

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2020-150949
(P2020-150949A)

(43) 公開日 令和2年9月24日(2020.9.24)

(51) Int.Cl.	F I	テーマコード (参考)
C 1 2 Q 1/6886 (2018.01)	C 1 2 Q 1/6886	Z 2 G 0 4 5
G O 1 N 33/574 (2006.01)	G O 1 N 33/574	A 4 B 0 2 9
G O 1 N 33/50 (2006.01)	G O 1 N 33/50	P 4 B 0 6 3
C 1 2 M 1/00 (2006.01)	C 1 2 M 1/00	A
C 1 2 N 15/12 (2006.01)	C 1 2 N 15/12	

審査請求 有 請求項の数 1 O L 外国語出願 (全 37 頁)

(21) 出願番号 特願2020-97666 (P2020-97666)
 (22) 出願日 令和2年6月4日(2020.6.4)
 (62) 分割の表示 特願2018-170779 (P2018-170779)
 の分割
 原出願日 平成20年5月23日(2008.5.23)
 (31) 優先権主張番号 555363
 (32) 優先日 平成19年5月24日(2007.5.24)
 (33) 優先権主張国・地域又は機関
 ニュージーランド(NZ)

(71) 出願人 504350120
 パシフィック エッジ リミテッド
 ニュージーランド国、ダニディン、モレー
 ・プレイス 481、オタゴ・ハウス、8
 ティーエイチ・フロアー、シー/オー・デ
 ロイト
 (71) 出願人 503412377
 ルートヴィヒ・インスティテュート・フォ
 ー・キャンサー・リサーチ
 スイス国8001 チューリヒ、スタデル
 ホッフエル シュトラーセ 22
 (74) 代理人 100078662
 弁理士 津国 肇

最終頁に続く

(54) 【発明の名称】 メラノーマ癌の予後予測

(57) 【要約】 (修正有)

【課題】患者における癌、特にメラノーマ、の予後を決定するための予後マーカー、予後シグネチャー、組成物および方法の提供。

【解決手段】2つ以上のメラノーマ予後マーカー(MPM)を含む、メラノーマの進行の危険を決定するための予後シグネチャー、表面に1つ以上のロケーションを有する基材を含み、各ロケーションはそこに2つ以上のオリゴヌクレオチドを有し、各オリゴヌクレオチドは1つ以上のMPMから選ばれる、メラノーマの予後を決定するための装置、および薬物試験に対するメラノーマ患者の適性を決定するための方法。

【選択図】なし

【特許請求の範囲】**【請求項 1】**

2つ以上のメラノーマ予後マーカー（MPM）を含む、メラノーマの進行の危険を決定するための予後シグネチャー。

【請求項 2】

MPMが表1から選ばれる、請求項1に記載のシグネチャー。

【請求項 3】

表面に1つ以上のロケーションを有する基材を含み、各ロケーションはそこに2つ以上のオリゴヌクレオチドを有し、各オリゴヌクレオチドは1つ以上のMPMから選ばれる、メラノーマの予後を決定するための装置。

10

【請求項 4】

前記2つ以上のオリゴヌクレオチドが表1から選ばれるMPMである、請求項3に記載の装置。

【請求項 5】

患者におけるメラノーマの予後を決定するための方法であって、下記工程：

(i) 患者からのメラノーマ腫瘍サンプルにおいて、MPMの発現レベルまたは2つ以上のMPMを含む予後シグネチャーの発現レベルを決定する工程、

(ii) 予後が良好な腫瘍サンプルおよび予後が悪い腫瘍サンプルにおけるMPMの発現レベルまたは予後シグネチャーの発現レベルに予測方法をアプライすることにより確立された予測モデルをアプライする工程、

20

(iii) 予後を確立する工程、

を含む方法。

【請求項 6】

薬物試験に対するメラノーマ患者の適性を決定するための方法であって、下記工程：

(i) 患者からのメラノーマ腫瘍サンプルにおいて、MPMの発現レベルまたは2つ以上のMPMを含む予後シグネチャーの発現レベルを決定する工程、

(ii) 予後が良好な腫瘍サンプルおよび予後が悪い腫瘍サンプルにおけるMPMの発現レベルまたは予後シグネチャーの発現レベルに予測方法をアプライすることにより確立された予測モデルをアプライする工程、

30

(iii) 試験に対する患者の適合性を確立する工程、

を含む方法。

【請求項 7】

MPMが表1から選ばれる、請求項5または請求項6に記載の方法。

【請求項 8】

前記予測方法が、線形モデル、サポートベクターマシン、ニューラルネットワーク、分類および回帰木、アンサンブル学習方法、判別分析、最近傍法、ベイジアンネットワーク、独立成分分析からなる群より選ばれる、請求項5に記載の方法。

