に、この発現データを、予後転移スコアを決定するために提供されたアルゴリズムに挿入することができる。

20

【0011】

本発明の一局面では、乳癌患者のエストロゲン受容体陽性の腫瘍細胞におけるCENPA、PKMYT1、MELK、MYBL2、BUB1、RACGAP1、TK1、UBE2S、DC13、RFC4、PRR11、DIAPH3、ORC6LおよびCCNB1として知られる遺伝子のmRNA発現を測定し、これらの遺伝子のmRNA発現レベルに基づく腫瘍転移のリスクを予測することを含む、乳癌患者の腫瘍転移と関連したリスクを測定する方法に関する。

30

【0012】

本発明の別の局面では、乳癌患者のエストロゲン受容体陽性の腫瘍細胞における遺伝子CENPA、PKMYT1、MELK、MYBL2、BUB1、RACGAP1、TK1、UBE2S、DC13、RFC4、PRR11、DIAPH3、ORC6LおよびCCNB1の発現レベルを測定し、それにより、これらの遺伝子の発現レベルに基づく転移スコア(MS)を得て、この転移スコアと所定の転移スコアカットポイント(MS閾値)とを比較することによってこの乳癌患者について腫瘍転移のリスクを測定することを含む、乳癌患者の腫瘍転移と関連したリスクを測定する方法に関する。

40

【0013】

本発明の更なる局面では、乳癌患者は、そのMSが所定のMS閾値よりも高い場合に、腫瘍転移のリスクが増加していると決定される。

【0014】

本発明の別の局面では、乳癌患者は、そのMSが所定のMS閾値よりも低い場合に、腫瘍転移のリスクが減少していると決定される。

【0015】

本発明の一局面では、乳癌患者の腫瘍転移と関連したリスクを測定する方法に関し、その場合、14個の遺伝子サインのmRNAはER陽性の腫瘍細胞から得られ、cDNAに

50

逆転写され、ポリメラーゼ鎖反応増幅によって検出される。

【 0 0 1 6 】

本発明の別の局面では、乳癌患者の腫瘍転移と関連したリスクを測定する方法に関し、その場合、E R 陽性の腫瘍細胞のm R N A は逆転写され、表 3 に記載される、配列番号 1 ~ 3 4 の各遺伝子と関連した 2 つのプライマーによって増幅される。

【 0 0 1 7 】

本発明の別の局面では、乳癌患者の腫瘍転移と関連したリスクを測定する方法に関し、その場合、E R 陽性の腫瘍細胞からのm R N A 発現の測定は、内在性対照（単数または複数）として、N U P 2 1 4、P P I G および S L U 7、またはそれらの組み合わせとして知られる遺伝子のいずれか 1 つのm R N A 発現に対して標準化される。

10

【 0 0 1 8 】

本発明の別の局面では、乳癌患者の腫瘍転移と関連したリスクを測定する方法に関し、その場合、E R 陽性の腫瘍細胞からのm R N A 発現は、マイクロアレイによって検出される。

【 0 0 1 9 】

本発明の別の局面では、乳癌患者の腫瘍転移と関連したリスクを測定する方法に関し、その場合、m R N A 発現は、下記：

【 0 0 2 0 】

【数 1】

$$MS = a0 + \sum_{i=1}^M ai * Gi$$

20

（式中、

M は 1 4 であり、G i は 1 4 個の遺伝子の各遺伝子（i）の標準化された発現レベルであり、a 0 は 0 . 0 2 2 であり、a i は 1 4 個の遺伝子サインにおける遺伝子の各々について表 2 に示されている値に対応する）

によって転移スコア（MS）に計算される。

【 0 0 2 1 】

本発明の別の局面では、乳癌患者の腫瘍転移と関連したリスクを測定する方法に関し、その場合、m R N A 発現は、下記：

30

【 0 0 2 2 】

【数 2】

$$MS = a0 + b * \left[\sum_{i=1}^M ai * Gi \right] \quad \text{等式 1}$$

等式 1

（式中、

M は 1 4 であり、G i は 1 4 個の遺伝子の各遺伝子（i）の標準化された発現レベルであり、a 0 は 0 . 0 2 2 であり、b は - 0 . 2 5 1 であり、a i は 1 4 個の遺伝子サインに含まれる遺伝子の各々について表 2 に示されている値に対応する）

40

によって転移スコア（MS）に計算される。標準化された発現レベルは、（C t）において測定される発現レベルからトレーニングセットの遺伝子の平均発現を差し引くことによって得られ、次に、その遺伝子の遺伝子発現の標準偏差によって割られる。トレーニングセットの各遺伝子についての遺伝子発現の平均および標準偏差は表 4 に示される。等式 1 は、実施例 1、2 および 3 において用いられた。

【 0 0 2 3 】

本発明の更なる局面では、MS 式は、下記の定義：

【 0 0 2 4 】

【数 3】

$$MS = a_0 + b * \left[\sum_{i=1}^M a_i * G_i \right] \quad \text{等式 2}$$

等式 2

(式中、

Mは14であり、 G_i は14個の遺伝子の各遺伝子(i)の (C_t)において測定される発現レベルであり、 a_0 が0.8657であり、 b が-0.04778であり、 a_i が全ての遺伝子について1である)

を推定することができる。等式 2 は、実施例 4 および 5 において用いられた。

10

【0025】

本発明の更なる局面では、MS式は、 $a = 0$ 、 $b = -1$ 、 $a_i = 1$ であり得る。

【0026】

本発明の別の局面では、上述された14個の遺伝子の発現プロファイリングを用いて、乳癌患者の腫瘍転移と関連したリスクを測定する方法に関し、その場合、各遺伝子(i)の発現レベル G_i は、下記：

$$(C_t) = (C_{t_{GOI}} - C_{t_{EC}})_{testRNA} - (C_{t_{GOI}} - C_{t_{EC}})_{refRNA} \quad \text{等式 4}$$

(式中、

C_t は指数関数的な標的増幅のPCR閾値サイクルであり、GOIは対象とする遺伝子であり、ECは内在性対照であり、 $testRNA$ は患者の試料RNAであり、 $refRNA$ は参照RNAである)

20

によって遺伝子発現値 G_i に計算される。

【0027】

本発明の別の局面では、上述した14個の遺伝子の発現プロファイリングを用いて、乳癌患者の腫瘍転移と関連したリスクを測定する方法に関し、その場合、各遺伝子(i)の発現レベル G_i は、MSスコアの1つの値に組み合わされ、ここで、関連のあるMS閾値またはカットポイントよりも高いMSスコアを有する患者は、腫瘍転移に関して高いリスクにある。

【0028】

30

本発明の一局面では、遺伝子CENPA、PKMYT1、MELK、MYBL2、BUB1、RACGAP1、TK1、UBE2S、DC13、RFC4、PRR11、DIAPH3、ORC6LおよびCCNB1の発現レベルを検出するための試薬、酵素；およびバッファーを含むキットに関する。

【0029】

本発明の別の局面では、遺伝子CENPA、PKMYT1、MELK、MYBL2、BUB1、RACGAP1、TK1、UBE2S、DC13、RFC4、PRR11、DIAPH3、ORC6LおよびCCNB1にハイブリダイズするポリヌクレオチドを含むマイクロアレイに関する。

本発明はまた、以下の項目を提供する。

40

(項目1)

乳癌患者の腫瘍転移と関連したリスクを測定する方法であって、

(a) 上記乳癌患者のエストロゲン受容体陽性の腫瘍細胞における遺伝子CENPA、PKMYT1、MELK、MYBL2、BUB1、RACGAP1、TK1、UBE2S、DC13、RFC4、PRR11、DIAPH3、ORC6LおよびCCNB1の発現レベルを測定し、それによって、上記遺伝子の発現レベルに基づく転移スコア(MS)を得て；

(b) 上記転移スコアと所定の転移スコアのカットオフ閾値(MS閾値)とを比較することによって、上記乳癌患者の腫瘍転移のリスクを測定することを含む方法。

50

(項目2)

工程(b)において、上記乳癌患者のMSが所定のMS閾値よりも高い場合に、上記乳癌患者は、腫瘍転移のリスクが増加していると決定される、項目1に記載の方法。

(項目3)

工程(b)において、そのMSが所定のMS閾値よりも低い場合に、上記乳癌患者は、腫瘍転移のリスクが減少していると決定される、項目1に記載の方法。

(項目4)

工程(b)において、2以上のMS閾値が、上記乳癌患者について腫瘍転移のリスクの決定に用いられる、項目1に記載の方法。

(項目5)

上記乳癌患者が、リンパ節において検出可能な腫瘍細胞を有さない、項目1に記載の方法。

(項目6)

上記乳癌患者が、腫瘍転移を発症するリスクが増加している、項目1に記載方法。

(項目7)

上記エストロゲン受容体陽性の腫瘍細胞が、ホルマリン固定されたパラフィン包埋された切片に由来する、項目1に記載の方法。

(項目8)

上記エストロゲン受容体陽性の腫瘍細胞が、腫瘍生検由来である、項目1に記載の方法。

(項目9)

上記エストロゲン受容体陽性の腫瘍細胞が、凍結腫瘍組織由来である、項目1に記載の方法。

(項目10)

上記遺伝子のmRNAが、エストロゲン受容体陽性の腫瘍細胞から得られ、cDNAに逆転写され、上記cDNAのポリメラーゼ連鎖反応増幅によって検出される、項目1に記載の方法。

(項目11)

上記mRNAが、逆転写およびPCR増幅前に濃縮される、項目10に記載の方法。

(項目12)

上記遺伝子の各々のmRNAが、逆転写され、表3の配列番号1~30に示される対応する遺伝子と関連した2つのプライマーによって増幅される、項目1に記載の方法。

(項目13)

上記遺伝子の発現レベルが、あるハウスキーピング遺伝子または2以上のハウスキーピング遺伝子の平均の発現レベルに対して標準化される、項目1に記載の方法。

(項目14)

上記ハウスキーピング遺伝子が、NUP214、PP1GおよびSLU7で構成される群から選択される、項目13に記載の方法。

(項目15)

上記ハウスキーピング遺伝子のmRNAが、逆転写され、表3に列挙される、配列番号29~34の上記ハウスキーピング遺伝子と関連した2つのプライマーによって増幅される、項目13に記載の方法。

(項目16)

上記発現レベルが、マイクロアレイによって検出される、項目1に記載の方法。

(項目17)

上記転移スコア(MS)が、下記：

【数8】

$$MS = a0 + \sum_{i=1}^M ai * Gi$$

10

20

30

40

50

(式中、

Mは14であり、 G_i は14個の上記遺伝子の各遺伝子(i)の発現レベルであり、 a_0 は0.022であり、 a_i は上記遺伝子の各々について表2にある値に対応する)によって計算される、項目1に記載の方法。

(項目18)

上記各々の遺伝子(i)の発現レベルが、下記：

$$(Ct) = (Ct_{GOI} - Ct_{EC})_{testRNA} - (Ct_{GOI} - Ct_{EC})_{refRNA}$$

(式中、

Ct は指数関数的な標的増幅のPCR閾値サイクルであり、GOIは対象とする遺伝子であり、ECは内在性対照であり、 $testRNA$ は患者の試料RNAであり、 $refRNA$ は参照RNAである)

によって遺伝子発現値 G_i に算定される、項目17に記載の方法。

(項目19)

a_0 が0であり、 a_i が1である、項目17に記載の方法。

(項目20)

上記転移スコア(MS)が、下記：

【数9】

$$MS = a_0 + b * \left[\sum_{i=1}^M a_i * G_i \right]$$

(式中、

Mが14であり、 G_i が14個の上記遺伝子の各遺伝子(i)の基準化された発現レベルであり、 a_0 が0.022であり、bが-0.251であり、 a_i が上記遺伝子の各々について表2に示される値に対応する)

によって計算される、項目1に記載の方法。

(項目21)

上記転移スコア(MS)が、下記：

【数10】

$$MS = a_0 + b * \left[\sum_{i=1}^M a_i * G_i \right]$$

(式中、

Mが14であり、 G_i が14個の上記遺伝子の各遺伝子(i)の発現レベルであり、 a_0 が0.8657であり、bが-0.04778であり、 a_i が上記遺伝子の各々について1である)

によって計算される、項目1に記載の方法。

(項目22)

上記転移スコア(MS)が、下記：

【数11】

$$MS(新規) = -1/14 * \left[\sum_{i=1}^{14} G_i \right]$$

(式中、 G_i が14個の上記遺伝子の各遺伝子(i)の発現レベルである)

によって計算される、項目1に記載の方法。

(項目 23)

上記転移スコア (MS) が、14 個の上記遺伝子の各遺伝子の発現レベルを一緒に加えることによって計算される、項目 1 に記載の方法。

(項目 24)

遺伝子 CENPA、PKMYT1、MELK、MYBL2、BUB1、RACGAP1、TK1、UBE2S、DC13、RFC4、PRR11、DIAPH3、ORC6L および CCNB1 の発現レベルを検出するための試薬 (複数)、および酵素；並びにバッファーを含むキット。

(項目 25)

1 以上のハウスキーピング遺伝子の発現レベルを検出するための試薬 (複数) またはその組み合わせをさらに含む、項目 24 に記載のキット。

10

(項目 26)

上記ハウスキーピング遺伝子が、NUP214、PPIG および SLU7 で構成される群から選択される、項目 25 に記載のキット。

(項目 27)

遺伝子 CENPA、PKMYT1、MELK、MYBL2、BUB1、RACGAP1、TK1、UBE2S、DC13、RFC4、PRR11、DIAPH3、ORC6L および CCNB1 にハイブリダイズするポリヌクレオチドを含むマイクロアレイ。

(項目 28)

ハウスキーピング遺伝子にハイブリダイズするポリヌクレオチドをさらに含む、項目 27 に記載のマイクロアレイ。

20

(項目 29)

上記ハウスキーピング遺伝子が、NUP214、PPIG および SLU7 からなる群から選択される、項目 28 に記載のマイクロアレイ。

配列リスト

添付された配列リストは、全体として参照により本明細書中に援用される。配列リストは、表 3 に示されるオリゴヌクレオチド (配列番号 1 ~ 34) を与える。これらのオリゴヌクレオチドは、表 3 に列挙された遺伝子の RT-PCR 増幅における例示的なプライマーである。

30

【図面の簡単な説明】

【0030】

【図 1 a】a) 遠隔転移までの時間、b) CPMC のトレーニングセットに関する全生存についての Kaplan-Meier 曲線を示し、そこでは、高リスクおよび低リスク群は、カットポイントとしてゼロを用いた MS (CV) によって特定された。

【図 1 b】a) 遠隔転移までの時間、b) CPMC のトレーニングセットに関する全生存についての Kaplan-Meier 曲線を示し、そこでは、高リスクおよび低リスク群は、カットポイントとしてゼロを用いた MS (CV) によって特定された。

【図 2】CPMC のトレーニングセットにおける 5 年以内の MS (CV) による遠隔転移を予測するための ROC 曲線である。AUC = 0.76 (0.65 - 0.87)。

40

【図 3 a】ガイ病院からの 280 人の未処置患者の遺伝子サインおよびアジュバント！によって特定されたリスク群による Kaplan-Meier 曲線を示す。具体的には、図 3 a) および b) は、14 個の遺伝子サインを用いた結果を記載し、図 3 c) および d) は、アジュバント！因子を用いた結果を示す。a) MS リスク群による遠隔転移 (DMFS) までの時間、b) MS リスク群による全生存、c) アジュバント！リスク群による遠隔転移 (DMFS) までの時間、d) アジュバント！リスク群による全生存。

【図 3 b】ガイ病院からの 280 人の未処置患者の遺伝子サインおよびアジュバント！によって特定されたリスク群による Kaplan-Meier 曲線を示す。具体的には、図 3 a) および b) は、14 個の遺伝子サインを用いた結果を記載し、図 3 c) および d) は、アジュバント！因子を用いた結果を示す。a) MS リスク群による遠隔転移 (DMFS) ま

50

での時間、b) MS リスク群による全生存、c) アジュバント！リスク群による遠隔転移 (DMFS) までの時間、d) アジュバント！リスク群による全生存。

【図3c】ガイ病院からの280人の未処置患者の遺伝子サインおよびアジュバント！によって特定されたリスク群による Kaplan-Meier 曲線を示す。具体的には、図3a) および b) は、14個の遺伝子サインを用いた結果を記載し、図3c) および d) は、アジュバント！因子を用いた結果を示す。a) MS リスク群による遠隔転移 (DMFS) までの時間、b) MS リスク群による全生存、c) アジュバント！リスク群による遠隔転移 (DMFS) までの時間、d) アジュバント！リスク群による全生存。

【図3d】ガイ病院からの280人の未処置患者の遺伝子サインおよびアジュバント！によって特定されたリスク群による Kaplan-Meier 曲線を示す。具体的には、図3a) および b) は、14個の遺伝子サインを用いた結果を記載し、図3c) および d) は、アジュバント！因子を用いた結果を示す。a) MS リスク群による遠隔転移 (DMFS) までの時間、b) MS リスク群による全生存、c) アジュバント！リスク群による遠隔転移 (DMFS) までの時間、d) アジュバント！リスク群による全生存。

【図4a】遺伝子サインおよびオンラインプログラムアジュバント！のレシーバー操作特徴 (Receiver operating characteristic: ROC) 曲線を示す。a) 遺伝子サインに関する5年以内の遠隔転移についての ROC 曲線、b) 遺伝子サインに関する10年以内の遠隔転移についての ROC 曲線、c) 遺伝子サインに関する10年以内の死亡に関する ROC 曲線、d) アジュバント！に関する5年以内の転移についての ROC 曲線、e) アジュバント！に関する10年以内の転移についての ROC 曲線、f) ガイ病院からの未処置患者についてのアジュバント！に関する10年以内の死亡についての ROC 曲線。

【図4b】遺伝子サインおよびオンラインプログラムアジュバント！のレシーバー操作特徴 (Receiver operating characteristic: ROC) 曲線を示す。a) 遺伝子サインに関する5年以内の遠隔転移についての ROC 曲線、b) 遺伝子サインに関する10年以内の遠隔転移についての ROC 曲線、c) 遺伝子サインに関する10年以内の死亡に関する ROC 曲線、d) アジュバント！に関する5年以内の転移についての ROC 曲線、e) アジュバント！に関する10年以内の転移についての ROC 曲線、f) ガイ病院からの未処置患者についてのアジュバント！に関する10年以内の死亡についての ROC 曲線。

【図4c】遺伝子サインおよびオンラインプログラムアジュバント！のレシーバー操作特徴 (Receiver operating characteristic: ROC) 曲線を示す。a) 遺伝子サインに関する5年以内の遠隔転移についての ROC 曲線、b) 遺伝子サインに関する10年以内の遠隔転移についての ROC 曲線、c) 遺伝子サインに関する10年以内の死亡に関する ROC 曲線、d) アジュバント！に関する5年以内の転移についての ROC 曲線、e) アジュバント！に関する10年以内の転移についての ROC 曲線、f) ガイ病院からの未処置患者についてのアジュバント！に関する10年以内の死亡についての ROC 曲線。

【図4d】遺伝子サインおよびオンラインプログラムアジュバント！のレシーバー操作特徴 (Receiver operating characteristic: ROC) 曲線を示す。a) 遺伝子サインに関する5年以内の遠隔転移についての ROC 曲線、b) 遺伝子サインに関する10年以内の遠隔転移についての ROC 曲線、c) 遺伝子サインに関する10年以内の死亡に関する ROC 曲線、d) アジュバント！に関する5年以内の転移についての ROC 曲線、e) アジュバント！に関する10年以内の転移についての ROC 曲線、f) ガイ病院からの未処置患者についてのアジュバント！に関する10年以内の死亡についての ROC 曲線。

【図4e】遺伝子サインおよびオンラインプログラムアジュバント！のレシーバー操作特徴 (Receiver operating characteristic: ROC) 曲線を示す。a) 遺伝子サインに関する5年以内の遠隔転移についての ROC 曲線、b) 遺伝子サインに関する10年以内の遠隔転移についての ROC 曲線、c) 遺伝子サインに

10

20

30

40

50

関する10年以内の死亡に関するROC曲線、d) アジュバント！に関する5年以内の転移についてのROC曲線、e) アジュバント！に関する10年以内の転移についてのROC曲線、f) ガイ病院からの未処理患者についてのアジュバント！に関する10年以内の死亡についてのROC曲線。

【図4f】遺伝子サインおよびオンラインプログラムアジュバント！のレシーバー操作特徴 (Receiver operating characteristic: ROC) 曲線を示す。a) 遺伝子サインに関する5年以内の遠隔転移についてのROC曲線、b) 遺伝子サインに関する10年以内の遠隔転移についてのROC曲線、c) 遺伝子サインに関する10年以内の死亡に関するROC曲線、d) アジュバント！に関する5年以内の転移についてのROC曲線、e) アジュバント！に関する10年以内の転移についてのROC曲線、f) ガイ病院からの未処理患者についてのアジュバント！に関する10年以内の死亡についてのROC曲線。

10

【図5】280人のガイの未処理患者からの5年および10年以内の遠隔転移の可能性に対する転移スコア (MS) の比較である。

【図6】14個の遺伝子サインからの10年以内の遠隔転移の可能性に対する、ガイ病院からの未処置患者についてのアジュバント！からの10年の再発率の比較である。

【図7A】ミュエンスター大学の患者における遠隔転移のない生存についての Kaplan-Meier 曲線を示す。

【図7B】ミュエンスター大学の患者における遠隔転移のない生存についての Kaplan-Meier 曲線を示す。

20

【図8】ガイ病院からの205人の処置患者に関する3MS群 (高、中、低) の遠隔転移なしの生存の Kaplan-Meier 曲線を示す。

【図9】ガイ病院からの205人の処置患者によって測定された2つのリスク群 (高および低) における遠隔転移なしの生存の Kaplan-Meier 曲線を示す。

【図10】ガイの処置試料に関する5年以内の遠隔転移を予測するためのMSのROC曲線を示す。AUC = 0.7 (0.57 - 0.87)。

【図11A】ガイ処置試料におけるMSによる高対低のリスク群のハザード率に依存した時間を示す。

【図11B】ガイ処置試料におけるMSによる高対低のリスク群のハザード率に依存した時間を示す。

30

【図12】234人の日本人試料における3つのMS群 (高、中、および低) に関する遠隔転移なしの生存 (DMFS) の Kaplan-Meier 曲線を示す。

【図13】234人の日本人試料における2つのリスク群 (高MSは高リスクを有し、中および低MSは低リスクを有する) についての遠隔転移なしの生存 (DMFS) の Kaplan-Meier 曲線を示す。

【図14】日本人患者に関する5年以内の遠隔転移を予測するためのMSのROC曲線を示す。AUC = 0.73 (0.63 - 0.84)。

【図15A】時間の関数として、MS群に対する年換算でのハザード率、および高対低リスク群のハザード率を示す。

【図15B】時間の関数として、MS群に対する年換算でのハザード率、および高対低リスク群のハザード率を示す。

40

【発明を実施するための形態】

【0031】

本発明は、乳癌転移を予測するために用いることができる多重遺伝子サイン、本明細書に開示されている遺伝子を検出するための方法および試薬、並びにこのような試薬を利用するアッセイまたはキットを提供する。本明細書に開示されている乳癌転移に関連した遺伝子は、乳癌患者のER陽性腫瘍の遠隔転移の可能性を診断、スクリーニング、および評価するのに有用である。

【0032】

表2に開示されている分子サインの14個の遺伝子の発現プロファイリングは、遠隔転

50

移の予後を容易に推測することができる。表 2 に示されている情報は、参照配列 (R e f S e q) を含み、国立衛生研究所 / 国立医学図書館の全米バイオテクノロジー情報センター (N C B I) から得られ、記載されている各遺伝子の 1 つの変異転写配列を特定する。この変異体の配列に基づいて、試薬は、14 個の遺伝子サインの各遺伝子の全ての変異体を特定するために設計されてもよい。表 3 は、各遺伝子の全ての変異体が増幅されるように、14 個の遺伝子サインの各遺伝子を検出するために用いることができる例示的なプライマーセットを提供する。このようにして、本発明は、本明細書に開示されている全ての遺伝子の全ての知られている転移変異体の発現プロファイリングを提供する。

【 0 0 3 3 】

また、表 2 には、各 R e f S e q のヌクレオチド配列を公開する参考文献が示されている。これらの参考文献の全ては、全体として参照により本明細書に援用される。また、表 2 には、各遺伝子の記述がある。参考文献および記述はともに N C B I によって提供された。

10

【 0 0 3 4 】

C E N P A 遺伝子は、参照配列 N M _ 0 0 1 8 0 9 によって特定され、B l a c k , B . E . , F o l t z , D . R . ら , 2 0 0 4 , N a t u r e 4 3 0 (6 9 9 9) : 5 7 8 - 5 8 2 に開示されている。この参照配列および参考文献は、全体として参照により本明細書に援用される。

【 0 0 3 5 】

P K M Y T 1 遺伝子は、参照配列 N M _ 0 0 4 2 0 3 によって特定され、B r y a n , B . A . , D y s o n , O . F . ら , 2 0 0 6 , J . G e n . V i r o l . 8 7 (P T 3) , 5 1 9 - 5 2 9 に開示されている。この参照配列および参考文献は、全体として参照により本明細書に援用される。

20

【 0 0 3 6 】

M E L K 遺伝子は、参照配列 N M _ 0 1 4 7 9 1 によって特定され、B e u l l e n s , M . , V a n c a u w e n b e r g h , S . ら , 2 0 0 5 , J . B i o l . C h e m . 2 8 0 (4 8) , 4 0 0 0 3 - 4 0 0 1 1 に開示されている。この参照配列および参考文献は、全体として参照により本明細書に援用される。

【 0 0 3 7 】

M Y B L 2 遺伝子は、参照配列 N M _ 0 0 2 4 6 6 によって特定され、B r y a n , B . A . , D y s o n , O . F . ら , 2 0 0 6 , J . G e n . V i r o l . 8 7 (P T 3) , 5 1 9 - 5 2 9 に開示されている。この参照配列および参考文献は、全体として参照により本明細書に援用される。

30

【 0 0 3 8 】

B U B 1 遺伝子は、参照配列 N M _ 0 0 4 3 3 6 によって特定され、M o r r o w , C . J . , T i g h e , A . ら , 2 0 0 5 , J . C e l l . S c i . 1 1 8 (P T 1 6) , 3 6 3 9 - 3 6 5 2 に開示されている。この参照配列および参考文献は、全体として参照により本明細書に援用される。

【 0 0 3 9 】

R A C G A P 1 遺伝子は、参照配列 N M _ 0 1 3 2 7 7 によって特定され、N i i y a , F . , X i e , X . ら , 2 0 0 5 , J . B i o l . C h e m . 2 8 0 (4 3) , 3 6 5 0 2 - 3 6 5 0 9 に開示されている。この参照配列および参考文献は、全体として参照により本明細書に援用される。

40

【 0 0 4 0 】

T K 1 遺伝子は、参照配列 N M _ 0 0 3 2 5 8 によって特定され、K a r b o w n i k , M . , B r z e z i a n s k a , E . ら , 2 0 0 5 , C a n c e r L e t t . 2 2 5 (2) , 2 6 7 - 2 7 3 に開示されている。この参照配列および参考文献は、全体として参照により本明細書に援用される。

【 0 0 4 1 】

U B E 2 S 遺伝子は、参照配列 N M _ 0 1 4 5 0 1 によって特定され、L i u , Z . ,

50

D i a z , L . A . ら , 1 9 9 2 , J . B i o l . C h e m . 2 6 7 (2 2) , 1 5 8 2 9 - 1 5 8 3 5 に開示されている。この参照配列および参考文献は、全体として参照により本明細書に援用される。

【 0 0 4 2 】

D C 1 3 遺伝子は、参照配列 A F 2 0 1 9 3 5 によって特定され、G u , Y . , P e n g , Y . ら , D i r e c t S u b m i s s i o n , S u b m i t t e d N o v . 5 , 1 9 9 9 、中国 国家ヒトゲノムセンター、上海、351 G u o S h o u j i n g R o a d , Z h a n g j i a n g H i - T e c h P a r k , P u d o n g , S h a n g h a i (上海市浦東新區張江高科技園區郭守敬路351號) 201203、中華人民共和國に開示されている。この参照配列および参考文献は、全体として参照により本明細書に援用される。

10

【 0 0 4 3 】

R F C 4 遺伝子は、参照配列 N M _ 0 0 2 9 1 6 によって特定され、G u p t e , R . S . , W e n g , Y . ら , 2 0 0 5 , C e l l C y c l e 4 (2) , 3 2 3 - 3 2 9 に開示されている。この参照配列および参考文献は、全体として参照により本明細書に援用される。

【 0 0 4 4 】

P R R 1 1 遺伝子は、参照配列 N M _ 0 1 8 3 0 4 によって特定され、W e i n m a n n , A . S . , Y a n , P . S . ら , 2 0 0 2 , G e n e s D e v . 1 6 (2) , 2 3 5 - 2 4 4 に開示されている。この参照配列および参考文献は、全体として参照により本明細書に援用される。

20

【 0 0 4 5 】

D I A P H 3 遺伝子は、参照配列 N M _ 0 3 0 9 3 2 によって特定され、K a t o h , M . および K a t o h , M . , 2 0 0 4 , I n t . J . M o l . M e d . 1 3 (3) , 4 7 3 - 4 7 8 に開示されている。この参照配列および参考文献は、全体として参照により本明細書に援用される。

【 0 0 4 6 】

O R C 6 L 遺伝子は、参照配列 N M _ 0 1 4 3 2 1 によって特定され、S i b a n i , S . , P r i c e , G . B . ら , 2 0 0 5 , B i o c h e m i s t r y 4 4 (2 1) , 7 8 8 5 - 7 8 9 6 に開示されている。この参照配列および参考文献は、全体として参照により本明細書に援用される。

30

【 0 0 4 7 】

C C N B 1 遺伝子は、参照配列 N M _ 0 3 1 9 6 6 によって特定され、Z h a o , M . , K i m , Y . T . ら , 2 0 0 6 , E x p O n c o l 2 8 (1) , 4 4 - 4 8 に開示されている。この参照配列および参考文献は、全体として参照により本明細書に援用される。

【 0 0 4 8 】

P P I G 遺伝子は、参照配列 N M _ 0 0 4 7 9 2 によって特定され、L i n , C . L . , L e u , S . ら , 2 0 0 4 , B i o c h e m . B i o p h y s . R e s . C o m m u n . 3 2 1 (3) , 6 3 8 - 6 4 7 に開示されている。この参照配列および参考文献は、全体として参照により本明細書に援用される。

40

【 0 0 4 9 】

N U P 2 1 4 遺伝子は、参照配列 N M _ 0 0 5 0 8 5 によって特定され、G r a u x , C , C o o l s , J . ら , 2 0 0 4 , N a t . G e n e t . 3 6 (1 0) , 1 0 8 4 - 1 0 8 9 に開示されている。この参照配列および参考文献は、全体として参照により本明細書に援用される。

【 0 0 5 0 】

S L U 7 遺伝子は、参照配列 N M _ 0 0 6 4 2 5 によって特定され、S h o m r o n , N . , A l b e r s t e i n , M . ら , 2 0 0 5 , J . C e l l . S c i . 1 1 8 (P T 6) , 1 1 5 1 - 1 1 5 9 に開示されている。この参照配列および参考文献は、全体とし

50

て参照により本明細書に援用される。

【 0 0 5 1 】

また、表 2 には、各遺伝子 i に対する転移スコアを決定するために必要とされる一定の a_i についての値が示され、それは、その遺伝子について得られた発現プロファイリングの結果に基づいている。転移スコアの誘導およびその使用、遺伝子発現プロファイリングの方法、そこから得られたデータの使用が下記に記載されている。

【 0 0 5 2 】

このようにして、本発明は、乳癌転移についてともに予後となる 14 個の別々の遺伝子、試験試料にあるこれらの遺伝子の発現レベルを決定する方法、遠隔転移を発祥する個体の可能性を決定する方法、処置戦略を選択するための開示されている遺伝子の使用方法を提供する。

10

【 0 0 5 3 】

本発明は、従来、当該技術分野において知られていなかった 14 個の遺伝子サインの独自の組み合わせを提供する。したがって、本発明は、本明細書に開示されている遺伝子に基づく新規な方法を提供し、また、乳癌転移に関連した（例えば、乳癌転移の予後に関する）方法において、知られてはいるが、従来無関係とされた遺伝子を用いる方法を提供する。

【 0 0 5 4 】

当業者は、核酸分子が二重鎖分子であってもよく、1 本鎖の特定配列への言及が、同様に、相補鎖の対応する部位への言及であることを容易に認識する。ヌクレオチド配列の定義において、核酸分子の 1 本鎖上の特定部位でのアデニン、チミジン（ウリジン）、シトシン、またはグアニンへの言及も、核酸分子の相補鎖上の対応する部位でのチミジン（ウリジン）、アデニン、グアニン、またはシトシン（それぞれ）を定義する。つまり、特定のヌクレオチド配列に言及するために、いずれかの鎖に言及がなされてもよい。プローブおよびプライマーは、いずれかの鎖にハイブリダイズするように設計されてもよく、本明細書に開示されている遺伝子発現プロファイリング法は、一般に、いずれかの鎖を標的としてもよい。

20

腫瘍組織供給源および RNA 抽出

本発明では、標的ポリヌクレオチド分子は、乳癌に苦しんでいる個体から採取された試料から抽出される。試料は、任意の臨床的に許容される方法で回収されてもよいが、遺伝子特異的なポリヌクレオチド（即ち、転写 RNA、またはメッセージ）は保存されるように回収されなければならない。次に、このようにして試料から得られた mRNA または核酸は、さらに分析されてもよい。例えば、遺伝子（例えば、表 2 に示されている遺伝子）に特異的なオリゴヌクレオチド対は、試料に含まれる特異的な mRNA（単数または複数）を増幅させるために用いられてもよい。次に、各メッセージの量が測定されるか、またはプロファイルすることができ、疾患予後と関連付けることができる。あるいは、mRNA またはそれから誘導される核酸（即ち、cDNA、増幅された DNA または濃縮された RNA）は、標準または対照のポリヌクレオチド分子から区別されるように標識されてもよく、両者は、上述されるマーカーのいくつかもしくは全部、またはマーカーセットもしくはサブセットを含むマイクロアレイに、同時にまたは個別にハイブリダイズされてもよい。あるいは、mRNA またはそこから誘導される核酸は、標準または対照のポリヌクレオチド分子と同じ標識で標識されてもよく、ここで、特定のプローブで各々のハイブリダイゼーションの強度が比較される。

30

40

【 0 0 5 5 】

試料は、いずれかの臨床的に関連する組織試料を含んでもよく、例えば、ホルマリン固定されたパラフィン包埋試料、凍結試料、腫瘍検体もしくは細針吸引物、または、ER 陽性の腫瘍細胞を含む体液試料、例えば血液、血漿、血清、リンパ液、腹水もしくは囊胞液、尿、または乳頭浸出液である。

【 0 0 5 6 】

トータルおよびポリ (A) + RNA を調製する方法は周知であり、一般に、Sambrook

50

ookら, MOLECULAR CLONING - A LABORATORY MANUAL (第2版), Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York (1989))、Ausubelら, Current Protocols in Molecular Biology vol. 2, Current Protocols Publishing, New York (1994))に記載されている。RNAは、細胞の溶解、そこに含まれているタンパク質の変性を含む、当該技術分野において周知であるいずれかの手法によって、ER陽性の腫瘍細胞から単離されてよい。

【0057】

組織試料からRNAを調製するための一例として、RNAは、当該技術分野において周知の技術を用いて、ホルマリン固定されたパラフィン包埋 (FFPE) 組織から単離されてもよい。この目的のための市販キットは、例えば、Zymo Research、Ambion、Qiagen、またはStratageneから得ることができる。FFPE組織からトータルRNAを単離する例示的な方法は、Pinpoint Slide RNA Isolation System (Zymo Research, Orange, Calif.) に従えば、下記のとおりである。要約すると、Zymoキットから得られる溶液は、対象とする選択されたFFPE組織領域全体に塗布され、乾燥される。次に、包埋された組織は、スライドから剥がされ、プロテイナーゼKを含む遠心管に入れられる。組織を数時間インキュベート後、細胞溶解物を遠心分離し、上清を別の管に移す。RNAは、グアニジニウムチオシネート / メルカプトエタノール溶液にエタノールを加え、混合した溶液を用いて、溶解物から抽出される。試料をスピncラムに適用し、1分間回転される。エタノールおよびTris / EDTAを含むバッファーを用いてカラムを洗浄する。DNaseをカラムに添加し、インキュベートする。加熱したRNAaseを含まない水をカラムに添加し、遠心分離することによって、カラムからRNAを抽出する。純粋なトータルRNAは溶出物に存在する。

【0058】

更なる工程は、汚染しているDNAを除去するために用いられてもよく、例えば、上述されるように、スピncラムにDNaseが添加される。細胞溶解物は、非イオン性界面活性剤を用い、その後、核、したがって、細胞性DNAのバルクを除去するために、マイクロ遠心分離によって達成されてもよい。一態様では、RNAは、グアニジニウムチオシネートの存在下で細胞溶解後、CsCl遠心分離によりDNAからRNAを分離することによって対象とする種々のタイプの細胞から抽出される (Chirgwinら, Biochemistry 18:5294-5299 (1979))。ポリ(A)+RNAは、オリゴ-dTセルロース (Sambrookら, MOLECULAR CLONING - A LABORATORY MANUAL (第2版), Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York (1989) を参照されたい)。あるいは、DNAからのRNAの分離は、有機抽出、例えば、熱フェノール / クロロホルム / イソアミルアルコールを用いて達成することができる。

【0059】

必要に応じて、RNAase阻害剤を溶解バッファーに添加してもよい。同様に、ある種の細胞型については、タンパク質変性 / 消化工程をプロトコールに加えることが望ましい場合がある。

【0060】

多くの応用に関して、転移RNA (tRNA) およびリボゾームRNA (rRNA) などの、細胞から抽出された他の細胞性RNAに関して、選択的にmRNAを濃縮することが望ましい。大部分のmRNAは、それらの3'末端でポリ(A)テールを含む。これは、アフィニティークロマトグラフィーによって；例えば、固体支持体、例えばセルロースまたはSephadex (商標) に結合されたオリゴ(dT) またはポリ(U) を用いて、濃縮が可能になる (Ausubelら, CURRENT PROTOCOLS IN

10

20

30

40

50

MOLECULAR BIOLOGY, vol. 2, Current Protocols Publishing, New York (1994)を参照されたい)。このようにして結合後、ポリ(A)+mRNAは、2mM EDTA/0.1%SDSを用いてアフィニティークラムから溶出される。

【0061】

RNA試料は、複数の異なるmRNA分子を含んでもよく、各mRNA分子は、種々のヌクレオチド配列を有する。特定の態様では、RNA試料のmRNA分子は、本明細書に開示されている14個の遺伝子の各々に対応するmRNAを含む。更なる特定の態様では、細胞からのトータルRNAまたはmRNAは、本発明の方法に用いられる。RNAの供給源は、任意のER陽性の腫瘍細胞由来の細胞であってもよい。特定の態様では、本発明の方法は、 1×10^6 細胞以下からトータルmRNAまたはトータルRNAを含む試料とともに用いられる。

10

遺伝子発現を測定するための試薬

本発明は、遺伝子発現プロファイリング、乳癌転移の予後の測定に用いることができる核酸分子を提供する。本明細書に記載されている14個の遺伝子サインの遺伝子発現プロファイリングにおいて、プライマーとして用いることができる例示的な核酸分子は、表3に示されている。

【0062】

表3に示されるように：

遺伝子BUB1は逆転写され、上流プライマー(5')としての配列番号1、下流プライマー(3')としての配列番号2を用いて増幅される。

20

【0063】

遺伝子CCNB1は逆転写され、上流プライマー(5')としての配列番号3、下流プライマー(3')としての配列番号4を用いて増幅される。

【0064】

遺伝子CENPAは逆転写され、上流プライマー(5')としての配列番号5、下流プライマー(3')としての配列番号6を用いて増幅される。

【0065】

遺伝子DC13は逆転写され、上流プライマー(5')としての配列番号7、下流プライマー(3')としての配列番号8を用いて増幅される。

30

【0066】

遺伝子DIAPH3は逆転写され、上流プライマー(5')としての配列番号9、下流プライマー(3')としての配列番号10を用いて増幅される。

【0067】

遺伝子MELKは逆転写され、上流プライマー(5')としての配列番号11、下流プライマー(3')としての配列番号12を用いて増幅される。

【0068】

遺伝子MYBL2は逆転写され、上流プライマー(5')としての配列番号13、下流プライマー(3')としての配列番号14を用いて増幅される。

【0069】

遺伝子NUP214は逆転写され、上流プライマー(5')としての配列番号29、下流プライマー(3')としての配列番号30を用いて増幅される。

40

【0070】

遺伝子ORC6Lは逆転写され、上流プライマー(5')としての配列番号15、下流プライマー(3')としての配列番号16を用いて増幅される。

【0071】

遺伝子PKMYT1は逆転写され、上流プライマー(5')としての配列番号17、下流プライマー(3')としての配列番号18を用いて増幅される。

【0072】

遺伝子PPIGは逆転写され、上流プライマー(5')としての配列番号31、下流プ

50

ライマー（３'）としての配列番号３２を用いて増幅される。

【００７３】

遺伝子 P R R １１は逆転写され、上流プライマー（５'）としての配列番号１９、下流プライマー（３'）としての配列番号２０を用いて増幅される。

【００７４】

遺伝子 R A C G A P １は逆転写され、上流プライマー（５'）としての配列番号２１、下流プライマー（３'）としての配列番号２２を用いて増幅される。

【００７５】

遺伝子 R F C ４は逆転写され、上流プライマー（５'）としての配列番号２３、下流プライマー（３'）としての配列番号２４を用いて増幅される。

【００７６】

遺伝子 S L U ７は逆転写され、上流プライマー（５'）としての配列番号３３、下流プライマー（３'）としての配列番号３４を用いて増幅される。

【００７７】

遺伝子 T K １は逆転写され、上流プライマー（５'）としての配列番号２５、下流プライマー（３'）としての配列番号２６を用いて増幅される。

【００７８】

遺伝子 U B E ２ S は逆転写され、上流プライマー（５'）としての配列番号２７、下流プライマー（３'）としての配列番号２８を用いて増幅される。

【００７９】

表２に示されている各遺伝子の完全なヌクレオチド配列に基づいて、当業者は、本明細書に記載されている１４個の遺伝子サインの増幅および／または検出に用いることができる更なるプライマーおよび／またはプローブを容易に設計し、合成することができる。

【００８０】

本発明の特定の局面では、表３に開示されている配列は、遺伝子発現プロファイリング試薬として用いることができる。本明細書中で使用するとき、「遺伝子発現プロファイリング試薬」とは、配列が m R N A または c D N A であることを問わず、本明細書に記載されている遺伝子の特定の標的遺伝子のヌクレオチド配列を増幅および／または検出する方法に特に有用である試薬を指す。例えば、プロファイリング試薬は、好ましくは、異なる代替の遺伝子ヌクレオチド配列間を区別することができ、それによって、測定されるべきヌクレオチド配列の同定および定量を可能にする。典型的には、このようなプロファイリング試薬は、配列特異的な方法で相補的な塩基対によって標的核酸分子にハイブリダイズすることができ、試験試料に含まれる従来知られている形態などの他の核酸配列から標的配列を区別する。検出試薬の例には、表３に示されている配列の１つと実質的に相補的であるヌクレオチド配列を含む標的核酸にハイブリダイズするプローブが挙げられる。好ましい態様では、このようなプローブは、異なる遺伝子の核酸間で区別することができる。検出試薬の別の例には、逆転写または P C R などの、標的ポリヌクレオチドの相補鎖に沿ったヌクレオチド伸長の開始点として作用するプライマーが挙げられる。また、本明細書に示されている配列情報は、例えば、本発明のいずれかの遺伝子を逆転写および／または増幅（例えば、P C R を用いて）するプライマーを設計するために有用でもある。

【００８１】

本発明の１つの好ましい態様では、検出試薬は、表２に開示されている遺伝子のいずれかに対応する標的核酸分子のセグメントにハイブリダイズする、単離されたもしくは合成の D N A もしくは R N A ポリヌクレオチドプローブまたはプライマーまたは P N A オリゴマー、あるいは D N A 、 R N A および／または P N A の組み合わせである。ポリヌクレオチドの形態での検出試薬は、場合により、修飾された塩基類似体、インターカレート剤または副溝結合剤を含んでもよい。プローブなどの多重検出試薬は、例えば、固体支持体（例えば、アレイまたはビーズ）に付され、または溶液（例えば、P C R 、 R T - P C R 、 T a q M a n アッセイ、またはプライマー伸長反応）で供給されて、発現プロファイリングキットを形成してもよい。

10

20

30

40

50

【 0 0 8 2 】

プローブまたはプライマーは、典型的には、実質的に精製されたオリゴヌクレオチドまたは PNA オリゴマーである。このようなオリゴヌクレオチドは、典型的には、ストリンジェントな条件下で、標的核酸分子と少なくとも約 8、10、12、16、18、20、22、25、30、40、50、55、60、65、70、80、90、100、120（または中間の任意の他の数）、またはそれを超える連続したヌクレオチドにハイブリダイズする相補的ヌクレオチド配列の領域を含む。

【 0 0 8 3 】

他の好ましいプライマーおよびプローブ配列は、表 2 に開示されている遺伝子のヌクレオチド配列を用いて容易に決定することができる。このようなプライマーおよびプローブは、本発明の遺伝子の発現プロファイリングのための試薬として直接的に有用であり、いずれかのキット/システムフォーマットに導入することができることは当業者に明確である。

10

【 0 0 8 4 】

標的遺伝子配列に特異的なプローブまたはプライマーを生成するために、遺伝子/転写配列は、典型的には、ヌクレオチド配列の 5' 末端または 3' 末端で開始するコンピュータアルゴリズムを用いて調べられる。次に、典型的なアルゴリズムは、遺伝子配列に独特であり、ハイブリダイゼーションに適した範囲内の GC 含有量を有し、ハイブリダイゼーションを妨げる可能性のある予測される二次構造を欠如し、および/または他の所望の特徴を有し、または他の望ましくない特徴を欠如している、所定の長さのオリゴマーを同定する。

20

【 0 0 8 5 】

本発明のプライマーまたはプローブは、典型的には、長さにして、少なくとも約 8 ヌクレオチドである。本発明の一態様では、プライマーまたはプローブは、長さにして少なくとも約 10 ヌクレオチドである。好ましい態様では、プライマーまたはプローブは、長さにして少なくとも約 12 ヌクレオチドである。より好ましい態様では、プライマーまたはプローブは、長さにして少なくとも約 16、17、18、19、20、21、22、23、24 または 25 ヌクレオチドである。プローブの最大の長さは、用いられるアッセイのタイプに応じて、検出されるべき標的配列と同程度の長さであり得るが、典型的には、長さにして、約 50、60、65、または 70 ヌクレオチド未満である。プライマーの場合には、典型的には、長さにして約 30 ヌクレオチド未満である。本発明の特定の好ましい態様では、プライマーまたはプローブは、約 18 と約 28 ヌクレオチドの長さの範囲内である。しかしながら、他の態様では、核酸アレイ、プローブが基材に付される他の態様などのように、プローブはより長くてもよく、例えば、長さにして、30 ~ 70、75、80、90、100、またはより長いヌクレオチドである。

30

【 0 0 8 6 】

本発明は、修飾された、合成のもしくは天然に存在しないヌクレオチドまたは構成成分、あるいは当該技術分野において知られている代替の/修飾された核酸化学を含む核酸類似体を包含する。このような核酸類似体は、例えば、表 2 において特定されている遺伝子の 1 以上を検出するための検出試薬（例えば、プライマー/プローブ）として有用である。さらに、これらの類似体を含むキット/システム（例えば、ビーズ、アレイなど）も本発明によって包含される。例えば、本発明の多型配列に基づく PNA オリゴマーが、特に意図される。PNA オリゴマーは、DNA の類似体であり、そこでは、リン酸塩骨格は、ペプチド様骨格と置換される（Lagriffoulら, Bioorganic & Medicinal Chemistry Letters 4:1081-1082 [1994], Petersenら, Bioorganic & Medicinal Chemistry Letters 6:793-796 [1996], Kumarら, Organic Letters 3[9]:1269-1272 [2001], 国際公開第 WO96/04000 号）。PNA は、従来のオリゴヌクレオチドおよびオリゴヌクレオチド類似体よりも高いアフィニティーおよび特異性で相補的 RNA または DNA にハイ

40

50

ブリダイズする。PNAの特性は、伝統的なオリゴヌクレオチドおよびペプチドを用いては達成不可能な新しい分子生物学および生化学応用を可能にする。

【0087】

核酸の結合特性および/または安定性を改善する核酸修飾の追加の例には、塩基類似体、例えばイノソン、インターカレート剤（米国特許第4,835,263号）、例えばエチジウムブロマイドおよびSYBR（登録商標）Green、副溝結合剤（米国特許第5,801,115号）の使用を含む。このようにして、核酸分子、発現プロファイリング試薬（例えば、プローブおよびプライマー）、オリゴヌクレオチド/ポリヌクレオチドへの本明細書中での言及には、PNAオリゴマーおよび他の核酸類似体が含まれる。核酸類似体、当該技術分野において知られている代替の/修飾された核酸化学の他の例は、Current Protocols in Nucleic Acid Chemistry, John Wiley & Sons, New York (2002)に記載されている。

10

【0088】

各々の対立遺伝子に特異的なプライマーまたはプローブの設計は、標的核酸分子のヌクレオチド配列の正確な組成、プライマーまたはプローブの長さなどの変数に依存するが、プライマーおよびプローブの使用における別の因子は、プローブまたはプライマーと標的配列との間のハイブリダイゼーションが実行される条件のストリンジェンシーである。高いストリンジェンシーな条件は、低イオン強度および/または高い反応温度を有するバッファーを利用し、安定な二重鎖を形成するために、プローブ/プライマーと標的配列との間の密接な適合を必要とする傾向がある。しかしながら、ストリンジェンシーが高すぎると、ハイブリダイゼーションは全く起こらない場合もある。対照的に、低いストリンジェンシー条件は、高イオン強度および/または低反応温度を有するバッファーを有し、プローブ/プライマーと標的配列との間のより不適合な塩基を有する安定な二重鎖の形成を許す。一例であって、限定されないが、対立遺伝子に特異的なプローブを用いた高ストリンジェンシーハイブリダイゼーション条件のための例示的な条件は下記のとおりである：5×標準的なリン酸生理食塩水EDTA (SSPE)、0.5% NaDodSO₄ (SDS)を含む溶液を55℃で用いたプレハイブリダイゼーション、プローブと同溶液中で、同温度の標的核酸分子とのインキュベーション、その後、2×SSPE、0.1% SDSを含む溶液を55℃または室温で用いた洗浄。

20

30

【0089】

中程度のストリンジェンシーなハイブリダイゼーション条件は、例えば、約50 mM KClを含む溶液を約46℃で用いるプライマー伸長反応に対して使用されてもよい。あるいは、この反応は、60℃などの高温で実行されてもよい。別の態様では、中程度にストリンジェントなハイブリダイゼーション条件は、オリゴヌクレオチドの連結アッセイ (OLA) 反応に適し、ここで、2つのプローブは、標的配列の完全に相補的である場合に連結され、46℃の温度で約100 mM KClの溶液を利用してもよい。