【請求項 9】

MPMまたは予後シグネチャーの発現レベルを決定する工程が、各遺伝子のmRNAの発現レベルを検出することにより行われる、請求項5～8のいずれかに記載の方法。

40

【請求項 10】

MPMまたは予後シグネチャーの発現レベルを決定する工程が、各遺伝子のcDNAの発現レベルを検出することにより行われる、請求項5～8のいずれかに記載の方法。

【請求項 11】

MPMまたは予後シグネチャーの発現レベルを決定する工程が、前記cDNAの少なくとも一部に対して相補的なヌクレオチドを使用して行われる、請求項10に記載の方法。

【請求項 12】

MPMまたは予後シグネチャーの発現レベルを決定する工程が、フォワードプライマーおよびリバースプライマーを使用するqPCR法を使用して行われる、請求項9に記載の方法。

50

【請求項 13】

M P Mまたは予後シグネチャーの発現レベルを決定する工程が、請求項 3 または請求項 4 に記載の装置を使用して行われる、請求項 8 に記載の方法。

【請求項 14】

M P Mまたは予後シグネチャーの発現レベルを決定する工程が、各マーカ-のタンパク質の発現レベルを検出することにより行われる、請求項 5 ~ 8 のいずれかに記載の方法。

【請求項 15】

M P Mまたは予後シグネチャーの発現レベルを決定する工程が、各マーカ-のペプチドの発現レベルを検出することにより行われる、請求項 5 ~ 8 のいずれかに記載の方法。

【請求項 16】

前記検出する工程が、各マーカ-に向けられた抗体を使用して行われる、請求項 13 または請求項 14 に記載の方法。

【請求項 17】

前記検出する工程がサンドイッチ型イムノアッセイ法を使用して行われる、請求項 13 ~ 15 のいずれかに記載の方法。

【請求項 18】

前記抗体がモノクローナル抗体である、請求項 13 ~ 16 のいずれかに記載の方法。

【請求項 19】

前記抗体がポリクローナル抗血清である、請求項 13 ~ 16 のいずれかに記載の方法。

【発明の詳細な説明】

【技術分野】

【0001】

関連出願

本出願は、2007年5月24日に提出されたニュージーランド仮特許出願第555363号の利益を主張し、それは全体において参照によって組み込まれる。

【0002】

本発明は、患者における癌、特にメラノーマの進行を決定するための方法および組成物に関する。詳しくは、本発明は、予後シグネチャー(prognostic signatures)に基づいて、メラノーマなどの癌の予後を決定するための遺伝子マ-およびプロテオームマ-の使用に関する。

【背景技術】

【0003】

先進国では、メラノーマの発生率はこれまでの25年間にわたり絶えず上昇しており、オーストラリアでの発生率が世界で最も高い¹。認められた「メラノーマ流行」は、多分薄いメラノーマの増加した検出を表すけれども²、メラノーマは、主としてより若年のグループ冒して実りある人生の年月を損失させる。この年月は児童悪性疾患および精巣癌によってのみ超えられる^{3, 4}。メラノーマは細胞傷害性化学療法⁵、生物学的作用物質^{6, 7}および種々のワクチン接種ストラテジー⁸に殆ど応答しない。小さなサブグループの患者は、生物学的および/または細胞傷害性化学療法から利益を受けると思われるが、これらの患者を推測的に鑑別することは現在不可能であり、これは、低い利益の確率で実質的な毒性に多くの患者をさらすことを必要とする。

【0004】

メラノーマが一旦局所的リンパ節に転移すると、患者の70%は5年以内に死ぬであろう⁹。長い生存を有するサブグループの患者は独特のコホ-トを表す。現在のアジュバント治療は、全般的生存利益を与えず、そしていくらかの臨床医は、無病生存を改良するためにインターフェロンを与えるが¹⁰、多くの国際的な総合施設は、臨床試験の範囲外では活発なアジュバント処置を与えない。どの患者がアジュバント治療の使用にかかわらず良くなりそうかを予測することは、不必要な毒性を防止し、そしてより多く利益を得るような患者をターゲティングする、より良好な治療ストラテジーの開発を可能とする。アジュバント臨床試験における患者のより良好な階層化は、第一種の過誤および第二種の過誤

10

20

30

40

50

の両方を減少させるであろう。ECOG 1684 研究の後の12年の更新および他の無作為化研究は、インターフェロンが、TTPを改良するが、段階IIIメラノーマにおける全生存期間(overall survival)を改良しないことを証明した^{5, 10, 11}。現在よく認識されているがコントロールされることができない患者母集団内の固有の不均一性は、初期のECOG 1684 研究¹⁰ および他のより少ない第II相の研究において見られた生存に関する有望な効果を混乱させたかもしれない。再発の可能性のより高いこれらの患者を階層化することは、この不均一性をバランスさせそして処置がより正確に比較されることを可能とすることができる。

【0005】

メラノーマの予後を予測するための更なるツールに対する要求がある。本発明は、癌の予後および処置を助けるための、予後癌マーカー、特にメラノーマ予後マーカーに基づく方法、組成物、キットおよび装置を提供する。