【0090】

ハイブリダイゼーションに基づくアッセイでは、1つの遺伝子配列の標的DNAのセグメントにハイブリダイズするが、他の遺伝子に由来する配列にはハイブリダイズしない特異的なプローブを設計することができる。ハイブリダイゼーション条件は、遺伝子間のハイブリダイゼーション強度、好ましくは本質的に2成分応答において有意に検出可能な相違があるように十分にストリンジェントでなければならない。それによって、プローブは、遺伝子配列のうちのただ1つにハイブリダイズするかまたは1つの遺伝子配列の有意により強力にハイブリダイズする。標的部位がプローブの配列に沿ってどこでも整列されるように、プローブが特定の遺伝子の標的配列にハイブリダイズするように設計されてもよいが、このプローブは、好ましくは、プローブの中心位置（例えば、プローブのいずれかの末端からの少なくとも3ヌクレオチドであるプローブ内の位置）で遺伝子配列が整列されるように標的配列のセグメントにハイブリダイズするように設計される。このプローブの設計は、一般に、異なる遺伝子間のハイブリダイゼーションにおいて良好な区別を達成

40

50

する。

【0091】

オリゴヌクレオチドプローブおよびプライマーは、当該技術分野において周知の方法によって調製されてもよい。化学合成法には、限定されないが、Narangら, Methods in Enzymology 68:90 [1979]によって記述されているホストリエステル法; Brownら, Methods in Enzymology 68:109 [1979]によって記述されているホスホジエステル法、Beaucageら, Tetrahedron Letters 22:1859 [1981]によって記述されているジエチルホスホアミデート法; 米国特許第4,458,066号に記述されている固体支持法が含まれる。アレイの場合には、複数のプローブは、複数の異なる遺伝子配列の同時分析のための同じ支持体上に固定化されてもよい。

10

【0092】

PCRに基づくアッセイの1タイプでは、遺伝子特異的なプライマーは、遺伝子配列に重複する標的核酸分子上の領域にハイブリダイズし、プライマーが完全な相補性を示す遺伝子配列の増幅をプライムするだけである(Gibbs, Nucleic Acid Res. 17:2427-2448 [1989])。典型的には、プライマーの3'最大ヌクレオチドは、標的核酸分子と整列し、相補的となる。このプライマーは、遠位部位をハイブリダイズする第2プライマーとともに用いられる。増幅は、2つのプライマーから開始し、どの遺伝子/転写物が試験試料に存在するかを指示する検出可能な生成物を生成する。このPCRに基づくアッセイは、下記に示されるTaqManアッセイの一部として利用することができる。

20

【0093】

本明細書に記載されている14個の遺伝子サインの遺伝子は、核酸試料に含まれる対象とするポリヌクレオチドのコピー数を増加させるために用いられる、種々の核酸増幅法のうちのいずれか1つによって検出することができる。このような増幅法は、当該技術分野において周知であり、それらには、限定されないが、ポリメラーゼ連鎖反応(PCR)(米国特許第4,683,195号および同第4,683,202号; PCR Technology: Principles and Applications for DNA Amplification, 編集, H.A. Erlich, Freeman Press, New York, New York [1992])、リガーゼ連鎖反応(LCR)(WuおよびWallace, Genomics 4:560 [1989]; Landegrenら, Science 241:1077 [1988])、鎖置換増幅(SDA)(米国特許第5,270,184号および同第5,422,252号)、転写媒介増幅(TMA)(米国特許第5,399,491号)、連結リンカー増幅(LLA)(米国特許第6,027,923号)など、等温増幅法、例えば、核酸配列に基づく増幅(NASBA)、自立した配列複製(Guattelliら, Proc. Natl. Acad. Sci. USA 87:1874 [1990])が含まれる。このような方法論に基づいて、当業者は、本明細書に開示されている遺伝子を増幅させるために、対象とする遺伝子配列の任意の適切な領域5'および3'において容易にプライマーを設計することができる。このようなプライマーは、任意の長さのDNAを逆転写し、増幅させるために用いることができ、その結果、その配列に対象とする遺伝子を含む。

30

40

【0094】

一般に、増幅されたポリヌクレオチドは、長さにして少なくとも16ヌクレオチドである。より典型的には、増幅されたポリヌクレオチドは、長さにして少なくとも約20ヌクレオチドである。本発明の好ましい態様では、増幅されたポリヌクレオチドは、長さにして少なくとも約30ヌクレオチドである。本発明のより好ましい態様では、増幅されたポリヌクレオチドは、長さにして少なくとも約32、40、45、50、または60ヌクレオチドである。本発明のなお別の好ましい態様では、増幅されたポリヌクレオチドは、長さにして少なくとも約100、200、300、400、または500ヌクレオチドである。本発明の増幅されたポリヌクレオチドの全長は、エクソン、イントロンまたは全遺伝

50

子と同程度の長さであり得るが、増幅産物は、典型的には、長さにして最大約 1,000ヌクレオチドである（しかし、ある種の増幅方法は、長さにして1,000ヌクレオチドよりも大きい増幅産物を生じてよい）。より好ましくは、増幅されたポリヌクレオチドは、長さにして約150～250ヌクレオチドよりも大きくはない。

【0095】

本発明の態様では、本発明の遺伝子発現プロファイリング試薬は、検出可能なシグナルを発光する蛍光レポーター色素を用いて標識される。好ましいレポーター色素は蛍光色素であるが、オリゴヌクレオチドプローブまたはプライマーなどの検出試薬に結合することができる任意のレポーター色素は本発明における使用に適している。このような色素には、限定されないが、Acridine、AMCA、BODIPY、Cascade Blue、Cy2、Cy3、Cy5、Cy7、Dabcyl、Edans、Eosin、Erythrosin、Fluorescein、6-Fam、Tet、Joe、Hex、Oregon Green、Rhodamine、Rhodol Green、Tamra、Rox、およびTexas Redが含まれる。

【0096】

本発明のなお別の態様では、検出試薬は、特に、その試薬がTaqMan（米国特許第5,210,015号および同第5,538,848号）などの自己クエンチングプローブ、または分子指標プローブ（米国特許第5,118,801号および同第5,312,728号）、あるいは他のステムレスまたは線状指標プローブ（Livakら, PCR Method Appl. 4:357-362[1995]; Tyagiら, Nature Biotechnology 14:303-308 [1996]; Nazarenkoら, Nucl. Acids Res. 25:2516-2521[1997]; 米国特許第5,866,336号および同第6,117,635号）として用いられる場合に、Tamraなどのクエンチャー色素を用いてさらに標識されてもよい。

【0097】

また、本発明の検出試薬は、他の標識を含んでもよく、限定されないが、ストレプトアビジン結合に対するビオチン、抗体結合に対するハプテン、ジップコードの対などの別の相補性オリゴヌクレオチドへの結合に対するオリゴヌクレオチドが含まれる。

遺伝子発現キットおよびシステム

当業者は、本明細書に開示されている遺伝子および配列情報に基づいて、発現プロファイリング試薬が、個別にまたは組み合わせて本発明のいずれかの遺伝子を評価するために作製され、用いることができ、このような検出試薬が、当該技術分野において周知である確立したキットまたはシステムの1つに容易に組み込むことができることを認識する。用語「キット」および「システム」は、遺伝子発現プロファイリング試薬との関連において本明細書中で使用するとき、複数の遺伝子発現プロファイリング試薬の組み合わせ、または1以上のタイプのエレメントもしくは成分（例えば、他のタイプの生化学試薬、容器、包装、例えば市販用に意図された包装、遺伝子プロファイリング試薬が結合する基材、電子機器コンポーネントなど）と組み合わせた1以上の遺伝子発現プロファイリング試薬などのようなものを指すことが意図される。したがって、本発明は、さらに、遺伝子発現プロファイリングキットおよびシステムを提供し、限定されないが、パッケージ化されたプローブおよびプライマーセット（例えば、TaqManプローブ/プライマーセット）、核酸分子のアレイ/マイクロアレイ、1以上のプローブ、プライマー、または本発明の1以上の遺伝子をプロファイルするための他の検出試薬が挙げられる。キット/システムは、場合により、種々の電子機器コンポーネントを含むことができる；例えば、種々の製造業者から提供されるアレイ（「DNAチップ」）および微小流体システム（「lab-on-a-chip」システム）は電子機器コンポーネントを含む。他のキット/システム（例えば、プローブ/プライマーセット）は、電子機器コンポーネントを含まなくてもよいが、例えば、1以上の容器に包装された1以上の遺伝子発現プロファイリング試薬（並びに、場合により、他の生化学試薬）から構成されてもよい。

【0098】

ある態様では、遺伝子発現プロファイリングキットは、典型的には、1以上の検出試薬、並びに、遺伝子を含む核酸分子の逆転写、増幅および/または検出などのアッセイまたは反応を行うために必要な他の成分（例えば、バッファー、酵素、例えば逆転写酵素、DNAポリメラーゼまたはリガーゼ、逆転写および鎖伸長ヌクレオチド、例えばデオキシヌクレオチド三リン酸、サンガータイプのDNAシーケンシング反応の場合には、連鎖停止ヌクレオチド、陽性対照配列、陰性対照配列など）を含む。さらに、キットは、標的核酸の量を測定するための手段、その量と標準とを比較するための手段を含み、対象とする遺伝子を含む核酸分子を検出するキットを用いるための指示書を含むことができる。本発明の一態様では、本明細書に開示されている1以上の遺伝子の発現をプロファイルする1以上のアッセイを行うために必要な試薬を含むキットが提供される。本発明の好ましい態様では、遺伝子発現プロファイリングキット/システムは、核酸アレイ、または区分化されたキットの形態であり、微小流体/ lab-on-a-chip システムが含まれる。

【0099】

遺伝子発現プロファイリングキット/システムは、例えば、各々の標的遺伝子配列の位置でまたはその近傍で核酸分子にハイブリダイズする1以上のプローブ、またはプローブ対を含んでもよい。遺伝子特異的なプローブの複数の対は、大多数の遺伝子を同時に評価するキット/システムに含むことができ、それらの遺伝子のうち少なくとも1つの遺伝子は、本発明の遺伝子である。あるキット/システムでは、遺伝子特異的なプローブは、アレイまたはビーズなどの基材に固定されている。例えば、同基材は、表2に示されている遺伝子の少なくとも1つもしくは実質的に全部の遺伝子、またはその中間の任意の他の数において検出するための遺伝子特異的なプローブを含むことができる。

【0100】

用語「アレイ」、「マイクロアレイ」、および「DNAチップ」は、本明細書において、基材（例えば、ガラス、プラスチック、紙、ナイロンまたは他のタイプのメンブレン、フィルタ、チップ、もしくは他の適切な固体支持体）に付された別個のポリヌクレオチド群のアレイを指すように互換可能に使用されている。ポリヌクレオチドは、上記基材上で直接合成されるか、または基材とは別個に合成された後、その基材に付されてもよい。一態様では、マイクロアレイは、米国特許第5,837,832号（Cheer）、PCT国際出願第WO95/11995号（Cheer）、Lockhart, D. J. ら（Nat. Biotech. 14:1675~1680 [1996]）およびSchena, M. ら（Proc. Natl. Acad. Sci. 93:10614~10619 [1996]）に記載される方法に従って、調製され、使用される。これらの文献の全ては、全体として参照により本明細書に援用される。他の態様では、このようなアレイは、Brown ら、米国特許第5,807,522号により記載される方法により生成される。

【0101】

核酸アレイは、下記の参考文献に概説されている：Zammatteo ら、"New chips for molecular biology and diagnostics," Biotechnol. Annu. Rev. 8:85-101 (2002); Sosnowski ら、"Active microelectronic array system for DNA hybridization, genotyping and pharmacogenomic applications," Psychiatr. Genet. 12(4):181-92 (Dec. 2002); Heller, "DNA microarray technology: devices, systems, and applications," Annu. Rev. Biomed. Eng. 4:129-53 (2002); Epub Mar. 22 2002; Kolchinsky ら、"Analysis of SNPs and other genomic variations using gel-based chips," Hum. Mutat. 19(4):343-60 (Apr. 2002); McGall ら、"High-density genechip oligonucleotide probe arrays," Adv. Biochem. Eng. Biotechnol. 77:

21 - 42 (2002)。

【0102】

任意数のプローブ、例えば遺伝子特異的プローブは、アレイにおいて実施されてもよく、各プローブまたはプローブ対は、異なる遺伝子配列位置にハイブリダイズすることができる。ポリヌクレオチドプローブの場合、それらは、光指向性化学プロセスを用いて、基材上の指定領域で合成することができる（または、別個に合成された後、指定領域に付することができる）。各DNAチップは、格子様パターンで整列され（例えば、10セント硬貨の大きさに）小型化された、例えば、数千から数億個の個々の合成ポリヌクレオチドプローブを含むことができる。好ましくは、プローブは、順序立ったアドレス指定可能なアレイで固体支持体に付される。

10

【0103】

マイクロアレイは、固体支持体に固定された多数の独特の一本鎖ポリヌクレオチド、通常、合成アンチセンスポリヌクレオチドまたはcDNAフラグメントのいずれかから構成され得る。典型的なポリヌクレオチドは、好ましくは長さにして約6～60ヌクレオチド、より好ましくは長さにして約15～30ヌクレオチド、最も好ましくは長さにして約18～25ヌクレオチドである。ある種のタイプのマイクロアレイまたは他の検出キット/システムに関しては、長さにしてほんの約7～20ヌクレオチドであるオリゴヌクレオチドを用いることが好ましい場合がある。他のタイプのアレイ、例えば化学発光検出技術と組み合わせて使用されるアレイにおいて、好ましいプローブの長さは、例えば、長さにして約15～80ヌクレオチド、好ましくは長さにして約50～70ヌクレオチド、より好ましくは長さにして約55～65ヌクレオチド、最も好ましくは長さにして約60ヌクレオチドであってもよい。このマイクロアレイまたは検出キットは、遺伝子/転写物の既知の5'配列または3'配列を網羅するポリヌクレオチド、遺伝子/転写物の全長配列を網羅する連続するポリヌクレオチド；または標的遺伝子/転写物配列の長さに沿った特定の領域、特に、表2に開示されている1以上の遺伝子に対応する領域から選択される独特のポリヌクレオチドを含むことができる。このマイクロアレイまたは検出キットにおいて用いられるポリヌクレオチドは、対象とする遺伝子または複数の遺伝子に対して特異的（例えば、標的遺伝子配列内の特定のサイン配列に対して特異的、または複数の異なる配列部位にある特定の遺伝子配列に対して特異的）であり得るか、あるいは対象とする多型遺伝子/転写物または複数の遺伝子/複数の転写物に対して特異的であり得る。

20

30

【0104】

ポリヌクレオチドアレイに基づくハイブリダイゼーションアッセイは、完全一致した標的配列およびミスマッチした標的配列に対するプローブのハイブリダイゼーション安定性の差異に依存する。

【0105】

他の態様では、アッセイは、化学発光検出技術と組み合わせて用いられる。下記の特許および特許出願は、その全てが全体として参照により本明細書中に援用され、化学発光検出に関するさらなる情報を提供する：米国特許出願第10/620332号および同第10/620333号は、マイクロアレイ検出のための化学発光アプローチを記載し；米国特許第6124478号、同第6107024号、同第5994073号、同第5981768号、同第5871938号、同第5843681号、同第5800999号、および同第5773628号は、化学発光検出を実施するための方法およびジオキセタンの組成物を記載し；米国出願公開第US2002/0110828号は、マイクロアレイコントロールのための方法および組成物を開示する。

40

【0106】

本発明の一態様では、核酸アレイは、長さにして約15～25ヌクレオチドのプローブのアレイを含むことができる。さらなる態様では、核酸アレイは、任意数のプローブを含むことができ、ここで、少なくとも1つのプローブは、表2に開示されている1以上の遺伝子を検出することができ、および/または少なくとも1つのプローブは、表2に開示されている配列からなる群より選択される遺伝子配列、並びにそれに相補的な配列のうちの

50

1つのフラグメントを含み、そのフラグメントは、少なくとも約8個の連続するヌクレオチド、好ましくは、10個、12個、15個、16個、18個、20個、より好ましくは22個、25個、30個、40個、47個、50個、55個、60個、65個、70個、80個、90個、100個（またはその中間の他の任意の個数）以上連続するヌクレオチド、表2に開示されている遺伝子の配列（またはそれに相補的である）を含む。ある態様では、この遺伝子部位に対して相補的なヌクレオチドは、上記プローブの中心から5、4、3、2、または1ヌクレオチド内、より好ましくはそのプローブの中心にある。

【0107】

ポリヌクレオチドプローブは、全体として参照により本明細書中に援用されるPCT国際出願第WO95/251116号（Baldeschweilerら）に記載されるように、化学カップリング手順およびインクジェット適用装置を用いることによって、基材表面上で合成され得る。別の局面では、ドット（またはスロット）プロットと類似する「格子状」アレイが、真空系、熱的結合手順、UV結合手順、機械的結合手順、または化学結合手順を用いて、基材表面にcDNAフラグメントまたはオリゴヌクレオチドを整列させて結合するために使用され得る。アレイ、例えば、上記のアレイは、手動によって、あるいは利用可能なデバイス（スロットプロット装置またはドットプロット装置）、材料（適切な任意の固体支持体）、および機器（ロボット機器を含む）を用いることによって生成することができ、8個、24個、96個、384個、1536個、6144個以上のポリヌクレオチド、または市販の機器の効率的使用のために適する他の任意の数のポリヌクレオチドを含んでもよい。

【0108】

このようなアレイまたは他のキット/システムを用いて、本発明は、試験試料中に含まれる本明細書に開示されている遺伝子を同定し、その発現をプロファイリングするための方法を提供する。このような方法は、典型的には、核酸の試験試料を、本発明の少なくとも1つの遺伝子配列の位置に対応する1以上のプローブを含むアレイとともにインキュベートすること、並びにこの試験試料由来の核酸とプローブの1以上との結合について評価することを伴う。遺伝子発現プロファイリング試薬（またはこのような1以上の遺伝子発現プロファイリング試薬を用いるキット/システム）を試験試料とともにインキュベートする条件は変化する。このインキュベーション条件は、アッセイにおいて用いられるフォーマット、用いられるプロファイリング方法、並びにアッセイにおいて用いられるプロファイリング試薬のタイプおよび性質などの因子に依存する。当業者は、市販のハイブリダイゼーション、増幅およびアレイアッセイフォーマットのうちのいずれか1つが、本明細書に開示されている遺伝子を検出するために容易に適合され得ることを認識する。

【0109】

本発明の遺伝子発現プロファイリングキット/システムは、試験試料から核酸を調製し、その後、逆転写、RNA濃縮、遺伝子配列を含む核酸分子の増幅および/または検出のために用いられる成分を含んでもよい。このような試料調製成分は、いずれかの腫瘍組織供給源、例えば、限定されないが、新鮮な腫瘍生検、凍結もしくはホルマリン固定したパラフィン包埋（FFPE）の組織標本、またはいずれかの方法によって回収および保存された腫瘍由来の核酸抽出物（DNA、cDNAおよび/またはRNA）を生成するために用いることができる。上述される方法において用いられる試験試料は、アッセイフォーマット、プロファイリング方法の性質、アッセイされるべき試験試料として用いられる特定の組織、細胞または抽出物などの因子に基づいて変更する。核酸を調製する方法は、当該技術分野において周知であり、利用されるシステムと適合可能な試料を得るために容易に適用され得る。試験試料から核酸を抽出するための自動化された試料調製システムは市販され、例えば、Qiagen's BioRobot 9600, Applied BiosystemsのPRISM 6700、Roche Molecular SystemsのCOBAS AmpliPrep Systemが挙げられる。

【0110】

本発明によって意図されるキットの別の形態は、区分化されたキットである。区分化さ

10

20

30

40

50

れたキットには、試薬が別々の容器に含められているいずれかのキットが含まれる。このような容器には、例えば、小さなガラス容器、プラスチック容器、プラスチック、ガラスもしくはペーパーのストリップ、またはシリカなどのアレイ用材料が含まれる。このような容器は、試験試料および試薬が二次汚染しないように1つの区分から別の区分に、あるいは1つの容器からキットに含まれない別の入れ物に、試薬を効率的に移すことができる。このような容器には、例えば、試験試料を受け入れる1以上の容器、本発明の1以上の遺伝子の発現をプロファイリングするための少なくとも1つのプローブまたは他の遺伝子発現プロファイリング試薬を含む1以上の容器、洗浄試薬（例えば、リン酸緩衝生理食塩水、T r i s バッファーなど）を含む1以上の容器、結合したプローブまたは他の遺伝子発現プロファイリング試薬の存在を示すために用いられる試薬を含む1以上の容器が含まれてもよい。このキットは、場合により、さらに、区分および/または試薬、例えば、逆転写、RNA濃縮、核酸増幅または他の酵素反応、例えばプライマー伸長反応、ハイブリダイゼーション、連結、電気泳動（好ましくはキャピラリー電気泳動）、質量分析、および/またはレーザー誘導の蛍光検出のための試薬を含む。また、キットは、そのキットを用いるための指示書を含んでもよい。例示的な区分化されたキットは、当該技術分野において知られている微小流体デバイスを含む（例えば、Weiglら, "Lab-on-a-chip for drug development," Adv. Drug Deliv. Rev. 24, 55 [3]: 349 - 77 [2003年2月]を参照されたい）。このような微小流体デバイスでは、容器は、例えば、微小流体「区分」、「チャンバー」、または「チャネル」とも呼ばれる場合がある。

10

20

遺伝子発現プロファイリング試薬の使用

本発明の表3に示される核酸分子は、特に、乳癌転移の予後における種々の用途を有する。例えば、核酸分子は、増幅プライマーまたはハイブリダイゼーションプローブとして、例えば、メッセンジャーRNA、転写RNA、cDNA、ゲノムDNA、増幅されたDNAまたは他の核酸分子を用いた発現プロファイリングのために、並びに、表2に開示されている遺伝子およびそれらの相同体をコードする全長cDNAおよびゲノムクローンを単離するために有用である。

【0111】

プローブは、核酸分子の全長に沿って、いずれかのヌクレオチド配列にハイブリダイズすることができる。好ましくは、本発明のプローブは、表2に指示されている遺伝子の遺伝子配列を包含する標的配列の領域にハイブリダイズする。より好ましくは、プローブは、配列特異的な方法で、遺伝子を含む標的配列にハイブリダイズし、その結果、標的配列と、標的配列とは異なる他のヌクレオチド配列とを区別する。このようなプローブは、試験試料に含まれる遺伝子を含む核酸の存在の検出に特に有用である。

30

【0112】

このようにして、本発明の核酸分子は、本明細書に開示されている遺伝子の発現レベルを検出およびプロファイリングするために、ハイブリダイゼーションプローブ、逆転写および/または増幅プライマーとしても用いることができ、それにより、乳癌およびその発現プロファイリングを有する個体は遠隔転移に対するリスクにあることを決定する。開示されている発現プロファイリングは、遠隔転移に対する予後ツールを提供する。

40

転移スコアの作成

表2に開示されている14個の遺伝子の発現レベルは、転移リスクを予測する転移スコア(MS)を誘導するために用いることができる。発現レベルは、 (C_t) 法によって計算されてもよく、ここで、 C_t は、標的増幅のための閾値サイクル；即ち、標的の時間の指数関数的な増幅が始まるPCRにおけるサイクル数である。(KJ Livak and TD Schmittgen, 2001, Methods 25: 402 - 408)。14個のプロファイルされた遺伝子の各々のmRNAのレベルは、

$$(C_t) = (C_{t_{GOI}} - C_{t_{EC}})_{testRNA} - (C_{t_{GOI}} - C_{t_{EC}})_{refRNA}$$

として定義されてもよく、ここで、GOIは対象とする遺伝子（14個のサイン遺伝子の

50

各々)であり、*test RNA*は患者試料から得られたRNAであり、*ref RNA*は校正参照RNAであり、*EC*は内在性対照である。各サイン遺伝子の発現レベルは、初めに、表2(*EC*)に列挙されている3個の内在性対照遺伝子に標準化されてもよい。3個の内在性対照(*Ct_{EC}*)の増幅から得られた、*Ct*(複数)の平均の*Ct*を用いて、ただ1つの対照遺伝子が用いられる場合に発生する標準化バイアスのリスクを最小化することができる。(T. Suzuki, P. J. Higginsら, 2000, *Biotechniques* 29:332-337)。内在性対照遺伝子を増幅するために選択的に用いられてもよいプライマーを表3に列挙している;しかし、これらの内在性対照を増幅することができるプライマーは、これらの開示されたオリゴヌクレオチドに限定されない。対象とする遺伝子の調整された発現レベルは、校正参照RNAプール、*ref RNA*(ユニバーサルヒト参照RNA、*Stratagene, La Jolla, Calif.*)にさらに標準化されてもよい。これは、種々の機械から得られた発現結果を標準化するために用いられてもよい。

10

(*Ct*)値は、14個のサイン遺伝子の各々についての遺伝子発現プロファイリングにおいて得られ、転移スコア(*MS*)を作成するための下記の式:

【0113】

【数4】

$$MS = a0 + \sum_{i=1}^M ai * Gi$$

20

(式中、*Gi*は14個の遺伝子予後サインの各遺伝子(*i*)の発現レベルを表す)に用いられてもよい。*Gi*の値は、上述される発現プロファイリングにおいて得られる(*Ct*)である。各遺伝子*i*についての定数*ai*は、表2に与えられている。定数*a0*は0.022である;これは、*MS*を中心にし、それにより、その中央値はゼロとなる。*M*は、成分リストにおける遺伝子数である;この場合は14である。このようにして、*MS*は、表2に開示されている14個の遺伝子についての発現レベルの総和の測定であり、各々は、表2にもある特定の定数*ai*によって乗じられ、最終的には、この総和は、*MS*を誘導するために中心の定数0.022に加えられる。

【0114】

あるいは、(*Ct*)値は、14個のサイン遺伝子の各々についての遺伝子発現プロファイリングにおいて得られ、転移スコア(*MS*)を作成するための下記の式:

30

【0115】

【数5】

$$MS = a0 + b * \sum_{i=1}^M ai * Gi$$

(式中、*Gi*は14個の遺伝子予後サインの各遺伝子の標準化された発現レベルを表す)に用いられてもよい。*Gi*の値は、上述される発現プロファイリングにおいて得られる(*Ct*)において測定される元々の発現レベルから平均の遺伝子発現を差し引き、次に、トレーニングセットにおける遺伝子発現の標準偏差によって割ることにより得られる。各遺伝子*i*に対する定数*ai*は、表2に与えられている。定数*b*は、-0.251であった。正しいサインおよびスケールリングを得るには、予測材料として主成分を用いて一変量Coxモデルからであった。定数0は0.022である;これは、*MS*を中心にし、それにより、その中央値はゼロとなる。*M*は、成分リストにある遺伝子数である;この場合は14である。このようにして、*MS*は、表2に開示されている14個の遺伝子についての発現レベルの総和の測定であり、各々は、表2にも示されている特定の定数*ai*によって乗じられる。この総和は、定数*b*によって乗じられ、次に、中心定数0.022は、*MS*を誘導するために加えられる。

40

【0116】

いずれかの新しい試料は、その患者に関する14個の遺伝子発現プロファイリングデー

50

タからこの転移スコアを作成することによって評価されてもよく、このスコアから、患者についての遠隔転移の可能性を測定することができる。

【0117】

MSスコアは、上述される (Ct) の値の和を単純化することができ、この場合、MSの式は、 0 の値をゼロで置換することによって単純化され、定数 a_i は1であることに留意されたい。

【0118】

また、MSは、上述される (Ct) の値の和を単純化し、次に、正しいサインおよびスケールについて定数 - 0.04778によって乗じられ、それにより、遠隔転移リスクは、MSの増加とともに増えることに留意されたい。最後に、定数 0.8657は、MSがゼロになるように添加される。この代替の方法において誘導されるMSスコアは、14個全ての遺伝子の等しい計量を有する。遠隔転移のリスクは、MSが増加するにつれて増える。本明細書に記載されている2つの異なるMSスコアは、非常に高い相関を有し、ピアソン相関係数は0.999より大きい。

MSからの遠隔転移可能性の作成

いずれかの個体の患者に関する遠隔転移の可能性は、基礎生存関数としてWeibull分布を用いて、変更可能な時間点でMSから計算され得る。

【0119】

14個の遺伝子サインの発現プロファイリングから、上記で得られた転移スコア (MS) は、Cox比例ハザードモデルによって遠隔転移の可能性に転換された。Coxモデルは、基準のハザード機能を特定しないため、ハザードおよび生存機能は、初めに、パラメトリック抑制モデルを通じて構築された。パラメトリックな抑制モデルにおいて、遠隔転移なしの生存時間は結果であり、転移スコア (MS) は独立した可変インプットであった。イベント時間は、Weibull分布を有するように想定された；その2つのパラメータは、MSが誘導される生存データを用いて推定された。患者についてある種の時間内に遠隔転移の可能性を計算するために、MS値は、単に、生存機能についてその式に置換される。

リスク測定におけるMSスコアの臨床応用

転移に関するリスクの測定にMSスコアを用いるという1つの方法は、MS「カットポイント」としても知られている1以上のMS閾値を作成することである。このようなMS閾値は、このような患者が増加したリスクまたは減少したリスクのいずれかを有するかを決定するために、乳癌患者のMSスコアと比較した場合に、基準として用いることができる。MS閾値は、異なる方法によって測定することができ、転移スコアの異なる定義によって異なっている。実施例1、2および3で用いられる等式1において定義されているMSに関して、MS閾値は、高リスク群対低リスク群のハザード率から決定された。遠隔転移なしの生存についてのカプラン-マイヤー (KM) 曲線は、MSカットポイントによって特定された高および低リスクの患者群について作成される。カットポイントとしての中央値MSの選択は、MSの10パーセンタイルからMSの90パーセンタイルの異なるカットポイントを用いた、高リスク群対低リスク群のハザード率の計算に基づいている。中央値のカットポイントは、高および低リスク群において等しい数の個体が存在する場合に、カットポイントとして特定することができ、実施例1に記載されるトレーニング試料において最大に近いハザード率を生み出すことが見出される。中央値MSのカットポイントを用いたハザード率 (HR) および95%の信頼区間が計算され、報告され得る。ログラंक (log rank) 試験が行われ、ハザード率は、異なるカットポイントについて計算される。5年での遠隔転移の予測における14個の遺伝子サインの精度および値は、種々の手段によって評価することができる。 (XH Zhou, N. Obuchowskiら, 編集, 2002, Statistical Methods in Diagnostic Medicine, Wiley-Interscience, New York)。実施例4および5で用いられる等式2において定義されているMSに関して、MS閾値は、実施例2に記載されているガイ病院からの試料における5年以内の遠隔転移を予

10

20

30

40

50

測するためのMSの感受性および特異性から決定される。2つのカットポイントは、5年以内の遠隔転移を予測するためのMSの感受性が、第1のカットポイントが用いられると90%を超えるように選択される。第2のカットポイントは、5年以内の遠隔転移を予測するためのMSの感受性および特異性がともに70%であるように選択される。等式2において定義されているMSに関して、第1のMS閾値は-0.1186であり、第2のMS閾値は0.3019である。2つのMSカットポイントを用いた場合、高、中、および低MS群がある。ガイ病院からの処置された試料、日本の愛知県がんセンターからの試料では、高MS群は、高リスク群として設計され、中および低MS群は低リスク群として設計される。

【実施例】

10

【0120】

下記の実施例は、例証するために提供され、請求される発明を限定するものではない。
(実施例1)

14個の遺伝子予後サインのmRNA発現レベルが142人のリンパ節転移陰性のER陽性乳癌患者の遠隔転移に対するリスクを予測する

下記の実施例は、14個の遺伝子予後サインがどのようにして同定され、日常的な臨床検査室試験においてでさえ、乳癌患者の遠隔転移に関する予後の決定にどのように用いることができるかを例証する。臨床医は、生検、FFPE、凍結組織などの多くの手段から得られたRNAを用いて、本明細書に開示されている14個の遺伝子に基づくmRNA発現プロファイリングを行い、次に、予後転移スコアを測定するために本明細書に与えられたアルゴリズムに発現データを挿入することができる。

20

【0121】

リンパ節転移陰性のER陽性乳癌患者から得られたFFPE組織は、後述される実施例に用いられた。初期セットの200個の遺伝子を分析して、最終の14個の遺伝子サインを誘導した。文献で以前に報告された遺伝子は、このサインのための候補遺伝子として含められた。また、本実施例では、このサインと、日常的に使用されている予後因子およびツールとの重複の程度を測定した。

【0122】

リンパ節転移陰性のER陽性患者からの腫瘍は、この試験のために選択され、それは、リンパ節転移陰性患者に関する予後情報が処置戦略の指標に非常に価値があるからである。また、マイクロアレイ研究は、この腫瘍のサブセットが他のタイプの乳癌腫瘍と臨床的に区別されることを指示する。(T. Sorlie, CM Perou, 2001, Proc Natl Acad Sci USA 98:10869-10874; C. Sotiriou, SY Neo, 2003, Proc Natl Acad Sci USA 100:10393-10398)。FFPE切片において、H. Dai (H. Dai, L.J. van't Veer, 2005, Cancer Res 15:4059-4066)、L.J. van't Veer (L.J. van't Veer, H. Dai, 2002, Nature 415:530-536)、S.P. Paik (S.P. Paik, S. Shak, 2004, N Engl J Med 351:2817-2826)によって報告された遺伝子サインからの発現プロファイリングのために遺伝子が選択され、これらの遺伝子の構造安定性、並びに日常的に回収され、保存された臨床試料が転移の予後に用いられ得る程度を測定した。遺伝子発現データから、転移スコアを製作して、いずれかの期間、個々の患者における遠隔転移の可能性を推測した。

30

40

患者および試料

初期段階の乳癌を有する、合計142人のリンパ節転移陰性のER陽性患者が選択され、全ては、全身性アジュバント療法で処置されていない患者由来であった(表1のトレーニング試料)。この研究を乳癌症例のサブセットに限定することによって、分子サインは、転移とのより説得力のある関連性を伴って同定され、異なる試料セット全体でより堅固であり、日常的な臨床実施への転換を良好に促進するように少量の遺伝子を含んでいる。患者の平均年齢は、約62歳である(31~89歳の範囲)。

50

【0123】

非常に特徴付けられた乳房腫瘍試料セットが本研究のための試料の供給源として役割を果たした；このセットは、カリフォルニア パシフィック メディカルセンター (California Pacific Medical Center: CPMC) で1975から1986年まで生じた。初期の研究のための試験対象患者基準は、リンパ節転移陰性であり、全身療法を受けておらず、8年間フォローアップケアを受けた患者由来の腫瘍からの試料を含んでいた。

【0124】

各機関の医療倫理委員会によって、本研究における使用のための試料が承認された。試料を提供する患者は、ESR1遺伝子の発現レベルの測定に基づいて、ER陽性として分類された。ESR1遺伝子の発現レベルは、個体のER状態とよく相関する。(M. Cronin, M. Phor, 2004, Am J Pathol 164(1):35-42; JM Knowlden, JM Geer, 1997, Clin Cancer Res 3:2165-2172)。

【0125】

遠隔転移のない生存が、癌に関連した死亡に最も直接的に連結されるため、一次評価項目として選択された。二次評価項目は、全生存であった。

試料加工

各々のパラフィンブロックから4つの10 μ m切片をRNA抽出のために用いた。腫瘍領域は、ガイドスライドに基づいて取り出し、そこでは、癌の細胞領域は、病理学者によって標識され、Pinpoint Slide RNA Isolation System II (Zymo Research, Orange CA) を用いてRNAを抽出した。

【0126】

トータルRNAの収量は試料間で変化した。分析に利用可能なRNA量を増やすために、T7 RNAポリメラーゼ線状増幅が、抽出されたRNAに対して行われた。FFPE試料から単離されたRNAは、Message AmpliA RNA amplification kit (Ambion, Austin, Texas) を用いて、T7に基づくRNA増幅に供した。

【0127】

RNA増幅の前後で、遺伝子発現の一貫性を評価するために、多数の実験が、異なる試料に含まれる種々の遺伝子に対して行われた。初めに、0.1~100ngのトータルRNAを用いて、本研究の一部ではない67個のFFPE試料由来のRNAに対して増幅を行った。20個の遺伝子のプロファイリングは、得られた濃縮RNAおよび元々の濃縮されていないRNAを用いて行われた。これらの比較は、倍数濃縮が遺伝子間で変換することを示した；しかしながら、相対的な発現レベルは、67個の試料のうち20個全ての遺伝子において、RNA増幅の前後で一貫していた。

【0128】

本研究のためのRNAは、上述したように、Message AmpliA RNA増幅キットを用いた増幅によって濃縮された。トータルRNAは、分光光度測定(OD₂₆₀)を用いて定量された。

遺伝子発現プロファイリング

既刊文献の調査およびマイクロアレイに基づく遺伝子発現プロファイリング実験の結果に基づいて、最適な予後サインを決定するために、最初に、200個の候補遺伝子が分析用に選択された。このセットは、van't Veerら(LJ van't Veer, H. Dai, 2002, Nature 415:530-536)によって記載された70個の遺伝子の予後パネル、Daiら(H. Dai, LJ van't Veer, 2005, Cancer Res 15:4059-4066)によって分析された104個の遺伝子、Paikら(SP Paik, S. Shak, 2004, N Engl J Med 351:2817-2826)によって報告されたタモキシフェン処置

10

20

30

40

50

に対する応答のためのサインを含む16個の遺伝子のパネル、Westら(M. West, C. Blanchetteら, 2001, Proc Natl Acad Sci U S A 98: 11462 - 11467)によって報告された24個のER関連遺伝子由来の遺伝子を含んでいた。

【0129】

J. Vandesompele, K. De Preterら, Genome Biol 3(7): Research 0034.1 - 0034.11 (Epub 2002)に記載されている方法に従って、発現データを標準化するために内在性対照(EC)として、追加の遺伝子が選択された。また、内在性対照は、本明細書では「ハウスキーピング遺伝子」とも呼ばれる。6個の内在性対照遺伝子は、凍結された乳癌腫瘍の150個の試料におけるそれらの発現レベルの安定性について試験された。VandesompeleらのgeNormプログラムを用いて発現データを分析し、この場合、M値は、遺伝子発現レベルの安定性の測定として決定される。(J. Vandesompele, K. De Preterら, Genome Biol 3(7): Research 0034.1 - 0034.11, Epub 2002)。M値が低くなると、遺伝子はより安定になる。結果を表7に示す。M値は、PPIG、SLU7およびNUP214がこの試料セットにおいて最も安定な内在性対照であることを指示し、試験された試料全体で遺伝子発現の変化が最も小さかった。これらの3つの遺伝子の安定性は、138個の乳癌腫瘍FFPE試料で有効であった。結果を表8に示す。

【0130】

選択された200個の遺伝子、並びに3つのEC遺伝子の発現レベルは、142個のRNA試料においてプロファイルされた。遺伝子発現プロファイリングに関して、1工程の逆転写ポリメラーゼ連鎖反応(RT-PCR)を用いて相対的な定量化を行った。標的遺伝子の発現が参照の安定に発現される対照遺伝子のセットの発現と比較して評価される点で、定量化は「相対的」であった。SYBR(登録商標)Greenのインターカレーティング色素(Stragagene, La Jolla, Calif.)を用いて、リアルタイムPCR中に増幅産物を視覚化した。要約すると、反応ミックスは、抽出された試料RNAをcDNAに逆転写することを可能にした。次に、このcDNAは、後述する循環パラメータに従って、同じ反応管においてPCR増幅された。PCR条件は、表3に開示されている上流および下流のプライマーが、対象とする遺伝子の標的配列のそれぞれ5'および3'にハイブリダイズするのを可能にするように設計され、その後、これらのプライマーから伸長させ、ハイブリダイゼーションと伸長のサイクルを繰り返して、増幅産物を作製する。SYBR(登録商標)Greenの存在下でPCRを行った。この色素は、二本鎖DNAにインターカレートし、増幅産物を視覚できるようにする。RT-PCRは、Applied Biosystems Prism(登録商標)7900HT Sequence Detection System(Applied Biosystems, Foster City, CA)上で行い、生成物の間接的な測定として、インターカレートしたSYBR(登録商標)Greenの量を用いて、PCRを通じて周期的なサイクルで存在する増幅産物の量を検出した。(SYBR(登録商標)Greenの蛍光強度はDNAに結合すると100倍を超えて増大する。)PCRプライマーは、各遺伝子の全て知られているスプライス変異体を増幅するように設計され、その結果、全てのPCR生成物の大きさは、長さにして150塩基対より短くなり、分解して相対的に短い長さのRNAがFFPE試料に見出されるように期待される。本明細書に記載されている分子サインにおける14個の遺伝子、並びに3個の内在性対照遺伝子の増幅に用いられるプライマーは、表3に列挙されている。RT-PCR増幅は、2点測定で、384ウェル増幅プレートで行った。各ウェルは、15μlの反応ミックスを含んでいた。サイクルプロフィールは、50℃で2分、95℃で1分、60℃で30分、その後、95℃で15秒と60℃で30秒を45サイクルからなり、増幅産物の解離分析で終了する。PCR成分は、本質的には、L. Rogge, E. Bianchiら, 2000, Nat Genet 25: 96 - 101に記載されているとおりであった。

【0131】

遺伝子発現の相対的变化は、定量的PCRによって測定された。発現レベルは、 (C_t) 法によって計算され、ここで、 C_t は標的増幅に対する閾値サイクルであり；即ち、標的の時間の指数関数的増幅が開始するPCRのサイクル数である。(KJ Livak and TD Schmittgen, 2001, Methods 25:402-408)。対象とする遺伝子のmRNAの相対レベルは、下記：

$$(C_t) = (C_{t_{GOI}} - C_{t_{EC}})_{testRNA} - (C_{t_{GOI}} - C_{t_{EC}})_{refRNA}$$

(式中、GOIは対象とする遺伝子であり、testRNAは試料RNAであり、refRNAは校正参照RNAであり、ECは内在性対照である)

のように特定された。対象とする遺伝子ごとの発現レベルは、初めに、3つの内在性対照遺伝子に標準化された。3つの内在性対照($C_{t_{EC}}$)の平均を表す C_t を用いて、たった1つの対照遺伝子が用いられた場合に生じるであろう標準化バイアスのリスクを最小限にする。(T. Suzuki, PJ Higginsら, 2000, Biotechniques 29:332-337)。内在性対照を増幅するために用いられるプライマーは、表3に列挙されている。対象とする遺伝子の調整された発現レベルは、さらに、校正参照RNAプール、refRNA(ユニバーサルヒト参照RNA、Stratagene, La Jolla Calif.)に対して標準化された。これは、種々の機器から得られる発現結果を標準化するために用いられた。200個の遺伝子の発現プロファイリング実験で得られた (C_t) 値は、後述する統計学的分析に用いられ、本発明の14個の遺伝子予後サインを決定した。

14個の遺伝子サインの決定

元々の200個の遺伝子の発現プロファイリングからのデータ(即ち、上記で得られた (C_t) 値)を用いて、遠隔転移に対する生存時間を測定するための半教師あり主成分(supervised principal component: SPC)が、予後サインを含む遺伝子のリストを生じさせるために用いられた。(E. Bair and R. Tibshirani, 2004, PLoS Biology 2:0511-0522)。SPC計算は、R. Tibshirani, TJ Hastieら, 2002, PNAS 99:6567-6742の方法に従って、R. Tibshirani at Stanford University, Stanford, Calif.の研究室を経由したオンラインで利用可能であるPAMアプリケーションを用いて行われた。

【0132】

初めに、一変量のCox比例ハザードモデルを用いて、遠隔転移との遺伝子の関連性に従って、遺伝子をランク付けした。P値<0.05であるそれらの遺伝子が有意であると考えられた。Coxスコアのいずれかの切断に関して、SPCは、最適な閾値に到達した遺伝子の成分(即ち、主成分)を計算した：SPCは、最適な閾値を選択するためのCoxモデル(単一変数としての主成分を有する)と併せて内部の相互検証を用いた。正しいサインおよびスケールを決定するために、SPCによって得られた第一の主成分遺伝子リストが重要であり、一変量のCoxモデルにおける予測材料として用いられた。

【0133】

SPCによって生成される主成分遺伝子リストは、さらに、Lasso抑制法によって減らした。(R. Tibshirani, 1996, J Royal Statistical Soc B, 58:267-288)。Lasso抑制は、LARSアルゴリズムを用いて行われた。(B. Efron, T. Hastieら, 2004, Annals of Statistics 32:407-499; T. Hastie, R. Tibshiraniら, 編集, 2002, The Elements of Statistical Learning, Springer, New York)。LARSアルゴリズムに用いられた結果変数は、SPCによって選択された主成分であった。Lasso法は、予め特定された精度を用いて、このスコアを再生することができる遺伝子のサブセットを選択した。

転移スコア

転移スコア (MS) は下記の形式：

【 0 1 3 4 】

【 数 6 】

$$MS = a_0 + b * \sum_{i=1}^M a_i * G_i$$

を有する。G_i は、14個の遺伝子予後サインの各Lasso誘導の遺伝子(i)の標準的な発現レベルを表す。このG_i値は、上述される発現プロファイリングにおいて得られる (Ct) から全集団におけるその遺伝子の平均的遺伝子発現を差し引き、次に、その遺伝子の標準偏差で割ることによって計算される。定数a_iは、表2に列挙されている14個の遺伝子の第1の主成分への荷重である。各遺伝子iについてのa_iスコアは、表2に与えられている。定数bは-0.251であり、正しいサインおよびスケールリングを得るために、予測材料として主成分を含む一変量のCoxモデル由来であった。定数a₀は0.022である；これは、中央値がゼロになるようにMSを中心にする。Mは、成分リストにおける遺伝子数である；このケースでは14である。したがって、MSは表2に開示されている14個の遺伝子についての、一定量の発現レベルの総和であり、各々は、特定の定数a_iによって乗じる；次に、この総和は、定数bによって乗じられ、中心の定数0.022にこの総和を加算した。

【 0 1 3 5 】

このスコアは、本明細書においては、MS(全て)と呼ばれ、それは、試験された142人全てのER陽性個体の分析に基づくためである。いずれかの新しい試料は、その患者についての14個の遺伝子発現プロファイリングデータから、場合により、このスコアからこの転移スコアを生じることによって評価されてもよく、患者に対する遠隔転移の可能性を測定することができる。

MSからの遠隔転移可能性の発生

いずれかの個体の患者に関する遠隔転移の可能性は、可変時間点でのMSから計算することができ、基準生存関数としてWeibull分布を用いる。

【 0 1 3 6 】

14個の遺伝子サインの発現プロファイリングからの、上記で得られた転移スコア(MS)は、Cox比例ハザードモデルによって遠隔転移の可能性に転換された。Coxモデルは基準ハザード関数を特定しないため、ハザードおよび生存関数は、初めに、パラメータの回帰モデルを介して構築された。パラメータの回帰モデルでは、遠隔転移なしの生存時間は結果であり、転移スコア(MS)は独立した可変インプットとであった。イベント時間は、Weibull分布を有するように推定された；その2つのパラメータは、MSが誘導される生存データを用いて推定された。ある患者に対して、ある時間内の遠隔転移の可能性を計算するために、MS値は、単に、生存関数に対する式に置換される。

事前検証

この研究では、142個の試験患者が、無作為に10のサブセットに分けられた。1つのサブセットが取り除かれ、全SPC手法が残りの9個のサブセットの統合に対して行われた。遺伝子が選択され、9個のサブセットに構築された予後材料が、残りのサブセットに対して、相互検証転移スコアのMS(CV)を得るように適用された。この相互検証手法は、MS(CV)が全ての患者について記入されるまで、10回行われた。このようにしてMS(CV)を構築することによって、各々10分の1の一片は、直接的には、対応する生存する時間を使用せず、即ち、教師なしと考えることができる。

【 0 1 3 7 】

これは、試料における全ての個体に対して誘導された変数をもたらし、次に、その性能について試験され、他の臨床変数と比較することができた。しかしながら、MS(全て)は、試験された全ての142個の個体に構築され、同じように試験された場合、相当なバイアスを生成した。

【 0 1 3 8 】

R O C 曲線下の時間に依存した面積 (A U C) が計算されると、M S (C V) を用いて、14 個の遺伝子予後サインの精度を計算した (後述される)。また、M S (C V) は、14 個の遺伝子サインが臨床予測材料と組み合わせられると、C o x 抑制モデルに用いられた。M S (C V) には、通常の (事前検証されていない) 予測材料とは対照的に、1 つの自由度がなければならない。事前検証されていない予測材料は、より多くの自由度を有している。

M S の統計学的分析

遠隔転移なしの生存、並びに全生存についての Kaplan - M a y e r (K M) 曲線は、カットポイントとして M S (C V) の中央値 (即ち、M S (C V) の 5 0 パーセンタイル) を用いて、高リスクおよび低リスク患者群について作成された。カットポイントとしての中央値 M S の選択は、M S の 1 0 パーセンタイルから M S の 9 0 パーセンタイルまでの異なるカットポイントを用いて、高リスク対低リスク群のハザード率の計算に基づいた。高リスクおよび低リスクの個体の平衡数、並びに最大に近いハザード率は、カットポイントとして中央値を選択するための決定因子であった。

10

【 0 1 3 9 】

中央値 M S のカットポイントを用いたハザード率 (H R) および 9 5 % の信頼区間 (C I) が計算され、報告された。ログランク試験を行い、ハザード率を異なるカットポイントについて計算した。5 年以内での遠隔転移の予測における 14 個の遺伝子サインの精度および値は、種々の手段によって評価された。 (X H Z h o u , N . O b u c h o w s k i ら , 編集 , 2 0 0 2 , S t a t i s t i c a l M e t h o d s i n D i a g n o s t i c M e d i c i n e , W i l e y - I n t e r s c i e n c e , N e w Y o r k) 。

20

【 0 1 4 0 】

一変量および多変量の C o x 比例ハザード抑制は、年齢、腫瘍サイズ、腫瘍悪性度、および 14 個の遺伝子サインを用いて行われた。サインに関する臨床サブグループ分析も行われた。統計学的分析は、本明細書に記されている統計パッケージを除いて、S A S (登録商標) 9 . 1 統計学的ソフトウェア (S A S I n s t i t u t e , I n c . , C a r y , N . C .) を用いて行われた。

多重遺伝子サイン

30

研究された 2 0 0 個の候補遺伝子のうち、44 個は、一変量の C o x 比例ハザード抑制において、調整されていない P 値 < 0 . 0 5 であった。僅かな手に予後を有する患者は、37 個の遺伝子のアップレギュレーションを示し、7 個の遺伝子はダウンレギュレーションを示した。P A M における半教師あり主成分手法 (S P C) により 38 個の予後材料が得られた。この遺伝子リストは、さらに、L A R S アルゴリズムを介して、L a s s o 抑制を用いることによって、14 (表 2) まで減らされた。表 2 は、各遺伝子の記述を与える。種々のカットポイントでのハザード率 (H R) (パーセンタイル M S (C V)) を計算した。M S (C V) の中央値は、患者を低リスク群および高リスク群に分類するために選択された。

結果

40

低リスク群および高リスク群の遠隔転移なしの生存率並びに全生存率

トレーニングセットにおいて、相互検証した転移スコア、M S (C V) の中央値によって定義されるように、71 個の低リスクおよび 71 個の高リスク患者において 7 および 24 個の遠隔転移があった。Kaplan - M a y e r 推定 (図 1 a) は、ログランクの p 値が 0 . 0 0 0 2 8 である 2 つの群の間の遠隔転移なしの生存 (D M F S) における有意差を指示した。低リスク群の 5 年および 10 年の D M F S 比 (標準誤差) は、それぞれ 0 . 9 6 (0 . 0 2 5) および 0 . 9 0 (0 . 0 3 7) であった。高リスク群については、対応する比は、0 . 7 4 (0 . 0 5 3) および 0 . 6 2 (0 . 0 6 6) であった。全生存 (図 1 b) については、また、2 つの群間に有意差があった (ログランクの p 値 = 0 . 0 0 4 8)。低リスク群の 5 年および 10 年の O S 率 (標準誤差) は、0 . 9 0 (0 . 0 3 6) お

50

よび 0.78 (0.059) であり、対応する比は、高リスク群においては 0.79 (0.049) および 0.048 (0.070) であった (表 5)。

一変量および多変量の Cox 抑制モデルからのハザード率

DMFS を予測するための MS による高リスク群対低リスク群の未調整ハザード率は、一変量の Cox 抑制分析によって指示されるように、4.23 (95% CI = 1.82 ~ 9.85) であった。比較すると、腫瘍悪性度 (中 + 高悪性度対低悪性度) による高リスク群対低リスク群は、未調整のハザード率が 2.18 (1.04 ~ 4.59) であった。腫瘍サイズは、DMFS の予測において有意 ($p = 0.05$) であり、直径にして 1 cm 増加あたりハザードにして 7% の増加であるが、年齢は、この患者セットでは有意な因子ではなかった。多変量の Cox 抑制分析では、14 個の分子サインのリスク群は、ハザード率が 3.26 (1.26 ~ 8.38) であり、これは、外科的処置での年齢、腫瘍サイズおよび悪性度によって調整された。それは、多変量分析において、唯一の有意なリスク因子 ($p = 0.014$) であった。

10

診断精度および予測値

5 年以内の遠隔転移を予測するための 14 個の有意なリスク群の診断精度および予測値を表 9 に要約した。感度は、5 年以内に遠隔転移があった患者に対しては 0.86 であり、特異性は、そうでない患者に対しては 0.57 であった。陰性予測値 (NPV) は 96% であり、ほんの 4% の個体が 5 年以内に遠隔転移であったことを指示し、その場合、その個体が低リスク群にいたことをその遺伝子サインは指示した。それにもかかわらず、陽性適中率 (PPV) はほんの 26% であり、ほんの 26% の個体が遠隔転移を発症したことを指示し、分子サインは、その個体が高リスクにあることを指示した。高 NPV および低 PPV は、部分的には、5 年以内の遠隔転移の低い有病率に起因し、それは、現在の患者セットでは 0.15 であると推定された。

20

【0141】

さらに、5 年以内の遠隔転移を予測するための連続的な MS (CV) の受信者動作特性 (ROC) 曲線を図 2 に示した。AUC は、5 年以内の遠隔転移の予測については 0.76 (0.65 ~ 0.87) であった。10 年以内の死亡予測に対する AUC は、0.61 (0.49 ~ 0.73) であった。0.5 よりも大きい AUC についての片側試験は有意であり、転移および死亡の評価項目に対しては、対応する p 値は < 0.0001 および 0.04 であった。

30

検討

経路分析は、本明細書に開示されている予後サインにおける 14 個の遺伝子が、多様な生物学的機能に関与していることを表すが、大多数の遺伝子は、細胞増殖と関連している。14 個の遺伝子のうち 11 個は、不良転帰へと導く腫瘍において、協調的に過剰発現されることが見出された TP53 および TNF シグナル経路と関連している。BUB1、CCNB1、MYBL2、PKMYT1、PRR11 および ORC6L は、細胞周期に関連した遺伝子である。DIAPH は、アクチン細胞骨格および生物発生に関与する遺伝子である。DC13 は、シトクロムオキシダーゼの集合と関連することが期待されている。

【0142】

従来報告された研究は、RNA の供給源として凍結組織の使用に制限され、マイクロアレイ上にプロファイルされたが、本実施例に記載の発明は、FFPE 腫瘍試料における遺伝子プロファイリングに対して、リアルタイム RT-PCR を用いてもよいことを示す。このようにして、アジュバント療法の日常的な使用に先立って拡張された結果データを有する患者由来の保存された乳癌患者切片の使用を提供する。

40

【0143】

遠隔転移なしの生存は、本実施例において記載される研究のための予後評価項目である。教師あり主成分 (SPC) 法を用いて、本発明の 14 個の遺伝子サインパネルを構築した。このサインの集合において使用されるアプローチは、個体の発現プロフィールを、任意の所定時間、遠隔転移のリスクの測定に言い換えることができる転移スコア (MS) の誘導を可能にする。任意の期間、転移のリスクを定量する能力は、処置の決定における患

50

者および臨床医の間の非常に柔軟性のある予後情報を与え、これは、患者の間でリスク寛容および計画対象期間が変化するためである。

【0144】

14個の遺伝子の分子サインを用いて、高および低リスク群は、遠隔転移なしの生存率、並びに全生存率において有意差があった。このサインは、乳癌予後において日常的ではなく試験された増殖遺伝子を含む。14個の遺伝子サインは、H. Dai, L.J. van 't Veerら(2005, Cancer Res 15:4059-4066)によって記述される50個の遺伝子サインとは10個の遺伝子の重複がある。対照的に、ほんの6個の遺伝子は、Dai, van 't Veerら(2002, Nature 415:530-536)によって記述される70個の遺伝子サインと重複する。これは、ER陽性患者とER陰性患者の両方を含む、その研究が患者のより異質な群を分析したという事実によって説明することができる。本明細書に記載されているサインは、S.P. Paik, S. Shakら,(2004, N Engl J Med 351:2817-2826)によって記述される16個のサインと重複する2つの増幅遺伝子を有していた。

10

【0145】

本明細書に記載されている分子サインは、多変量分析から指示されるように、年齢、腫瘍サイズおよび悪性度などの伝統的なリスク因子より、独立した予後値を有する。このサインは、腫瘍悪性度が低い場合、さらにより有力な予後の測定を提供する。Daiらによって報告されているように、低い悪性度の腫瘍を有するこの患者群のサブセットは、従来推定されたものよりも実に高い転移のリスクにある可能性がある。(Daiら, 2005, Cancer Res 15:4059-4066)。また、本明細書に記載されているサインは、年齢にして55歳未満の女性からの試料を最初に用いたvan 't Veerら(2002, Nature 415:530-536)およびDaiら(2005, Cancer Res 15:4059-4066)によって初期に報告された予後遺伝子の信頼を広げている。それは、このサインが、乳癌患者の一般的な範囲に類似している広い年齢分布(中央値64歳)の患者において確認された。

20

【0146】

FFPE組織の試料への使用は、凍結した組織を用いるマイクロアレイよりもさらに少量の切片でさえ、異なる検出プラットフォーム(定量的PCR対マイクロアレイ)で一部の遺伝子を確認した。結果のこの反復は、転移能が腫瘍に含まれる少数の細胞よりはむしろ、その大部分の細胞の固有の特徴であるという、BernardsおよびWeinbergによって記述された概念と一致している。(R. Bernards and R.A. Weinberg, 2002, Nature 418:823)。

30

【0147】

また、本明細書に記載されている発明は、TibshiraniおよびEfronによって提唱された事前検証技術を用いることによる予後材料性能の客観的評価を提供する。(R. Tibshirani, B. Efron, 2002, Statistical Applications in Genetics and Molecular Biology 1:article 1)。研究者の中には、次第に巨大になり、特徴付けられたデータベースにおいて独立した検証の重要性に気付くものもあった。(R. Simon, J. Clin Oncol, 2005, 23:7332-7341; D.F. Ransohoff, 2004, Nat Rev Cancer 4:309-14; D.F. Hayes, B. Trockら, 1998, Breast Cancer Res 52:305-319; D.G. Airman and P. Royston, 2000, Stat Med 19:453-73)。

40

【0148】

本発明では、従来のマーカーおよびツールに明確な情報を与え、全身処置と混同されない独自の14個の遺伝子予後サインが提供される。このサインは、初期の段階のリンパ節転移陰性のER陽性患者について、FFPE切片およびRT-PCRを用いて生成されるが、それを当該技術分野において知られている任意の方法と併せて用い、そのサインにお

50

ける遺伝子のmRNA発現、並びに、限定されないが、FFPE切片、凍結された組織試料および新鮮な腫瘍生検を含む任意の腫瘍組織供給源から得られたmRNAを測定してもよい。本発明の14個の遺伝子サインのmRNA発現レベルに基づいて、転移スコアは、いずれかの個体の乳癌患者に対して遠隔転移を定量するために計算され得る。このようにして、本明細書に開示されている本発明は、いずれかの期間でER陽性の乳癌患者の日常的な臨床検査室試験における使用を可能にする。

(実施例2)

14個の遺伝子サインは280個のFFPE試料を用いた未処理のリンパ節転移陰性のER陽性乳癌患者における遠隔転移を予測する

全身処置を受けていない独立した試料セットにおける、リンパ節転移陰性(N-)、エストロゲン受容体陽性(ER+)の乳癌患者の遠隔転移を予測することができる14個の遺伝子発現サインを確認する努力がなされた。14個の遺伝子サインの有効性を評価するために用いられた、実施例1の実験プロトコルおよび統計学的分析について述べる。

患者 & 方法

ガイ病院の乳房組織およびデータバンクの遡及検索は、初期の乳癌であると診断され、明確な局所療法(乳房温存療法または乳腺切除)を有しているが、追加のアジュバント全身処置を受けていない患者の集団を特定するためになされた。この研究グループは、1975年~2001年に、臨床的な腫瘍サイズが3cm以下であり、病理学的に腋窩のリンパ節転移陰性のER陽性腫瘍を有し、5年フォローアップよりも長い女性に制限された。総数412人の患者が特定され、それらは、RNA抽出に使用可能である十分にホルマリン固定された、パラフィン包埋(FFPE)組織を有してもいた。本研究のために患者の資料およびデータの使用は、ガイ研究倫理委員会によって承認された。

【0149】

このグループから、300人の患者由来の腫瘍をプロファイルするには十分なmRNAの量と質があった。その後、さらに20ケースが、この試験から排除された。6人の患者は、遠隔転移前、両側に乳癌であり、6人は、遺伝子発現レベルでは欠損値であり、8つの腫瘍は、現在の技術を用いた再評価に基づいてER陰性であることがわかった。このようにして、総数280人の患者が分析に含められた(表1のガイ病院からの検証セット)。選択バイアスを評価するために、本発明者らは、試験対象患者基準を満足することが確認され、RNA抽出に十分なFFPE組織を有していた412人の患者を用いて分析した。2つのセット間で年齢、腫瘍サイズおよび組織学的悪性度における有意差がなかったため、選択バイアスを検出しなかった。

【0150】

このグループの患者におけるER状態は、同時期のIHCアッセイを用いて再評価された。Allredスコアが3以上は、受容体陽性と考えられた。腫瘍は、WHOガイドライン(World Health Organization, Geneva, Switzerland. Histological typing of breast tumours. Tumori 1982; 68: 181)に従って分類され、組織学的悪性度は、改変されたBloomおよびRichardson法(Elston CW & Ellis 10. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long term follow-up. Histopathology 1991; 19: 403-10)を用いて確かめられた。

【0151】

トレーニングセットと比較すると、検証セットは若年者であり、大きな腫瘍を有し、その大部分が高悪性度の腫瘍であった(表1)。それらの特徴における差異に関する試験は非常に有意である($p < 0.001$)。

転移スコア

10

20

30

40

50

実施例 1 のトレーニング試料において誘導された等式 1 の M S は、ガイ病院からの未処置患者に適用された。実施例 1 では、R N A は濃縮されたが、ガイ病院からの未処置患者の場合には、R N A 試料は濃縮されなかった。

【 0 1 5 2 】

濃縮していない試料に M S を適用するために、濃縮された試料と濃縮されていない試料との間の変換因子は、分子サインにおける遺伝子の各々に対して、実施例 1 の 9 3 個のトレーニング試料から得られた。

結果

低リスク群および高リスク群の遠隔転移なしの生存率および全生存率

それぞれ、71 人の低リスク患者においては 4 (5 . 6 %) の遠隔転移があり、209 人の M S の高リスク患者においては 6 2 (2 9 . 7 %) の遠隔転移があった。カプラン - マイヤー推定 (図 3 a) は、ログランクの p 値が $6 . 0 2 e - 5$ である 2 つの群の間の D M F S における有意差を指示した。低リスク群の 5 年および 1 0 年の D M F S 比 (標準誤差) は、それぞれ 0 . 9 9 (0 . 0 1 4) および 0 . 9 6 (0 . 0 2 5) であった。高リスク群については、対応する生存率は、0 . 8 6 (0 . 0 2 5) および 0 . 7 6 (0 . 0 3 1) であった (表 1 2) 。

【 0 1 5 3 】

アジュバント！リスク群については、D M F S のカプラン - マイヤー曲線を図 3 c に示した。5 年および 1 0 年の D M F S 比 (標準誤差) は、それぞれ、低リスク群については 0 . 9 6 (0 . 0 2 3) および 0 . 9 3 (0 . 0 3) であり、高リスク群については 0 . 8 7 (0 . 0 3) および 0 . 7 7 (0 . 0 3 1) であった。アジュバント！によって特定されるものよりも、M S によって特定される高リスク群と低リスク群との間の生存率において大きな差異があった。

【 0 1 5 4 】

全生存 (O S) に関して、カプラン - マイヤー曲線 (図 3 b) は、M S 低リスク群と高リスク群との間の O S 比に有意差を指示した (ログランクの p 値 = $0 . 0 0 0 2 8$) 。5 年および 1 0 年の O S 比は、それぞれ、低リスク群については 0 . 9 7 (0 . 0 2 0) および 0 . 9 4 (0 . 0 2 8) であり、高リスク群については 0 . 9 2 (0 . 0 1 9) および 0 . 7 1 (0 . 0 3 2) であった。図 3 d は、5 年および 1 0 年の全生存を予測するためのアジュバント！についてのカプラン - マイヤーを示した。M S およびアジュバント！は、全生存に関する類似の予後情報を提供する。

一変量および多変量の C o x 抑制モデルからのハザード率

遠隔転移までの時間を予測するための M S による高リスク群対低リスク群の未調整のハザード率は 6 . 1 2 (9 5 % C I = 2 . 2 3 ~ 1 6 . 8 3) であった (表 1 0) 。M S リスク群に対する未調整のハザード率は、アジュバント！、年齢、腫瘍サイズおよび組織学的悪性度によって特定される群のハザード率よりも高い。アジュバントは、第 2 の最大のハザード率が 2 . 6 3 であった (9 5 % C I 1 . 3 0 ~ 5 . 3 2) 。組織学的悪性度および腫瘍サイズによるリスク群は、D M F S の予測に有意であり、年齢群では有意でなかった。

【 0 1 5 5 】

年齢群は、O S の予測に最も有意な予後因子であり、未調整のハザード率は 2 . 9 (9 5 % C I 2 . 0 3 ~ 4 . 1 8) であった (表 1 0) 。それにもかかわらず、M S リスク群は、全生存を予測することができ、H R は 2 . 4 9 であった。

【 0 1 5 6 】

M S リスク群、年齢、腫瘍サイズおよび組織学的悪性度の臨床病理学的なリスク因子による遠隔転移までの時間の多変量の C o x 抑制 (表 1 1 a) では、年齢、腫瘍サイズおよび組織学的悪性度によって調整された M S リスク群のハザード率は 4 . 8 1 (1 . 7 1 ~ 1 3 . 5 3 , p = 0 . 0 0 3) である。M S リスク群は、多変量分析における唯一の有意なリスク因子であった。したがって、遺伝子サインは、伝統的な臨床病理学的なリスク因子と比較して D M F S に対して独立した予後値を有し、これらの因子内で情報の一部を捕

10

20

30

40

50

捉する。

【0157】

14個の遺伝子サインによるリスク群、アジュバント！によるリスク群を用いた遠隔転移までの時間の多変量Cox抑制(表11b)では、対応する調節されたハザード率は、5.32(1.92~14.73)および2.06(1.02~4.19)であった。MSとアジュバント！のリスク群は、DMFSを予測するために有意のままであった。これは、MSおよびアジュバント！が、互いに非常に独立した相補的な予後情報を担持することを指示する。

異なる臨床下位群における分子サインの性能

表13は、遺伝子サインが若年および老齢の閉経前および閉経後の女性における遠隔転移を予測することを示す。小さいサイズの腫瘍を有する患者では非常に予後的であるが(HR=14.16、 $p=0.009$)、2cmより大きな腫瘍を有する患者では有意でない。低悪性度の下位群におけるハザード率(HR=7.6)は、高悪性度(HR=4.6)よりも高いが、それは、小さな試料サイズ(60試料中、7イベント)のため、低い悪性度の下位群における有意性($p=0.06$)の傾向を示すだけである。

【0158】

種々のサブグループにおけるハザード率は、検定試料セットにおける遺伝子サインが低い悪性度、小サイズの腫瘍、若年であり閉経前の患者においてより予後的である。MSリスク群と臨床的変数との間の相互作用に関する公式な試験は有意ではなかった。しかしながら、このサインも、CPMCトレーニングセットにおける低悪性度の腫瘍においてより予後的であった。それにもかかわらず、相互作用分析は、多重試験が行われるため、予備的であると見なされなければならない。

診断精度および予測値

10年以内の遠隔転移を予測するためのMAおよびアジュバント！によるリスク群の診断精度および予測値を表4に示した。MSリスク群は、アジュバント！のリスク群の0.90(0.78~0.96)よりも感度が0.94(0.84~0.98)と高いが、特異性は類似している(MSについては0.3(0.24~0.37)対アジュバント！については0.31(0.26~0.38))。10年以内の遠隔転移の推測された有病率を用いると、MSリスク群に対するPPVおよびNPVは、それぞれ0.23(0.21~0.25)および0.97(0.88~0.99)であった。対応する値は、アジュバント！リスク群に対しては0.23(0.20~0.25)および0.93(0.85~0.97)であった。したがって、MSは、アジュバント！よりも10年以内に遠隔転移がないものを僅かに良好に予測することができ、10年以内の遠隔転移を有するものに対する予測値は、分子および臨床予後材料に対して同じであった。

【0159】

5年および10年以内の遠隔転移を予測するための連続的MSのROC曲線(図4)は、AUC(95%CI)が0.73(0.65~0.81)、0.70(0.63~0.78)であった。10年以内の死亡を予測するROC曲線は、AUCが0.68(0.61~0.78)であった。したがって、MSは、遠隔転移および死亡を予測する。

【0160】

対照的に、アジュバント！による5年以内、10年以内の遠隔転移、10年以内の死亡を予測するためのROC曲線のAUCは、0.63(0.53~0.72)、0.65(0.57~0.73)、および0.63(0.56~0.71)であり、それらは、MSによる対応する値よりも低かった。

遠隔転移の可能性の連続的予測材料としてのMS

図5は、転移スコアMSを用いて個体の患者についての5および10年での遠隔転移の可能性を示す。5年および10年の遠隔転移の可能性は、中央値(最小~最大)が8.2%(1.4%~31.2%)および15.2%(2.7%~50.9%)であった。リスク群を定義するためのカットポイントであるゼロMSでは、5年および10年遠隔転移可能性は、それぞれ5%および10%であった。

【 0 1 6 1 】

MS による 10 以内の遠隔転移の可能性は、アジュバント！による 10 年以内の再発の可能性と比較された（図 6）。決定係数（ R^2 ）は、0.15 であり、MS による遠隔転移の可能性におけるほんのごく一部の変動がアジュバント！によって説明可能であるところを指示した。MS による遠隔転移の可能性は、アジュバント！による再発の可能性よりも低かった。それは、全ての再発イベントがアジュバント！再発可能性に含まれたためであり、遠隔転移だけが遠隔転移の可能性の MS 推測におけるイベントとしてカウントされた。

検討

初めに、14 個の遺伝子の予後サインは、リンパ節転移陰性の初期段階の未処置乳癌トレーニングセットにおいて、遠隔転移のための定量的 RT-PCR を用いて、FFPE 切片からの mRNA 発現に基づいて開発された。得られたサインは、異なる期間で個体に関するリスクを定量する転移スコア（MS）を生じさせるために用いられ、高リスクおよび低リスクに試料セットを二分するために用いられた。最近の「事前検証」の統計学的な技術を用いたトレーニングセットの初期の内部検証後、本発明者らは、類似の独立した検証集団におけるトレーニングセットの正確に二分された切断を用いて、発現サインを検証した。トレーニングセットおよび検証セットにおけるサインの性能特徴は同じであった。一変量および多変量のハザード率は、DMFS を予測するために、それぞれ、検証セットについては 6.12 および 4.81 であり、トレーニングセットについては 4.23 および 3.26 であった。多変量分析では、転移スコアだけが、他の臨床病理学的因子の腫瘍サイズのみについて有意である傾向を伴って、有意のままであった。また、14 個の遺伝子の予後サインは、全生存を予測することができ、検証セットでは一変量ハザード率は 2.49 であった。5 年および 10 年以内の遠隔転移を予測し、10 年以内の死亡を予測するための連続 MS の ROC 曲線は、AUC がそれぞれ 0.73、0.70 および 0.68 であった。サインは、腫瘍悪性度が低い場合にはより有力な予後を提供した（ハザード率は低悪性度では 7.58 であり、高悪性度の腫瘍では 4.59 であった）。例えば、Dai らは、患者の連続体の反射ではなく、むしろ患者の別個の群における差異性能であるものとして、分類の予後力の変化を解釈した。

【 0 1 6 2 】

アジュバント！から計算されたリスクと比較した場合、ウェブに基づく決定支援では、MS との適度な相関だけがあった。多変量の Cox 抑制では、MS およびアジュバント！リスク群はともに、互いに調製されると、優位な予後因子を維持する。アジュバント！と比較した MS によって特定される高リスク群と低リスク群との間の生存率には大きな差異があった。MA は、アジュバント！よりも 10 年以内の遠隔転移を持たないものを良好に予測することができた。これらのデータは、分子サインが、乳癌の遠隔転移を予測するのに、日常的であるかまたはより最近である予後ツールに独立した情報を提供することを示す。

（実施例 3）

14 個の遺伝子サインは、96 個の FFPE 試料を用いて、処理および未処理のリンパ節転移陰性および ER 陽性乳癌患者における遠隔転移を予測する

処置および未処置の患者を有する独立した試料セットにおける、リンパ節転移陰性（N-）、エストロゲン受容体陽性（ER+）の乳癌患者における遠隔転移を予測することができる 14 個の遺伝子発現サインを確認する努力がなされた。14 個の遺伝子サインの有効性を分析および評価するために用いられた、実施例 1 の実験プロトコールおよび統計学的分析について述べる。

患者 & 方法

平均年齢が 56.7 歳である 96 人の N-、ER+ 乳癌患者の集団は、検証研究のために選択された（表 15）。検証研究における患者は、ミュンスター大学から選択された。それらのうち、15 人は未処置であり、54 人はタモキシフェン単独で処置され、6 人は化学療法だけで処置され、2 人はタモキシフェンと化学療法を用いて処置された。19

人の患者は、未知の処置状態であった。サインにおける 14 個の遺伝子は、定量的 RT-PCR を用いて FFPE 試料においてプロファイルされた。以前に誘導された転移スコア (MS) は、遺伝子発現レベルからの検証セットのために計算された。患者は、ゼロであった所定の MS カットポイントを用いて、2 つの群に層をなしていた。

転移スコアの検証

実施例 1 の試料を用いて誘導された等式 1 の MS は、ミュエンスター大学からの患者に適用された。この実施例では、また、腫瘍組織由来の RNA は濃縮されたが、実施例 1 のものよりは程度が小さかった。この試料に MS を適用するために、濃縮された試料と濃縮されていない試料との間の転換因子が、サインにおける 14 個の遺伝子の各々について、ミュエンスター大学からの 93 個の試料から得られた。また、濃縮された試料と濃縮されていない試料との間の転換因子は、実施例 1 の CPMC から 93 個のトレーニング試料から得られた。次に、CPMC とミュエンスター大学の遺伝子発現レベルとの間の転換因子は、転換因子のそれらの 2 つのセットを用いて計算された。

10

結果

遠隔転移なしの生存率率

カットポイントとして MS ゼロを用いて、患者は、高リスクと低リスクとして分類された。96 人全ての患者のうち、48 人の患者は、高リスクとして同定され、5 年の DMF 生存率 (標準誤差) は 0.61 (0.072) であり、48 人の低リスク患者は、対応する生存率が 0.88 (0.052) であった。62 人の処置された患者のうち、32 人の高リスクおよび 30 人の低リスク患者では、5 年の DMF 生存率は、それぞれ 0.66 (0.084) および 0.89 (0.060) であった。タモキシフェン処置単独を受けた 54 人の患者のうち、26 人の高リスク患者および 28 人の低リスク患者は、5 年の DMF 生存率は、それぞれ 0.65 (0.094) および 0.88 (0.065) であった (表 16)。

20

未調整のハザード率

全集団について、MS は、遠隔転移なしの (DMF) 生存と相関した。Cox 比例ハザード抑制は、MS の単位増加あたりの 2.67 (1.28 ~ 5.57) 倍に増加したハザードを指示した ($p = 0.0087$)。カットポイントとしてゼロを用いると、全集団における高リスクと低リスク患者のハザード率は、2.65 (1.16 - 6.06, $p = 0.021$) であった。61 人の処置された患者については、ハザード率は 3.08 (0.99 ~ 9.56, $p = 0.052$) であった。タモキシフェン単独で処置された 54 人の患者については、ハザード率は、2.93 (0.92 ~ 9.35, $p = 0.07$)。異なる群の患者についての生存率およびハザード率は、それぞれ、表 16 および図に要約されている。

30

【0163】

元々は未処置の患者由来である、RT-PCR に基づく 14 個の遺伝子サインは、FFPE 組織を用いた独立した試料セットにおいて、N-、ER+、タモキシフェン処置された患者の遠隔転移を予測することができる。高リスクおよび低リスクのタモキシフェンを単独で処理された患者の大きさ差異 DMFS 比 (0.65 対 0.88) があり、その場合、2 つの群は、カットポイントとしてゼロを用いた MS によって特定された。多様な MS の上位および下位の五分位数における差異リスクは、全ての患者、タモキシフェンを単独で処理された患者に対して、それぞれ 3.99 および 3.75 倍であった。予後サインは、全身処置と混同されない基準リスクを提供することができる。さらに、処置を受ける患者についての転移リスクを予測することができる。したがって、遺伝子サインは、その後の療法とは独立して、治療法の決定を導くための乏しい臨床転帰を用いて女性を特定するのに適切である。

40

(実施例 4)

14 個の遺伝子サインは、205 個の FFPE 試料を用いた、タモキシフェン処置のリンパ節転移陰性の ER 陽性乳癌患者における遠隔転移を予測する

患者

50

ガイ病院で1975年から2001年に外科的処置を受けたN-、ER+の乳癌を有する205人の女性の集団を選択した。フォローアップの中央値は9.3年である。そのうち、17人(8.9%)が遠隔転移であり、44人(21.5%)が死亡し、17人(8.9%)が局所および遠隔再発であった。

【0164】

138人(67.3%)の患者は、ステージIであり、67人(32.7%)がステージIIであった。全ての患者は、アジュバントのホルモン療法を受けたが、化学療法は受けなかった。この集団では、平均(SD)年齢は59.3(10.4)歳であった。64%は、55歳を超え、80.5%が閉経後であった。全ての腫瘍は、直径にして3cmであり、平均(SD)の腫瘍直径は1.67cm(1.0)であった。60人(29.3%)、98人(47.8%)、47人(22.9%)の患者は、腫瘍の組織学的悪性がそれぞれ1、2、3であった。(表17)

10

エンドポイント

本発明者らは、初期カットポイントとして、外科的処置から、遠隔転移なしの生存(DMFS)とも呼ばれる、遠隔転移までの時間を選択する。イベントは遠隔転移である。反対側の再発および再発なしの死亡が検査イベントであり、局所再発は、考慮されるべきイベントおよび検査イベントではなかった。DMFSカットポイント、そのイベントおよび検査ルールの定義は、予後分子マーカー研究(Paikら 2004)に関する全米外科アジュバント胸部および腸管プロジェクト(NSABP)によって採用された定義を用いて調整された。DMFSカットポイントは、癌に関連した死亡に最も直接的に連結されている。

20

遺伝子発現サインおよび転移スコア(MS)

14個の遺伝子サインは、従来、実施例1に記載されるカリフォルニア・パシフィック・メディカル・センター(CPMC)からのFFPE試料を用いて、RT-PCRによるプロファイリング研究を用いて開発された。インジェニユイティプログラムによる経路分析は、そのサインの14個の遺伝子のうち多数が細胞増殖と関連することを示した。14個の遺伝子のうち10個が、乏しい結果の腫瘍において協調的に過剰発現することが見出されたTP53シグナル経路と関連している。

【0165】

転移スコア(MS)は、各個体について計算された。この実施例におけるMSは、以前に記載されたサインのうち14個の遺伝子の遺伝子発現に基づいていた。しかしながら、本実施例におけるMSを計算するためのアルゴリズムは、従来の実施例について記載された実施例とは異なっていた。それにもかかわらず、新しいアルゴリズムを用いて誘導されたMSは、従来の方法を用いて誘導されたMSと非常に相関し、ピアソン相関係数は>0.99であった。さらに、本実施例では、以前の3つの実施例において低および高リスク群を分類するために、たった1つのカットポイントを用いたこととは対照的に、高、中くらい、低のMS群を分類するために、2つのカットポイントを使用した。

30

【0166】

サインの14個の遺伝子は、実施例1に記載される研究において選択され、新しいMSスコアおよびカットポイントは、実施例2に記載されるガイ病院からの未処置試料を用いて、この研究に基づいて測定された。新しいMSアルゴリズムは、実施例4および5で適用され、検証された。そこで、新しい転移スコア(MS(新規))が、14個の遺伝子の遺伝子発現レベルの平均の負数として計算される。この新しいスコアを用いて、14個の遺伝子は、同計量を与えた。高いMSが遠隔転移の高リスクに対応するように、-1の乗数を用いた。新しいMSは、下記の式：

40

【0167】

【数 7】

$$MS(\text{新規}) = -(1/14) * \left[\sum_{i=1}^{14} G_i \right] \quad \text{等式 3}$$

(式中、 G_i はサインにおける 14 個の遺伝子の中心の発現レベルである) において表すことができる。

【0168】

MS (新規) の 2 つのカットポイントは、高、中、および低 MS 群に患者を分類するために選択された。低いカットポイントは - 1 . 47 であり、上のカットポイントは - 0 . 843 であった。 - 1 . 47 よりも小さな MS を有する個体は、低 MS 群であった。 - 1 . 47 と - 0 . 843 との間の MS を有する個体は、低リスクであると考えられ、 - 0 . 843 よりも高い MS を有する個体は高 MS による 10 年以内群であった。低い MS を有するものは、低リスクであると考えられ、中間の MS および高い MS 群にあるものは、ガ이의未処理試料の高リスクとして考えられる場合 (換言すれば、 - 1 . 47 を超える MS を有するものは高リスクであると考えられる)、MS リスク群の感受性は 90 % を超える。一方、低 MS および中間 MS は、低リスクであると考えられ、高い MS を有するものは高リスクであると考えられる場合 (換言すれば、 - 0 . 843 よりも低い MS を有するものは、低リスクであると考えられ、 - 0 . 843 よりも高い MS を有するものは、高リスクであると考えられた)、ガ이의未処理試料における MS リスク群の感受性および特異性は、70 % 程度である。

【0169】

未処置の試料については、中間の MS 群は、高い MS 群のものと類似したリスクを有し、高い MS 群および中間の MS 群は、低 MS を有するものよりも高かった。しかしながら、ホルモン治療で処置された患者に関しては、中間の MS 群は、低 MS 群のものと同じリスクを有している。高い MS 群のリスクは、中間の MS および低 MS 群のリスクよりも高い。

【0170】

14 個の遺伝子サインを適用する別の方法は、下記のように、等式 2 を用いることによる。14 個の遺伝子サインは、以前に、実施例 1 に記載されるように、カリフォルニア パシフィック メディカル センター (CPMC) によるプロファイリング研究を用いて開発された。プログラムの創意による経路分析は、サインの 14 個の遺伝子のうち大部分が、細胞増殖と関連していることを示した。14 個の遺伝子のうち 10 個は、不良転帰の腫瘍における協調的に過剰発現することを見出された TP53 シグナル経路と関連している。

【0171】

転移スコア (MS) は、各個体について計算された。この実施例における MS は、前述したように、サインにおける 14 個の遺伝子の遺伝子発現に基づいていた。しかしながら、この実施例における MS を計算するためのアルゴリズムは、14 個の遺伝子が等しく計量される等式 2 に基づいていた。さらに、この実施例では、2 つのカットポイントは、前述の 3 つの実施例における低および高リスク群に患者を分類するためのたった 1 つカットポイントを用いるのとは対照的に、高、中および低 MS 群に患者をグループ化するために用いられた。サインの 14 個の遺伝子は、実施例 1 に記載されるように、この研究において選択され、新しい MS スコアおよびカットポイントは、実施例 2 に記載されるガイ病院からの未処置の試料を用いた研究に基づいて測定された。新しい MS アルゴリズムは、実施例 4 および 5 において適用され、検証された。

【0172】

MS (新しい) の 2 つのカットポイントは、高、中および低 MS 群に患者を分類するために選択された。カットポイントは、第 1 のカットポイントが - 0 . 119 である上記 MS を有する個体が 5 年以内に遠隔転移を有する高リスクの個体として分類される場合、M

S リスク群の感受性は 0.7 程度となるように決定された。MS が 0.302 である第 2 のカットポイントは、感受性および特異性が 0.7 となるように選択された。

【0173】

実施例 2 において測定された MS と、14 個全ての遺伝子の遺伝子発現の平均の負数としての MS は、互いに単に一次変換であることに留意しなければならない。それ自体は、それらは完全相関（ピアソン相関係数 = 1）を有し、高、中、及び低 MS への患者の分類は、記載されている対応するカットポイントを用いることは同じである。

統計学的分析

遠隔転移なしの生存に関するカプラン - マイヤー（KM）曲線は、高、中および低 MS 群について作成された。3 つの群の DMFS 比を調べると、中および低 MS 群は、高 MS を有する高リスク群と比較される低リスク群として組み合わせられた。ログランク試験を行った。

【0174】

DMFS カットポイントについての MS 群の一変量および多変量の Cox 比例ハザード抑制分析を行った。高リスク（高 MS）群対低リスク（組み合わせられた中間および低 MS 群）群のハザード率は、多変量分析における年齢（年）、腫瘍サイズ（cm）、および組織学的悪性度について調整された。

【0175】

MS と年齢、腫瘍サイズとの関連性は、ANOVA 試験を用いて調べられ、MS 群と組織学的悪性度との関連は、関連の強固性については Crammer's V によって、統計学的有意性については chi-sq によって評価された。

【0176】

また、MS リスク群のハザード率は、異なる臨床のサブグループについて計算された。それらの群には、閉経前対閉経後、年齢 55 歳対 > 55 歳、腫瘍サイズ > 2 cm 対 2 cm、組織学的悪性度 1 & 2 対悪性度 3 が含まれる。

【0177】

診断精度を評価するために、5 年以内の遠隔転移を予測する受信者動作特性（ROC）曲線をプロットした。ROC 曲線下の面積（AUC）を計算した。感受性、特異性、陽性予測値（PPV）および陰性予測値（NPV）は、MS による高対低リスク群に関する 95% 信頼区間（95% CI）を用いて計算された。

【0178】

MS 群のハザード率の時間依存性は、打ち切りデータを操作することができるスプライン曲線フィッティング技術を用いて年換算のハザードを評価することによって調査された。R2.4.1 における HEF T 手法を用いた。年換算のハザードは、MS 高および低リスク群について評価され、そこから、異なる時間のハザード率を計算した。

【0179】

カプラン - マイヤー推定および Cox 比例ハザードは、R2.4.1 および SAS 9.1 を用いて行われた。ROC 曲線および AUC は、Mayo Clinic's ROC プログラムを用いて推定された。AUC の信頼区間を推定するための DeLong 法を用いた。

結果

MS 低リスクおよび高リスク群における遠隔転移なしの生存率

136 人の個体の低 MS 群において 8 個の遠隔転移、29 人の個体の高 MS 群において 2 個の遠隔転移、30 人の個体の高 MS 群において 7 個の遠隔転移があった。10 年の DMFS 比（SE）は、低、中、および MS 群について、それぞれ 0.921（0.028）、0.966（0.034）および 0.804（0.068）であった。DMFS 比において有意差があり、ログランクの p 値は 0.04 であった。DMFS 比は、低および中の MS 群において同じであったので、それらを組み合わせ、低リスク群を形成した。低リスク群は、10 年の DMFS 比が 0.928（0.025）であり、高リスク群についての 0.804（0.068）の対応する比とは有意に異なっていた（表 18）。ログラ

10

20

30

40

50

ンクのp値は0.011であった。3つのMS群および2つのMSリスク群に対する遠隔転移なしの生存のカプラン-マイヤープロットは、それぞれ図8および図9に示された。

一変量および多変量のCox抑制モデルからのハザード率

遠隔転移までの時間を予測するためのMS高リスク対低リスク群の一変量ハザード率は、3.25(95%CI=1.24または8.54、p値=0.017)であった。年齢、腫瘍サイズおよび組織学的悪性度によって調整された場合、ハザード率は、表19において5.82(1.71-1.75)であった。MSリスク群は、多変量分析において有意である唯一のリスク因子であった。したがって、遺伝子サインは、伝統的な臨床病理学的なリスク因子と比較したDMFSに対する独立した予後値を有し、これらの因子の情報の一部を捕捉する。

10

MSリスク群と、他の臨床および病理学的特徴との関連性

MSリスク群は、組織学的悪性度との非常に有意な関連を有する(関連についてのchi-sqに関するCrammer's V=0.65、p<0.0001)。60個の悪性度1の腫瘍について、MS高リスクはなかった(0%)。対照的に、98個の悪性度2の腫瘍については9(9.2%)、47個の悪性度3の腫瘍については31(66%)であった。

【0180】

腫瘍サイズは、MS高リスク群において大きい、この差は、統計学的な差異ではなかった(MS高および低リスク群では、それぞれ1.87cmおよび1.61cm、ANOVA p値=0.14)。MSリスク群と年齢との関連性は有意ではなかった(平均年齢は、高および低リスク群では59.3であり、ANOVA p=0.34)。結果を表20に概説した。

20

異なる臨床サブグループにおける分子サインの性能

MSリスク群のハザード率は、2cmの腫瘍で3.7(1.1-12.6)であり、>2cmの腫瘍で3.0(0.6-14.8)であった。55歳よりも若い女性では、HRは、7.4(1.0-54.7)であり、55歳よりも上の女性では2.8(0.88-8.8)であった。組織学的悪性度1および2の腫瘍については、HRは、12.8(3.8-43.1)であり、悪性度3の腫瘍について、HRは1.53(0.16-14.7)であった(表21)。

診断精度および予測値

30

5年以内の遠隔転移を予測するためのMSリスク群の感受性、特異性、PPVおよびNPVを表22に示した。MSリスク群の感受性は、0.50(0.50-0.76)であり、特異性は、0.82(0.76-0.87)であった。0.05であると推定された5年の遠隔転移比を用いると、PPVおよびNPVは、それぞれ0.13(0.068-0.23)および0.97(0.94-0.98)と推定された(表22)。MSリスク群の高いNPVは、低リスクである患者の化学療法などのより積極的な処置を除外するために用いられることが重要である。

【0181】

5年以内の遠隔転移を予測するための連続MSのROC曲線を図10に示し、AUCは0.72(0.57-0.87)であると推定された。

40

予後サインの時間依存性

MS高および低リスク群についての年換算したハザード率を図11aに示し、2つの群の間のハザード率の時間依存性を図11bに示した。高リスク群については、年のハザード率は外科的処置から約3年で2.5%のピークに達し、その後、次の数年かけて徐々に減少した。しかしながら、低リスク群の年換算のハザード率は、フォローアップした10年間で、僅かであるが着実な増加を示した。その後、MSリスク群のハザード率は、時間依存的であった。2、5、および10年で、それぞれ4.6、3.6および2.1であった。

(実施例5)

14個の遺伝子サインは234個のFFPE試料を用いたアジュバントホルモンの処置

50

されたリンパ節転移陰性のE R陽性乳癌患者における遠隔転移を予測する患者

愛知県がんセンターにおいて1995年から2003年に外科的処置を受けたN -、E R + 乳癌の234人の日本人女性の集団を選択した。フォローアップの中央値は8.7年であった。それらの中には、31人(13%)が遠隔転移であり、19人(8.1%)が死亡し、46人(19.7%)が局所および遠隔再発であった。

【0182】

146人(62%)の患者がステージIであり、88人(38%)がステージIIであった。全ての患者はアジュバントホルモン処置を受けたが、化学療法は受けなかった。112人の閉経後の女性は、タモキシフェンで処置された。122人の閉経前の女性のうち、102人がタモキシフェンを受け、20人がゾラデックス処置を受けた。この手段は、平均(SD)年齢は53(11)歳であり、平均(SD)腫瘍直径は2.05cm(1.1)であった。74人(32%)、113人(48%)、および47人(20%)の患者は、組織学的悪性度がそれぞれ1、2および3の腫瘍であった。(表23)

遺伝子発現サインおよび転移スコア(MS)

14個の遺伝子発現サインは前述において作製され、F F E 試料を用いたR T - P C Rにより米国および欧州におけるプロファイリング研究において評価された。プログラムIngenuityによる経路分析は、サインの14個の遺伝子の大多数が、細胞増殖と関連していることを示した。14個の遺伝子のうち10個は、乏しい結果の腫瘍において協調的に過剰発現することが見出されたT P 5 3シグナル経路と関連している。

【0183】

転移スコア(MS)は、各個体について計算された。MSは、サインの14個の遺伝子の遺伝子発現レベル(Ctによる)の平均の負数に基づいていた。さらに、2つのカットポイントは、以前に、ガイ病院からの未処置の患者からの腫瘍組織を用いた研究から測定され(実施例2)、患者を高、中間および低MS群にグループ分けした。

統計学的分析

遠隔転移なしの全生存に関するカプラン - マイヤー(KM)曲線は、高、中間および低MS群について作製された。3つの群のDMFS比を調べると、中間および低MS群は、高リスク群と比較される低リスク群として、高MSと比較された。ログランク試験を行った。

【0184】

DMFSおよびOSエンドポイントに関するMS群の一変量および多変量のCox比例ハザード抑制分析を行った。高リスク(高MS)対低リスク(中間および低MS群の組み合わせ)群は、1つの多変量分析において、年齢(年)、腫瘍サイズ(cm)および組織学的悪性度について調整された。別の多変量分析では、閉経状態、処置、腫瘍サイズ、組織学的悪性度およびPgR状態について調整された。

【0185】

MS群と、年齢、腫瘍サイズとの関連性は、ANOVA試験を用いて調べられ、MS群と組織学的悪性度および腫瘍サブタイプとの関連性は、関連した強度についてのCramer's V、統計学的差異についてはchi-sqによって評価された。

【0186】

また、MSリスク群のハザード率は、異なる臨床サブグループについて計算された。それらのグループは、閉経前対閉経後、年齢55歳対>55歳、腫瘍サイズ<2cm対2cm、組織学的悪性度1&2対悪性度3、PgR+ve対-veを含む。

【0187】

診断精度を評価するために、5年以内の遠隔転移を予測するMSの受信者動作特性(ROC)曲線をプロットした。ROC曲線下の面積(AUC)を計算した。感受性、特異性、陽性予測値(PPV)および陰性予測値(NPV)は、MSによる高対低リスク群に関する95%信頼区間(95%CI)を用いて計算された。

【0188】

M S 群のハザード率の時間依存性は、打ち切りデータを操作することができるスプライン曲線フィッティング技術を用いて年換算のハザードを評価することによって調査された。R 2 . 4 . 1 における H E F T 手法を用いた。年換算のハザードは、M S 高および低リスク群について評価され、そこから、異なる時間のハザード率を計算した。

【 0 1 8 9 】

カプラン - マイヤー推定および C o x 比例ハザードは、R 2 . 4 1 . 1 および S A S 9 . 1 を用いて行われた。R O C 曲線および A U C は、M a y o C l i n i c ' s R O C プログラムを用いて推定された。A U C の信頼区間を推定するための D e l l o n g 法を用いた。

結果

10

M S 低リスクおよび高リスク群における遠隔転移なしの生存率

7 7 人の個体の低 M S 群において 6 個の遠隔転移、6 6 人の個体の中間 M S 群において 4 個の遠隔転移、9 5 人の個体の高 M S 群において 2 1 個の遠隔転移があった。1 0 年の D M F S 比 (S E) は、低、中、および高 M S 群について、それぞれ、0 . 8 9 (0 . 0 5)、0 . 9 1 (0 . 0 4) および 0 . 7 5 (0 . 0 5) であった。D M F S 比において有意差があり、ログランクの p 値は 0 . 0 0 4 であった。D M F S 比は、低および中 M D 群において同じであったので、それらを組み合わせて、低リスク群を形成した。低リスク群は、1 0 年の D M F S 比が 0 . 8 9 5 (0 . 0 3 4) であり、高リスク群についての 0 . 7 5 (0 . 0 5) の対応する比とは有意に異なっていた (表 2 4)。ログランクの p 値は 0 . 0 0 0 9 2 であった。3 つの M S 群に対する遠隔転移なしの生存のカプラン - マイ

20

ー変量および多変量の C o x 抑制モデルからのハザード率

遠隔転移までの時間を予測するための M S 高リスク対低リスク群の一変量ハザード率は、3 . 3 2 (9 5 % C I = 1 . 5 6 ~ 7 . 0 6、p 値 = 0 . 0 0 1 8) であった。年齢、腫瘍サイズおよび組織学的悪性度によって調整された場合、ハザード率は、3 . 7 9 (1 . 4 2 - 1 0 . 1、p = 0 . 0 0 7 8) であった。M S に加えて、腫瘍サイズは、多変量分析における唯一の他の有意因子であり、H R は 1 c m あたり 1 . 4 増加した (p = 0 . 0 0 7) (表 2 5)。別の多変量分析では、M S リスク群は、閉経状態、処置、腫瘍、腫瘍サイズ、組織学的悪性度および P g R 状態によって調整された。M S リスク群のこの調

30

整されたハザード率は、3 . 4 4 (1 . 2 7 - 9 . 3 4) であった (表 2 7)。さらに、腫瘍サイズは、この多変量分析において唯一の他の有意因子であった (H R = 1 c m あたり 1 . 4 5 の増加、p = 0 . 0 0 4 9)。したがって、遺伝子サインは、伝統的な臨床病理学的なリスク因子と比較した D M F S に対する独立した予後値を有し、これらの因子の情報の一部を捕捉する。

M S リスク群と、他の臨床および病理学的特徴との関連性

M S リスク群は、組織学的悪性度との非常に有意な関連を有する (関連についての c h i - s q に関する C r a m m e r ' s V = 0 . 5 4、p < 0 . 0 0 0 1)。7 4 個の悪性度 1 の腫瘍について、ほんの 4 個 (5 . 4 %) が M S 高リスクであった。対照的に、1 1 3 個の悪性度 2 の腫瘍については 5 4 (4 7 . 8 %)、4 7 個の悪性度 3 の腫瘍につい

40

ては 3 7 (7 8 . 7 %) が M S 高リスクであった。

【 0 1 9 0 】

また、M S リスク群は、腫瘍のサブタイプと関連した (C r a m m e r ' s V = 0 . 2 5、p = 0 . 0 2)。7 7 個の硬性腫瘍のうち 2 3 個 (2 9 . 8 %) が M S 高リスクであり、1 0 8 個の乳頭管腫瘍のうち 4 5 個 (4 1 . 7 %)、3 8 個の固定管の 2 4 個 (6 3 . 2 %) が M S 高リスクであった。

【 0 1 9 1 】

腫瘍サイズは、M S 高リスク群において大きい (M S 高および低リスク群においてそれぞれ 2 . 2 3 c m および 1 . 9 3 c m、A N O V A p 値 = 0 . 0 3 7)、M S リスク群と年齢との関連性は有意ではなかった (p = 0 . 2 9)。結果を表 2 6 に概説した。

50

異なる臨床サブグループにおける分子サインの性能

MS リスク群は、 ≤ 2 cm、低悪性度（悪性度 1 および 2）および PgR + ve である腫瘍を有する若年（年齢 55 歳）、閉経前の女性において遠隔転移を予測することができる。MS リスク群のハザード率は、 ≤ 2 cm の腫瘍では 4.5 (1.2 - 17.3) および > 2 cm の腫瘍では 2.3 (0.92 - 5.6) であった。閉経前の女性では、HR は、6.0 (1.6 - 23.3) であり、閉経後の女性では 2.1 (0.83 - 5.1) であった。組織学的悪性度 1 および 2 の腫瘍を有する女性については、HR は 3.6 (1.5 - 8.4) であり、悪性度 3 の腫瘍では、HR は 2.4 (0.29 - 18.8) であった。MS リスク群の HR は、PgR + ve 腫瘍では 3.5 (1.4 - 9.0) であり、PgR - ve 腫瘍では 2.1 (0.57 - 7.49) であった（表 28）。

10

診断精度および予測値

5 年以内の遠隔転移を予測するための MS リスク群の感受性、特異性、PPV および NPV を表に示した。MS リスク群の感受性は、0.81 (0.60 - 0.92) であり、特異性は、0.65 (0.58 - 0.71) であった。0.0095 であると推定された 5 年の遠隔転移比を用いると、PPV および NPV は、それぞれ 0.19 (0.15 - 0.24) および 0.97 (0.93 - 0.99) と推定された（表 29）。MS リスク群の高い NPV は、低リスクである患者の化学療法などのより積極的な処置を除外するために用いられることが重要であった。5 年以内の遠隔転移を予測するための連続 MS の ROC 曲線を図 14 に示し、AUC は 0.73 (0.63 - 0.84) であると推定された。

予後サインの時間依存性

20

MS 高および低リスク群についての年換算したハザード率を図 15 a に示し、2 つの群の間のハザード率の時間依存性を図 15 b に示した。高リスク群については、年のハザード率は外科的処置から約 3 年で 3.5 % のピークに達し、その後、次の数年かけて徐々に減少した。しかしながら、低リスク群の年換算のハザード率は、フォローアップした 10 年間で、僅かであるが着実な増加を示した。その後、MS リスク群のハザード率は、時間依存的であった。2、5、および 10 年で、それぞれ 4.8、3.4 および 1.8 であった。

【0192】

この実施例および前述の実施例から分かるように、14 個の遺伝子サインは、白人およびアジア民族系の乳癌患者における有効なリスク予測材料であることが示され、14 個の遺伝子予後サインの構造安定性を指示している。

30

【0193】

本明細書に引用されている全ての刊行物および特許は、全体として参照により本明細書に援用される。本発明の記載されている組成物、方法およびシステムの種々の修飾および変更は、本発明の範囲および精神から逸脱することなしに当業者に明確となる。本発明は、特定の好ましい態様およびある種の実施例と関連して記載されているが、請求される発明はこのような特定の態様に不当に限定されてはならないことは理解すべきである。実際に、分子生物学、遺伝学および関連した分野における当業者に明らかである、本発明を実施するための上述した様式の種々の修飾は、続く特許請求の範囲にあることが意図される。

40

【0194】

【表 1】

表1. CPMCおよびガイ病因からの患者の臨床的および病理学的特徴

リンパ節転移陰性のER陽性の未処置患者の臨床的および病理学的特徴		
特徴	トレーニング(UCSF) n=142	検証(ガイ) n=280
年齢		
≤ 55 歳	40 (28.2%)	144 (51.4%)
> 55 歳	102 (71.8%)	136 (48.6%)
平均(標準偏差)	62 歳 (12.6)	55.5 歳 (11.6)
最小-最大	31 歳 - 89 歳	29 歳 - 87 歳
腫瘍直径		
≤ 2 cm	126 (88.7%)	167 (59.6%)
> 2 cm	8 (5.6%)	113 (40.4%)
不明	8 (5.6%)	0 (0%)
平均(標準偏差)	1.28 cm (0.50)	1.93 cm (0.85)
腫瘍悪性度		
悪性度1	74 (52.1%)	60 (21.4%)
悪性度2	61 (43%)	166 (59.3%)
悪性度3	4 (2.8%)	54 (19.3%)
不明	3 (2.1%)	0 (0%)
ステージ		
I	117 (82.4%)	167 (59.6%)
IIA	25 (17.6%)	113 (40.4%)
遠隔再発		
はい	31 (21.8%)	66 (23.4%)
いいえ	111 (78.2%)	214 (76.6%)
全てのケースの死亡		
はい	56 (39.4%)	135 (48.2%)
いいえ	86 (60.6%)	145 (51.8%)
中央値フォローアップ	8.7年	15.6年

10

20

【 0 1 9 5 】

【表2】

表2. 14個の転移予後パネルおよび内在性対照を含む遺伝子

遺伝子	MS定常ai	参照配列	説明	参考資料
CENPA	0.29	NM_001809	セントロメアプロテインA、17kDa	Black,B.E., Foltz,D.R., et al., <i>Nature</i> 430(6999):578-582 (2004)
PKMYT1	0.29	NM_004203	膜関連チロシンおよびスレオニン特異的cdc-2阻害性キナーゼ	Bryan,B.A., Dyson,O.F. et al., <i>J. Gen. Virol.</i> 87 (PT 3), 519-529 (2006)
MELK	0.29	NM_014791	母体胚性ロシインジッパーキナーゼ	Beullens,M., Vancauwenbergh,S. et al., <i>J. Biol. Chem.</i> 280 (48), 40003-40011 (2005)
MYBL2	0.29	NM_002466	v-myb骨髄芽球症ウイルス癌原遺伝子相同体(鳥類)様2	Bryan,B.A., Dyson,O.F. et al., <i>J. Gen. Virol.</i> 87 (PT 3), 519-529 (2006)
BUB1	0.27	NM_004366	ペンズミダゾール1相同体によって阻害されないBUB1出芽	Morrow,C.J., Tighe,A. et al., <i>J. Cell. Sci.</i> 118 (PT 16), 3639-3652 (2005)
RACGAP1	0.29	NM_013277	Rac GTPase 活性化プロテイン1	Niiya,F., Xie,X. et al., <i>J. Biol. Chem.</i> 280 (43), 36502-36509 (2005)
TKI	0.27	NM_003258	チミジンキナーゼ1、可溶性	Karbownik,M., Brzezianska,E. et al., <i>Cancer Lett.</i> 225 (2), 267-273 (2005)
UBE2S	0.27	NM_014501	ユビキチン結合酵素E2S	Liu,Z., Diaz,L.A. et al., <i>J. Biol. Chem.</i> 267 (22), 15829-15835 (1992)
DC13	0.22	AF201935	DC13タンパク質	Gu,Y., Peng,Y. et al., Direct Submission, Submitted (05-NOV-1999) Chinese National Human Genome Center at Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China
RFC4	0.25	NM_002916	複製因子C (活性化因子1)4, 37kDa	Gupte,R.S., Weng,Y. et al., <i>Cell Cycle</i> 4 (2), 323-329 (2005)
PRR11 (FLJ11029)	0.26	NM_018304	プロリンリッチ11	Weinmann,A.S., Yan,P.S. et al., <i>Genes Dev.</i> 16 (2), 235-244 (2002)
DIAPH3	0.23	NM_030932	透明な相同体3 (ショウジョウバエ)	Katoh,M. and Katoh,M., <i>Int. J. Mol. Med.</i> 13 (3), 473-478 (2004)
ORC6L	0.28	NM_014321	原認識複合体、サブユニット6相同体様(酵母)	Sibani,S., Price,G.B. et al., <i>Biochemistry</i> 44 (21), 7885-7896 (2005)
CCNB1	0.23	NM_031966	サイクリンB1	Zhao,M., Kim,Y.T. et al., <i>Exp Oncol</i> 28 (1), 44-48 (2006)
PPIG	EC	NM_004792	ペプチジルプロリルイソメラーゼG	Lin,C.L., Leu,S. et al., <i>Biochem. Biophys. Res. Commun.</i> 321 (3), 638-647 (2004)
NUP214	EC	NM_005085	ヌクレオポリン 214kDa	Graux,C., Cools,J. et al., <i>Nat. Genet.</i> 36 (10), 1084-1089 (2004)
SLU7	EC	NM_006425	ステップII スプライシング因子	Shomron,N., Alberstein,M. et al., <i>J. Cell. Sci.</i> 118 (PT 6), 1151-1159 (2005)

注釈: 本明細書に開示されている全ての遺伝子の発現プロファイリングのためのPCRプライマーは、出願時に知られている全ての転写変異体を増幅するために設計された。

EC = 内在性対照

参照配列 = この遺伝子の1変異体についてのNCBI参照配列

【0196】

【表 3】

表3. 遺伝子発現プロファイリングに用いられるプライマー

遺伝子	配列	配列番号	方向	
BUB1	CATGGTGGTGCCTTCAA	配列番号 1	上流	
CCNB1	GCCAAATACCTGATGGAACATA	配列番号 3	上流	
CENPA	CAGTCGGCGGAGACAA	配列番号 5	上流	
DC13	AAAGTGACCTGTGAGAGATTGAA	配列番号 7	上流	
DIAPH3	TTATCCCATCGCCTTGAA	配列番号 9	上流	
MELK	AGAGACGGGCCCCAGAA	配列番号 11	上流	
MYBL2	GCGGAGCCCCATCAA	配列番号 13	上流	10
ORC6L	CACTTCTGCTGCACTGCTTT	配列番号 15	上流	
PKMYT1	CTACCTGCCCCCTGAGTT	配列番号 17	上流	
PRR11	TGTCCAAGCTGTGGTCAAA	配列番号 19	上流	
RACGAP1	GACTGCGAAAAGCTGGAA	配列番号 21	上流	
RFC4	TTTGGCAGCAGCTAGAGAA	配列番号 23	上流	
TK1	GATGGTTTCCACAGGAACAA	配列番号 25	上流	
UBE2S	CCTGCTGATCCACCCTAA	配列番号 27	上流	
NUP214	CACTGGATCCCAAGAGTGAA	配列番号 29	上流	
PPIG	TGGACAAGTAATCTCTGGTCAA	配列番号 31	上流	
SLU7	TGCCAATGCAGGAAAGAA	配列番号 33	上流	
BUB1	GCTGAATACATGTGAGACGACAA	配列番号 2	下流	
CCNB1	CTCCTGCTGCAATTTGAGAA	配列番号 4	下流	20
CENPA	AAGAGGTGTGTGCTCTTCTGAA	配列番号 6	下流	
DC13	CGCCCTGCCCAACAA	配列番号 8	下流	
DIAPH3	TGTCACACCATGTGTGTA	配列番号 10	下流	
MELK	CAACAGTTGATCTGGATTCATA	配列番号 12	下流	
MYBL2	CATCCTCATCCACAATGTCAA	配列番号 14	下流	
ORC6L	GGATGTGGCTACCATTTTGT	配列番号 16	下流	
PKMYT1	AGCATCATGACAAGGACAGAA	配列番号 18	下流	
PRR11	TCTCCAGGGGTGATCAGAA	配列番号 20	下流	
RACGAP1	TTGCTCCTCGCTTAGTTGAA	配列番号 22	下流	
RFC4	CACGTTTCATCAGATGCATTTAA	配列番号 24	下流	
TK1	GGATCCAAGTCCCAGCAA	配列番号 26	下流	
UBE2S	GCATACTCCTCGTAGTTCTCAA	配列番号 28	下流	
NUP214	TGATCCCACTCCAAGTCTAGAA	配列番号 30	下流	30
PPIG	GTATCCGTACCTCCGCAAA	配列番号 32	下流	
SLU7	TGGTATCTCCTGTGTACCTAACAA	配列番号 34	下流	

【 0 1 9 7 】

【表 4】

表4. CPMCからの142個の資料におけるサインの14個の遺伝子の遺伝子発現レベルの平均(μ)および標準偏差(σ)(aiは第1の主成分への荷重である)

遺伝子	ai	μ	σ
CENPA	0.29	-0.18	1.40
TK1	0.27	-1.07	1.68
BUB1	0.27	2.69	1.78
PRR11	0.26	5.25	1.90
UBE2S	0.27	2.64	1.41
DC13	0.22	0.47	1.15
DIAPH3	0.23	1.90	1.34
MELK	0.29	-1.77	1.41
MYBL2	0.29	2.54	1.88
PKMYT1	0.29	-0.35	1.73
RFC4	0.25	1.25	0.99
ORC6L	0.28	1.98	1.43
RACGAP1	0.29	4.00	1.09
CCNB1	0.23	3.20	1.66

【 0 1 9 8 】

【表 5】

表5

表5. CPMCからの患者に関する5年および10年での低リスクおよび高リスク予後群についての遠隔転移なしの生存および全生存					
転移なしの生存率			全生存率		
群	患者数	5年 (標準誤差)	10年 (標準誤差)	5年 (標準誤差)	10年 (標準誤差)
低リスク	71	0.96 (0.025)	0.90 (0.037)	0.90 (0.036)	0.78 (0.059)
高リスク	71	0.74 (0.053)	0.62 (0.066)	0.79 (0.049)	0.48 (0.070)
全て	142	0.85 (0.031)	0.76 (0.040)	0.84 (0.031)	0.63 (0.048)

10

【 0 1 9 9 】

【表 6】

表6

表6. CPMCからの患者に関する14個の遺伝子予後サイン、腫瘍サイズ、腫瘍悪性度および年齢の一変量および多変量のCox比例ハザード抑制分析				
一変量分析			多変量分析	
変数	ハザード率 (95% CI)	P値	ハザード率 (95% CI)	P値
14個の遺伝子サイン	4.23 (1.82 - 9.85)	0.0008	3.26 (1.26 - 8.38)	0.014
年齢	1.00 (0.97-1.03)	0.850	1.00 (0.98-1.03)	0.810
腫瘍サイズ	1.07 (1.00-1.14)	0.050	1.03 (0.96-1.10)	0.450
腫瘍悪性度 (中程度+高い)	2.18 (1.04-4.59)	0.040	1.26 (0.55 - 2.87)	0.580
14個の遺伝子予後サインについては、ハザード比は、患者を分類するために中央値MS (CV)を用いて高リスク対低リスク群を把握する。年齢については、ハザード比は、年齢の増加において各年についてのハザード増加として与えられる。腫瘍サイズについては、直径にして1センチメートルの増加あたりハザードとして与えられる。腫瘍悪性度については、中程度および高悪性度腫瘍対低悪性度腫瘍を有する群のハザード。				

20

30

【 0 2 0 0 】

【表 7】

表7

遺伝子	M値
PPIG	0.8046
SLU7	0.8741
NUP214	0.8886
PPP1CA	1.0256
TERF2	1.1907
EEF1A1	1.1994

40

【 0 2 0 1 】

【表 8】

表8

遺伝子	M値
PPIG	0.5697
NUP214	0.5520
SLU7	0.6075

【 0 2 0 2 】

10

【表 9】

表9

表9. CPMCからの患者に関する5年以内の遠隔転移、5年を超えた転移なしを予測するためのカットポイントとして、中央値MS(CV)を用いた分子予後サインの予後値		
(A). 5年以内お遠隔転移—予後対実際の結果		
群	5年以内の 遠隔転移	疾患なし >5年
	(n=21)	(n=95)
MSによる高リスク	18	39
MSによる低リスク	3	56
(B). 5年以内の遠隔転移を予測するための予後サインの診断測定基準		
	値	95% CI
感受性	0.86	0.65 - 0.95
特異性	0.59	0.49 - 0.63
オッズ比	8.62	2.37 - 31.26
PLR	2.09	1.55 - 2.81
NLR	0.24	0.084 - 0.70
PPV*	0.26	0.21 - 0.33
NPV*	0.96	0.89 - 0.99
PLR - 陽性可能性比 , NLR - 陰性可能性比 ,		
PPV - 陽性予測値 , NPV - 陰性予測値		
* 予測値は現在のデータから0.15であると推定された5年以内の遠隔転移の有病率を用いて計算された		
**5年を超えた遠隔転移を有する個体と5年前に検査されたものは含まれなかった。		

20

30

40

【 0 2 0 3 】

【表 10】

表10. ガイ病院からの未処置患者に関する種々のリスク分類についてのハザード比
(未小生) (95%の確実な間隔)およびp値

分類	遠隔転移までの時間	全生存
遺伝子サイン(高リスク 対 低リスク)	6.12 (2.23 -16.83) p=0.0004	2.49 (1.50 - 4.14) p=0.0004
アジュバント!(高リスク 対 低リスク)	2.63 (1.30 - 5.32) p=0.007	2.89 (1.71 - 4.87) p<0.0001
年齢(55歳未満 対 >55歳)	1.45 (0.89 - 2.36) p=0.13	2.91 (2.03 - 4.18) p<0.0001
腫瘍サイズ(T2 対 T1)	2.27 (1.39 - 3.69) p=0.001	1.60 (1.14 - 2.25) p=0.0062
組織学的悪性度(悪性度2+3 対 悪性度1)	2.56 (1.17 - 5.61) p=0.019	2.56 (1.44 - 4.54) p=0.0013

10

【 0 2 0 4 】

【表 11】

表11. ガイ病院からの未処置患者に関する14個の遺伝子予後サイン、腫瘍サイズ、腫瘍悪性度および年齢についての遠隔転移までの時間の一変量および多変量のCoxモデル

20

表11a 14個の遺伝子予後サイン、腫瘍サイズ、腫瘍悪性度および年齢についての
遠隔転移(DMFS)までの一変量および多変量のCoxモデル

変数	一変量分析		多変量分析	
	ハザード率 (95% CI)	p値	ハザード率 (95% CI)	p値
14個の遺伝子サイン	6.12 (2.23-16.83)	0.0004	4.81 (1.71-13.53)	0.003
年齢	1.03 (1.00-1.05)	0.024	1.01 (0.99-1.04)	0.251
腫瘍サイズ	1.74 (1.24-2.44)	0.001	1.39 (0.97-2.00)	0.076
悪性度2	2.45 (1.10-5.43)	0.028	1.43 (0.63-3.23)	0.390
悪性度3	2.98 (1.21-7.30)	0.017	1.40 (0.55-3.53)	0.478

30

14個の遺伝子予後サインに関しては、ハザード率は、患者を分類するためのカットポイントとして、以前に定義したゼロMSを用いた高リスク対低リスク群を含む。年齢に関しては、ハザード率は、年齢における各年の増加に対するハザード増加として与えられる。腫瘍サイズに関しては、ハザード率は、直径にして各センチメートル増加あたりのハザード増加として与えられる。腫瘍悪性度に関しては、悪性度2および悪性度3である群のハザードは、悪性度1の腫瘍と比較された。

表11b MSおよびアジュバント!によるリスク群に関する遠隔転移(DMFS)までの時間の
一変量および多変量のCoxモデル

変数	一変量分析		多変量分析	
	ハザード率 (95% CI)	p値	ハザード率 (95% CI)	p値
14個の遺伝子サイン	6.12 (2.23 - 16.83)	0.0004	5.32 (1.92 - 14.73)	0.001
アジュバント!	2.63 (1.30 - 5.32)	0.007	2.06 (1.02 - 4.19)	0.045

40

14個の遺伝子予後サインに関しては、ハザード率は、患者を分類するためのカットポイントとして、以前に定義したゼロMSを用いた高リスク対低リスク群を含む。アジュバント!に関しては、ハザード率は、アジュバント!オンラインプログラムによって計算されるように、10年における20%の再発可能性のカットポイントを用いた高リスク対低リスク群を含む。

【 0 2 0 5 】

【表 1 2】

表12. ガイ病院からの未処置患者に関して、5年および10年のMSによる低リスクおよび高リスク群についての遠隔転移なしの生存率および全生存率

5年および10年の低リスクおよび高リスク予後群についての 遠隔転移なしの生存率および全生存率					
群	患者数	遠隔転移なしの生存率		全生存率	
		5年(標準誤差)	10年(標準誤差)	5年(標準誤差)	10年(標準誤差)
低リスク	71	0.99 (0.014)	0.96 (0.025)	0.97 (0.020)	0.94 (0.028)
高リスク	209	0.86 (0.025)	0.76 (0.031)	0.92 (0.019)	0.71 (0.032)
全て	280	0.89 (0.019)	0.81 (0.024)	0.93 (0.015)	0.77 (0.026)

10

【 0 2 0 6 】

【表 1 3】

表13. サブグループ分析: ガイ病院からの未処置患者に関するアジュバント！の異なるサブグループにおける遠隔転移までの時間(DMFS)、組織学的悪性度、腫瘍サイズ、年齢および閉経状態についてのMSリスク群のハザード率

	患者数	イベント数	MSリスク群のハザード率 (95% CI)	p値
アジュバント！				
高リスク	205	57	3.72 (1.34-10.28)	0.0114
低リスク	75	9	無限*	0.01
腫瘍悪性度				
高い悪性度(悪性度2、3)	220	59	4.59 (1.44-14.67)	0.010
低い悪性度(悪性度1)	60	7	7.58 (0.91-63.08)	0.061
腫瘍サイズ				
≤2 cm	167	28	14.16 (1.92-104.22)	0.009
> 2 cm	113	38	2.64 (0.81-8.58)	0.110
年齢				
≤ 55歳	144	30	13.58 (1.85-99.74)	0.010
> 55歳	136	36	3.45 (1.06-11.26)	0.040
閉経状態				
閉経前	102	21	7.85 (1.05-58.49)	0.044
閉経後	157	40	4.88 (1.51-15.84)	0.008

*0イベントである34のMS低リスク、9イベントである41のMS高リスク

20

30

【 0 2 0 7 】

【表 1 4】

表14. ガイ病院からの未処置患者に関する10年以内の遠隔転移についてのMSおよびアジュバント！リスク群の診断精度および予測値

リスク分類	10年以内の遠隔転移			
	感受性 95% CI	特異性 95% CI	PPV 95% CI	NPV 95% CI
MSリスク群	0.94 (0.84 - 0.98)	0.3 (0.24 - 0.37)	0.23 (0.21 - 0.25)	0.97 (0.88 - 0.99)
アジュバント！リスク群	0.90 (0.78 - 0.96)	0.31 (0.26 - 0.38)	0.23 (0.20 - 0.25)	0.93 (0.85 - 0.97)

40

【 0 2 0 8 】

【表 1 5】

表15. ミュエンスター大学からの未処置および処置患者の臨床および病理学的特徴

特徴	全ての患者 N = 96	全ての処置 N = 62	タモキシフェンだけの処置 N = 54
年齢			
平均(SD)	56.64 (12.6)	56.43 (12.4)	57.63 (12.1)
>55歳	38 (40.6%)	26 (41.9%)	26 (48.1%)
<55歳	37 (38.5%)	27 (43.5%)	22 (40.7%)
不明	20 (20.8%)	9 (14.5%)	6 (11.1%)
Tステージ			
1	56 (58.3%)	37 (59.7%)	33 (61.1%)
1C	1 (1.0%)	1 (1.6%)	1 (1.9%)
2	37 (38.5%)	22 (35.5%)	18 (33.3%)
不明	2 (2.1%)	2 (3.2%)	2 (3.7%)
悪性度			
悪い	15 (15.6%)	11 (17.7%)	9 (16.7%)
中程度	41 (42.7%)	31 (50.0%)	25 (46.3%)
良好	9 (9.4%)	6 (9.7%)	6 (11.1%)
不明	31 (32.3%)	14 (22.6%)	14 (25.9%)
遠隔転移			
はい	27 (28.1%)	16 (25.8%)	14 (25.9%)
フォローアップ			
中央値(月)	60	70.4	66.5

10

20

【 0 2 0 9】

【表 1 6】

表16. ミュエンスター大学からのMS高リスクおよび低リスク患者における
遠隔転移なしの生存率

群	患者数	リスク群	患者数	5年のDMFS 率(SE)
全ての患者	96	高	48	0.61 (0.072)
		低	48	0.88 (0.052)
全て処置	62	高	32	0.66 (0.084)
		低	30	0.89 (0.060)
タモキシフェンだけの処置	54	高	26	0.65 (0.084)
		低	28	0.88 (0.065)

30

【 0 2 1 0】

【表 17】

表17. ガイ病院からの205人の処置患者の臨床および病理学的特徴

特徴	リンパ節転移陰性ER陽性 (n=205)
閉経状態	
閉経前	31 (15.12%)
閉経前後	4 (1.95%)
閉経後	165 (80.49%)
不明	5 (2.44%)
年齢	
≤ 55歳	74 (36.1%)
> 55歳	131 (63.9%)
平均(標準偏差)	59.3 (10.4)
最小-最大	33 - 86
腫瘍直径	
≤ 2 cm	138 (67.32%)
> 2 cm	67 (32.68%)
平均(標準偏差)	1.67 (1.0)
最小-最大	0 - 3.0
組織学的悪性度	
悪性度1	60 (29.27%)
悪性度2	98 (47.8%)
悪性度3	47 (22.93%)
ステージ	
I	138 (67.32%)
II	67 (32.68%)
サブタイプ	
管状NOS	164 (80.0%)
小葉標準	22 (10.73%)
小葉変異	3 (1.46%)
管状	8 (3.9%)
粘液	6 (2.93%)
乳頭	1 (0.49%)
アポクリン	1 (0.49%)
遠隔再発	
はい	17 (8.29%)
いいえ	188 (91.71%)
いずれかの再発	
はい	17 (8.9%)
いいえ	188 (91.71%)
死亡(任意の原因)	
はい	44 (21.46%)
いいえ	161 (78.54%)
死亡(乳癌)	
はい	16 (7.8%)
いいえ	189 (92.20%)
中央値フォローアップ	9.3年

【 0 2 1 1 】

【表 18】

表18. ガイ処置試料における高、中、および低MS群における5年および10年の遠隔転移なしの生存率

	患者数	DM	5年のDMFS (SE)	10年のDMFS (SE)
高	40	7	0.872 (0.054)	0.804 (0.068)
中	29	2	0.966 (0.034)	0.966 (0.034)
低	136	8	0.970 (0.015)	0.921 (0.028)
中+低	165	10	0.969 (0.014)	0.928 (0.025)
全て	205	17	0.950 (0.015)	0.904 (0.0239)

【 0 2 1 2 】

【表 19】

表19. ガイ処置試料におけるMSリスク群、年齢、腫瘍サイズ、および組織学的悪性度の
一変量および多変量のCox比例ハザード抑制

	一変量			多変量		
	ハザード率	95% CI	P値	ハザード率	95% CI	P値
MS高 対 低リスク	3.25	1.24 - 8.54	0.017	5.82	1.71 - 19.75	0.0047
年齢(1年あたり)	1.03	0.98 - 1.08	0.27	1.03	0.98 - 1.08	0.31
腫瘍サイズ(cmあたり)	1.28	0.77 - 2.14	0.34	1.17	0.68 - 2.03	0.57
悪性度2 対 悪性度1	3.50	0.78 - 15.79	0.10	2.64	0.56 - 12.45	0.22
悪性度3 対 悪性度1	2.72	0.50 - 14.86	0.25	0.62	0.081 - 4.76	0.65

10

【 0 2 1 3 】

【表 20】

表20. ガイの処置患者におけるMSリスク群と、年齢、腫瘍サイズおよび
組織学的悪性度の関連性

年齢				
	N	平均	標準偏差	
高	40	59.3	8.3	
中+低	165	59.3	10.8	
1方向ANOVA		p = 0.34		
腫瘍サイズ				
高	40	1.87	0.92	
中+低	165	1.61	1.01	
1方向ANOVA		p = 0.14		
組織学的悪性度				
	悪性度1	悪性度2	悪性度3	
高	0	9	31	
中+低	60	89	16	
Crammer's V = 0.65		関連性についてのchi-sq試験 p値<0.0001		

【 0 2 1 4 】

【表 2 1】

表21. 年齢、腫瘍サイズ、組織学的悪性度および閉経状態のサブグループにおけるMSリスク群(高 対 低リスク)の性能

	N患者	N DM	HR	L95%CI	U95%CL	P値
全て	205	17	3.25	1.24	8.54	0.017
組織学的悪性度						
悪性度1	60	2	NA		低スコアのみ	
悪性度2	98	11	10.72	3.00	38.36	0.0003
悪性度3	47	4	1.53	0.16	14.66	0.715
腫瘍サイズ						
<= 2cm	138	11	3.66	1.07	12.57	0.0392
> 2cm	67	6	2.99	0.60	14.82	0.18
年齢						
<= 55歳	74	5	7.38	1.00	54.66	0.0503
> 55歳	131	12	2.78	0.88	8.75	0.0813
閉経状態						
閉経前	31	1	NA			
閉経後	165	15	2.84	1.01	7.99	0.0476

10

20

【 0 2 1 5 】

【表 2 2】

表22. ガイの処置試料におけるMSリスク群(高対低リスク)の診断値

	5年	10年
感受性	0.50 (0.24 - 0.76)	0.44 (0.23 - 0.67)
特異性	0.82 (0.76 - 0.87)	0.85 (0.76 - 0.92)
PPV	0.128 (0.068 - 0.227)	0.24 (0.13 - 0.41)
NPV	0.969 (0.944 - 0.983)	0.93 (0.90 - 0.96)

30

【 0 2 1 6 】

【表 2 3】

表23. 234人の日本人試料の臨床病理学的特徴

	閉経後 112 (47.9%)	閉経前 122 (52.1%)	全て 234 (100%)
年齢			
<= 55 歳	32 (28.6%)	122 (100%)	154 (65.8%)
> 55 歳	80 (71.4%)	0 (0%)	80 (34.2%)
平均 (標準偏差)	60.8 (7.8)	44.8 (6.0)	52.5 (10.6)
最小-最大	43 - 81	25 - 54	25 - 81
腫瘍直径			
<= 2 cm	65 (58.0%)	81 (66.4%)	146 (62.4%)
> 2 cm	47 (42.0%)	41 (33.6%)	88 (37.6%)
平均 (標準偏差)	2.15 (1.0)	1.96 (1.2)	2.05 (1.1)
最小-最大	0.3 - 8.4	0.1 - 6.5	0.1 - 8.4
組織学的悪性度			
悪性度 1	28 (25%)	46 (37.7%)	74 (31.6%)
悪性度 2	56 (50%)	57 (46.7%)	113 (48.3%)
悪性度 3	28 (25%)	19 (15.6%)	47 (20.1%)
腫瘍サブタイプ			
II a 乳頭管状	42 (37.5%)	66 (54.1%)	108 (46.2%)
II a 硬性癌	41 (36.6%)	36 (29.5%)	77 (32.9%)
II a 固体管状	24 (21.4%)	14 (11.5%)	38 (16.2%)
II b 浸潤性小葉	2 (1.8%)	2 (1.6%)	4 (1.7%)
II b 延髄	1 (0.9%)	0 (0%)	1 (0.4%)
II b 粘液	2 (1.8%)	4 (3.3%)	6 (2.6%)
ステージ			
I	65 (58.0%)	81 (66.4%)	146 (62.4%)
IIA	44 (39.3%)	36 (29.5%)	80 (34.2%)
IIB	3 (2.7%)	5 (4.1%)	8 (3.4%)
PgR			
+	65 (58%)	113 (92.6%)	178 (76.1%)
-	47 (42%)	9 (7.4%)	56 (23.9%)
療法			
Tam/ Tam 組み合わせ	112 (100%)	102 (83.9%)	214 (91.5%)
ZOL	0 (0%)	20 (16.1%)	20 (8.5%)
遠隔転移			
はい	21 (18.8%)	10 (8.9%)	31 (13.3%)
いいえ	91 (81.3%)	112 (91.1%)	203 (86.8%)
任意の原因による死亡			
はい	10 (8.9%)	9 (7.3%)	19 (8.1%)
いいえ	102 (90.9%)	115 (92.7%)	215 (91.9%)
再発 (局所および遠隔)			
はい	30 (26.8%)	16 (13.1%)	46 (19.7%)
いいえ	82 (73.2%)	106 (86.9%)	188 (80.3%)
フォローアップ (年)			
中央値	9	8	8.7

【 0 2 1 7 】

【表 2 4】

表24. 日本人の患者における異なるMS群の5年および10年の遠隔転移なしの生存率

	患者数	5年のDMFS (SE)	10年のDMFS (SE)
高	95	0.808 (0.042)	0.747 (0.050)
中	62	0.965 (0.024)	0.912 (0.044)
低	77	0.974 (0.018)	0.887 (0.047)
中+低	139	0.97 (0.015)	0.895 (0.034)
全て	234	0.905 (0.020)	0.837 (0.029)

【 0 2 1 8 】

10

20

30

40

【表 2 5】

表25. 日本人の患者におけるMSリスク群についての遠隔転移までの時間、年齢、腫瘍サイズおよび組織学的悪性度の一変量および多変量のCox比例ハザードモデル

	一変量			多変量		
	ハザード率	95% CI	P値	ハザード率	95% CI	P値
MS高 対 中+低	3.32	1.56 - 7.06	0.0018	3.79	1.42 - 10.1	0.0078
年齢(年あたり)	1.04	1.00 - 1.08	0.032	1.03	0.99 - 1.07	0.087
組織サイズ(cmあたり)	1.45	1.14 - 1.83	0.0024	1.4	1.10 - 1.78	0.007
組織学的悪性度2 対 1	1.47	0.60 - 3.61	0.40	0.72	0.24 - 2.14	0.56
組織学的悪性度3 対 1	2.02	0.75 - 5.44	0.16	0.55	0.15 - 2.00	0.36

10

【 0 2 1 9 】

【表 2 6】

表26. 日本人の患者における、MSリスク群と、年齢、腫瘍サイズおよび組織学的悪性度の関連性

年齢	N	平均	標準偏差	
高	95	53.4	10.4	
中+低	139	51.9	10.7	
一方向ANOVA		p = 0.29		
腫瘍サイズ				
高	95	2.23	1.04	
中+低	139	1.93	1.13	
一方向ANOVA		p = 0.037		
組織学的悪性度	悪性度1	悪性度2	悪性度3	
高	4	54	37	
中+低	70	59	10	
Crammer's V = 0.54	関連性についてのchi-aq試験 p値<0. 0001			
サブタイプ	2a乳頭管状	2a硬性癌	2a固体的管状	2b浸潤性小葉
高	45	23	24	1
中+低	63	54	14	3
Crammer's V = 0.25	関連性についてのchi-aq p値=0. 01			

20

30

【 0 2 2 0 】

【表 2 7】

表27. 日本人の患者における、MSリスク群に対する遠隔転移までの時間、閉経状態、腫瘍サイズ、PgR状態および組織学的悪性度の一変量および多変量のCox比例ハザードモデル

	一変量			多変量		
	ハザード率	95% CI	P値	ハザード率	95% CI	P値
MS高 対 中+低	3.32	1.56 - 7.06	0.0018	3.44	1.27 - 9.34	0.015
Tam前 対 Tam後	0.33	0.13 - 0.81	0.016	0.45	0.17 - 1.19	0.11
ZOL前 対 Tam後	1.22	0.42 - 3.58	0.72	1.8	0.55 - 5.81	0.33
組織 サイズ(cmあたり)	1.45	1.14 - 1.83	0.0024	1.45	1.12 - 1.88	0.0049
PgR (-ve 対 +ve)	2.3	1.13 - 4.7	0.022	1.7	0.75 - 3.86	0.2
組織学的悪性度2 対 1	1.47	0.6 - 3.61	0.4	0.53	0.17 - 1.62	0.26
組織学的悪性度3 対 1	2.02	0.75 - 5.44	0.16	0.49	0.13 - 1.76	0.27

10

【 0 2 2 1】

【表 2 8】

表28. サブグループ分析: 腫瘍サイズ、年齢、閉経状態、組織学的悪性度およびPgR状態の異なるサブグループにおける遠隔転移までの時間(DMFS)のハザード率

20

層(Strata)	ハザード率	95% CI	P値
全て	3.32	1.56 - 7.06	0.0018
腫瘍 ≤ 2cm	4.48	1.16 - 17.34	0.030
腫瘍 > 2cm	2.27	0.92 - 5.62	0.077
年齢 ≤ 55	4.03	1.38 - 11.81	0.011
年齢 > 55	2.34	0.81 - 6.74	0.12
閉経後	2.06	0.83 - 5.09	0.12
閉経前	6.01	1.55 - 23.25	0.0094
悪性度1 & 2	3.57	1.52 - 8.35	0.0034
悪性度3	2.35	0.29 - 18.77	0.42
PgR +	3.48	1.35 - 9.0	0.0099
PgR -	2.06	0.57 - 7.49	0.27

30

40

【 0 2 2 2】

【表 29】

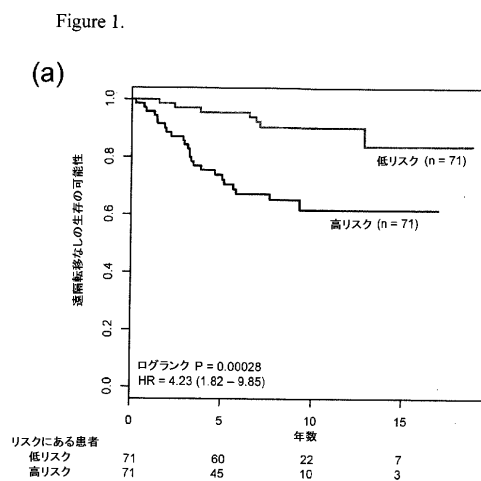
表29. 日本人の試料に関する5年以内の遠隔転移を予測するためのMSの診断値

カット1(中+低の組み合わせ)

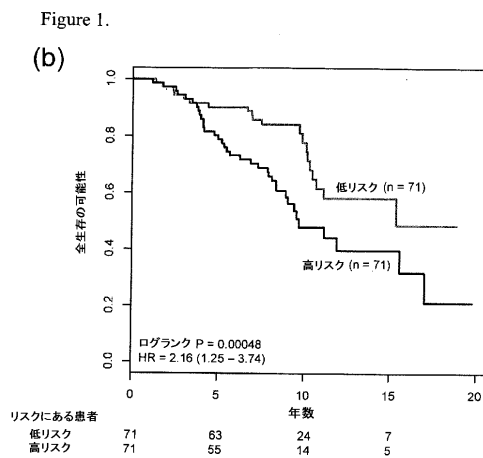
感受性	0.81 (0.60 - 0.92)
特異性	0.65 (0.58 - 0.71)
PPV	0.19 (0.15 - 0.24)
NPV	0.97 (0.93 - 0.99)

10

【図 1 a】

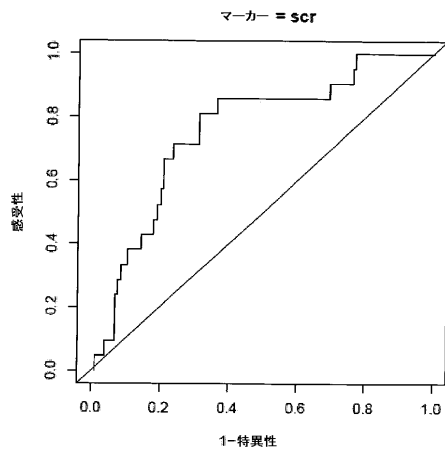


【図 1 b】



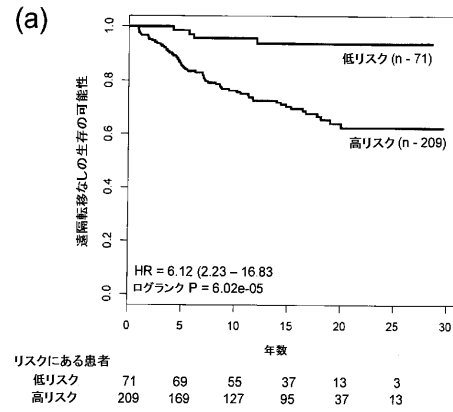
【図 2】

Figure 2.



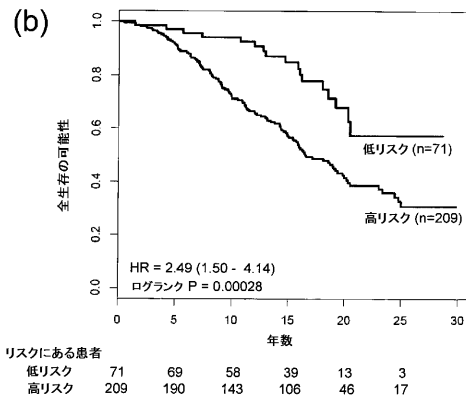
【図 3 a】

Figure 3.



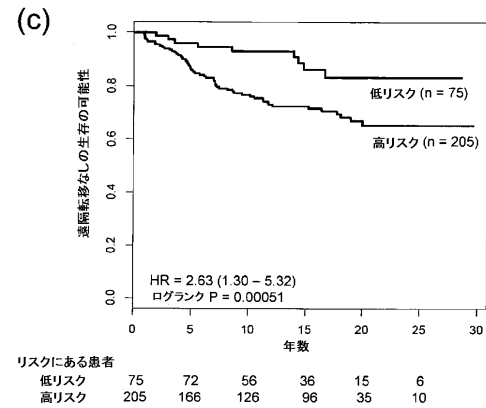
【図 3 b】

Figure 3.

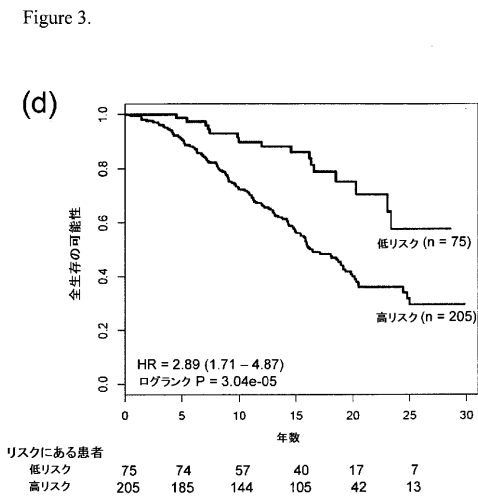


【図 3 c】

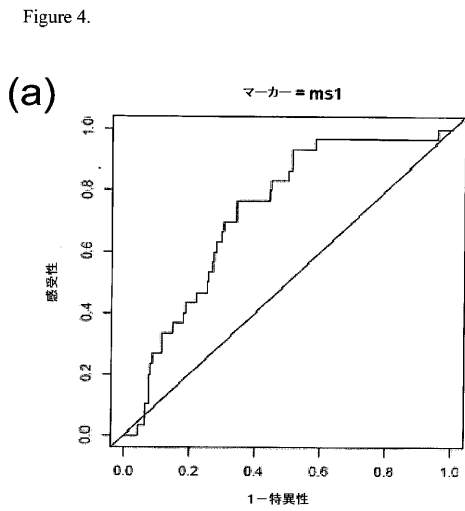
Figure 3.



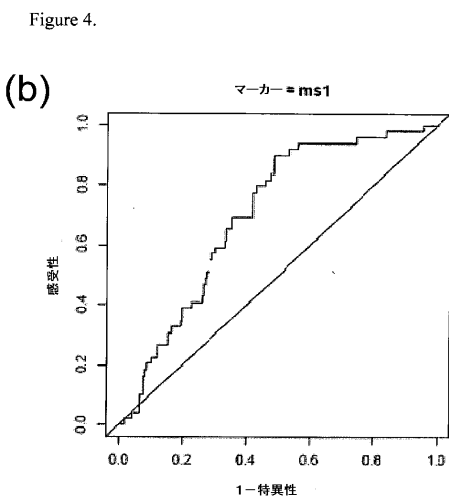
【図 3 d】



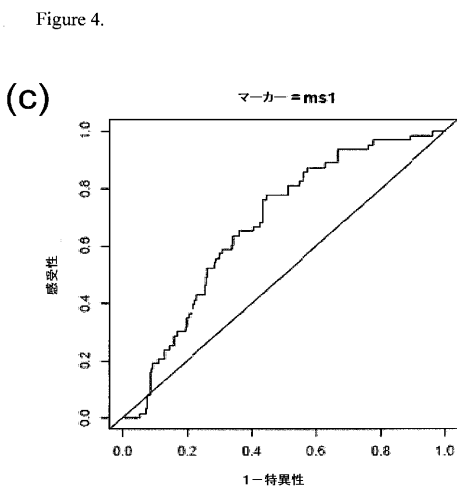
【図 4 a】



【図 4 b】

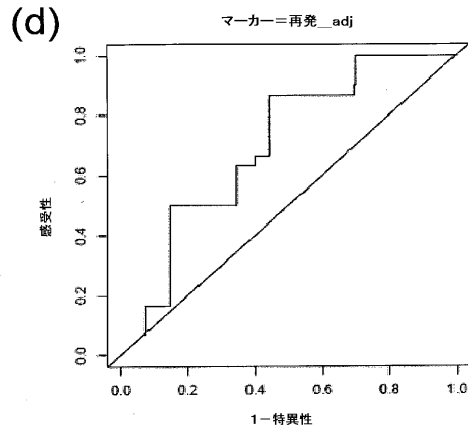


【図 4 c】



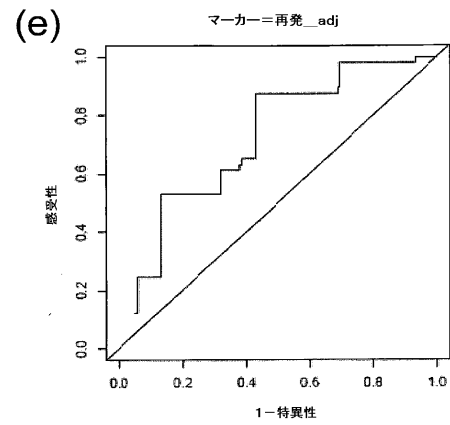
【図 4 d】

Figure 4.



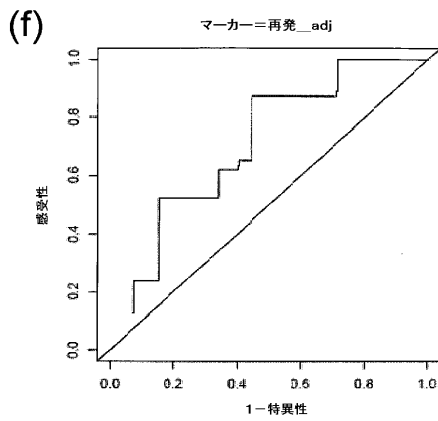
【図 4 e】

Figure 4.



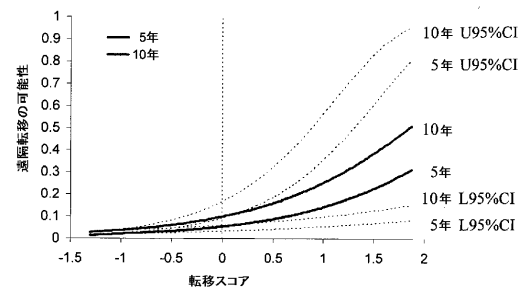
【図 4 f】

Figure 4.



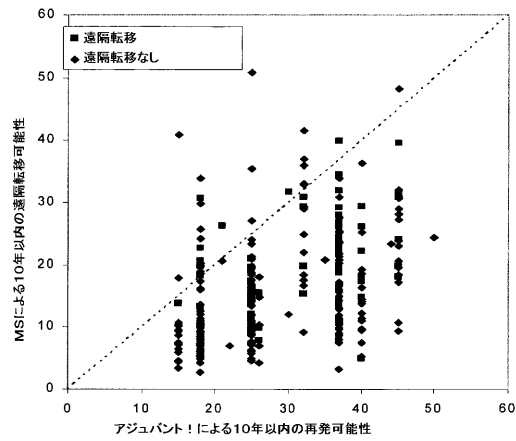
【図 5】

Figure 5.



【図 6】

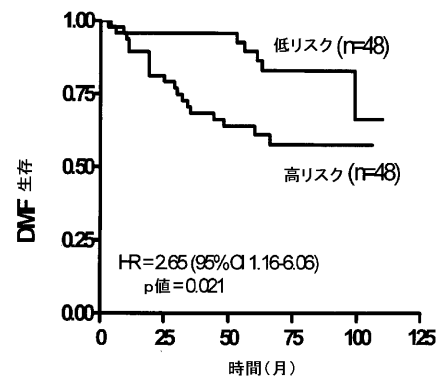
Figure 6.



【図 7 A】

Figure 7.

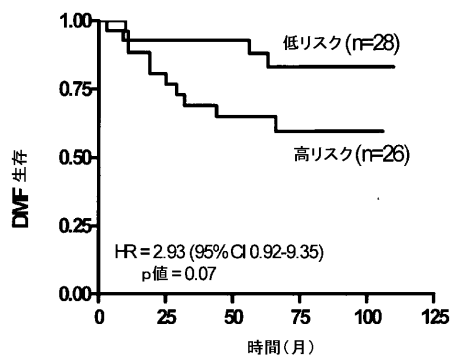
(A) 全ての患者



【図 7 B】

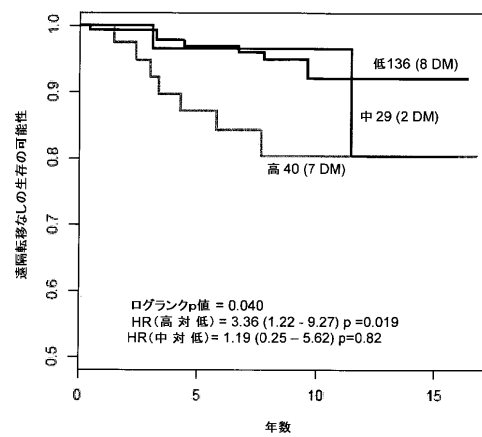
Figure 7.

(B) タモキシフェン処置



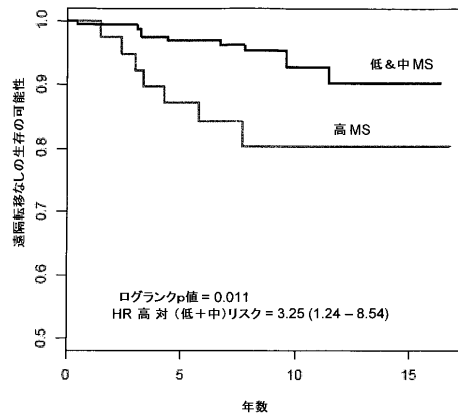
【図 8】

Figure 8.



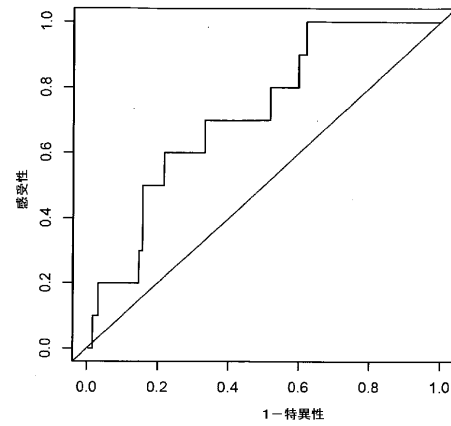
【図 9】

Figure 9.



【図 10】

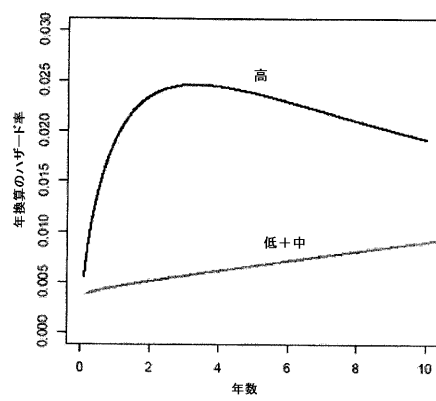
Figure 10.



【図 11 A】

Figure 11.

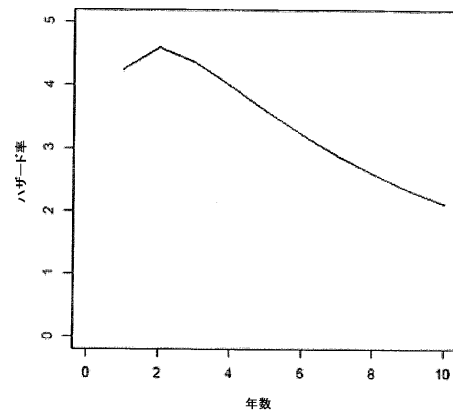
(A)



【図 11 B】

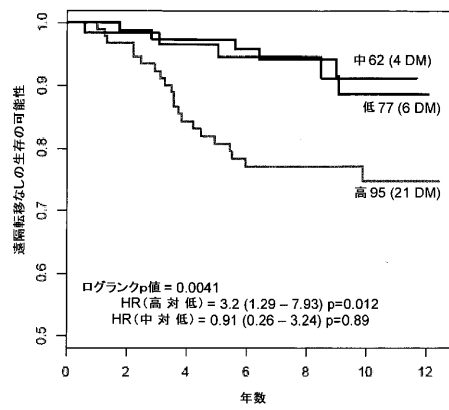
Figure 11.

(B)



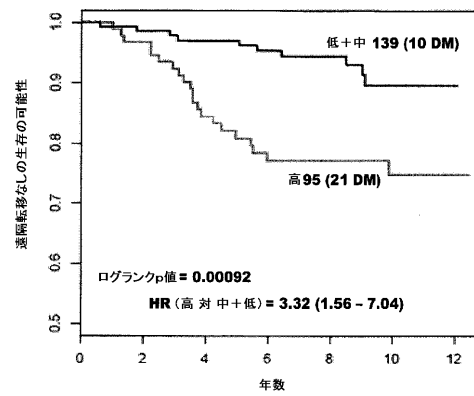
【図 12】

Figure 12.



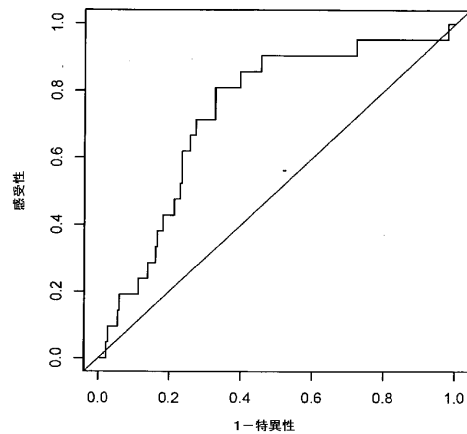
【図 13】

Figure 13.



【図 14】

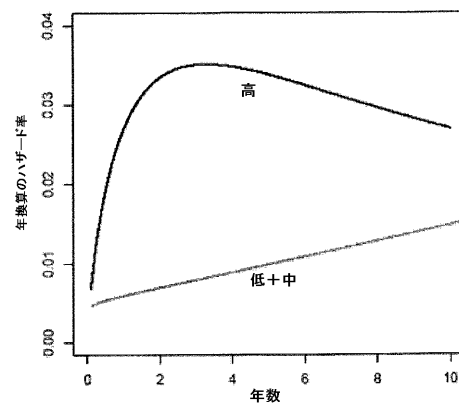
Figure 14.



【図 15 A】

Figure 15.

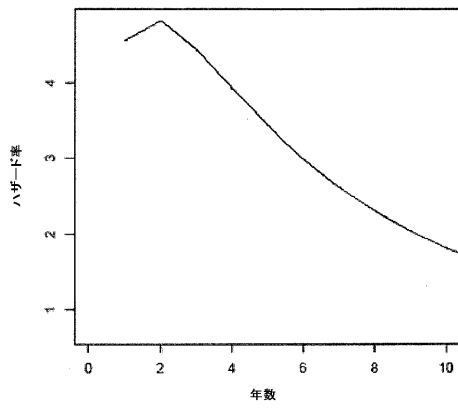
(A)



【図 15 B】

Figure 15.

(B)



【配列表】

0005913785000001.app

フロントページの続き

(51)Int.Cl. F I
G 0 1 N 37/00 1 0 2

(72)発明者 ワン, アリス
アメリカ合衆国 カリフォルニア 9 4 5 0 2, アラメダ, ハーバー ベイ パークウェイ
1 4 0 1, アン アブレラ ビジネス ユニット, セレラ

合議体

審判長 鈴木 恵理子

審判官 三原 健治

審判官 長井 啓子

(56)参考文献 N . E n g l . J . M e d . , 2 0 0 4 , V o l . 3 5 1 , p . 2 8 1 7 - 2 8 2 6
N a t u r e , 2 0 0 2 , V o l . 4 1 5 , p . 5 3 0 - 5 3 6
C a n c e r R e s . , 2 0 0 5 , V o l . 6 5 , N o . 1 0 , p . 4 0 5 9 - 4 0 6 6

(58)調査した分野(Int.Cl. , D B 名)

C12N 15/00-15/90

C12Q 1/68

C A p l u s / B I O S I S / M E D L I N E / E M B A S E / W P I D S (S T N)