【0006】

発明の要約

ある態様では、予後の良好なメラノーマおよび予後の悪い(poor)メラノーマにおいてディファレンシャルに発現されることが同定されたマーカー遺伝子のセットが提供される。この遺伝子のセットは、患者におけるメラノーマの進行速度を予測することができる2つ以上のマーカーを含む予後シグネチャーを発生させるのに使用することができる。

【0007】

個々のマーカーは、腫瘍が速く進行するか速く進行しないかに依存して、ディファレンシャルに発現されうる。マーカーを組み合わせて、単一遺伝子アッセイよりもはるかに有効な個々の検定を提供することができる予後シグネチャーとすることにより、予測の正確度を高めることができる。予測モデルを発生させるために統計学、機械学習、人口知能およびデータマイニングなどの技術を予後シグネチャーにアプリケーションすることも提供される。他の態様では、次いで患者の腫瘍における特定の予後シグネチャーのマーカーの発現レベルを予測モデルにアプリケーションして予後を決定することができる。

【0008】

ある態様では、マーカーの発現レベルは、マイクロアレイ法、定量的ポリメラーゼ連鎖反応(qPCR)またはイムノアッセイを使用して確立することができる。

【0009】

特に、本発明は、患者におけるメラノーマの予後を決定するための方法であって、
 (i) 患者からのメラノーマ腫瘍サンプルにおいてメラノーマ予後マーカー(MPM)の発現レベルまたは2つ以上のMPMを含む予後シグネチャーの発現レベルを決定し、
 (ii) 予後が良好な腫瘍サンプルおよび予後が悪い腫瘍サンプルにおけるMPMの発現レベルまたは予測シグネチャー(predictive signature)の発現レベルに予測方法をアプリケーションすることにより確立された予測モデルをアプリケーションし、
 (iii) 予後を確立する工程を含む方法を提供する。

【0010】

あるいは、本発明は、薬物試験に対するメラノーマ患者の適性を決定するための方法であって、

(i) 患者からのメラノーマ腫瘍サンプルにおいてMPMの発現レベルまたは2つ以上のMPMを含む予後シグネチャーの発現レベルを決定し、
 (ii) 予後が良好な腫瘍サンプルおよび予後が悪い腫瘍サンプルにおけるMPMの発現レベルまたは予測シグネチャーの発現レベルに予測方法をアプリケーションすることにより確立された予測モデルをアプリケーションし、
 (iii) 試験に対する患者の適性を確立する、工程を含む方法も提供する。

【0011】

本方法に従うMPMは、表1から選択されうる。予測方法は、線形モデル、サポートベ

10

20

30

40

50

クターマシン、ニューラルネットワーク、分類および回帰木、アンサンブル学習法、判別分析(discriminant analysis)、最近傍法(nearest neighbor method)、ベイジアンネットワーク、独立成分分析からなる群より選ばれる。

【0012】

M P Mの発現レベルまたは予後シグネチャーの発現レベルを決定することは、例えばフォアードプライマーおよびリバースプライマーを使用するq P C Rを使用して各遺伝子のm R N Aの発現レベルを検出することにより行うことができる。M P Mの発現レベルまたは予後シグネチャーの発現レベルを決定することは、各遺伝子のc D N Aの発現レベルを検出することにより、例えば各遺伝子のc D N Aの少なくとも一部に対して相補的なヌクレオチドを使用することにより行うこともできる。更に、M P Mの発現レベルまたは予後シグネチャーの発現レベルは、各マーカーのタンパク質の発現レベルを検出することまたは各マーカーのペプチドの発現レベルを検出すること、例えば、各マーカーに対して向けられた抗体、例えばモノクローナル抗体またはポリクローナル抗血清を使用することにより決定することができる。サンドイッチ型イムノアッセイ法またはE L I S Aアッセイを使用することができる。

10

【0013】

本発明は、2つ以上のメラノーマ予後マーカー(M P M)を含むメラノーマの進行の危険を決定するための予後シグネチャーも提供する。予後シグネチャーのM P Mは表1から選ぶことができる。

【0014】

他の局面では、本発明は、メラノーマの予後を決定するための装置であって、表面に1つ以上のロケーション有する基材を含み、各ロケーションはそこに2つ以上のオリゴヌクレオチドを有し、各オリゴヌクレオチドは1つ以上のM P Mから選ばれる、装置を提供する。

20

【0015】

2つ以上のオリゴヌクレオチドは表1から選ばれるM P Mであることができる。

【0016】

本発明は、患者におけるメラノーマの予後を予測するためのキットの製造において、M P Mの発現または2つ以上のM P Mを含む予後シグネチャーの発現を検出するための試薬の使用も提供する。M P Mは表1から選択されうる。

30

【0017】

試薬は、M P M m R N AまたはM P M c D N Aの発現を検出することにより1つ以上のM P Mの発現のレベルを検出することができる。試薬は、M P M m R N AまたはM P M c D N Aの少なくとも一部に対して相補的なオリゴヌクレオチドであることができる。または、試薬は、M P Mタンパク質またはペプチドの発現を検出することにより1つ以上のM P Mの発現のレベルを検出することができる。試薬は、抗体、例えば、モノクローナル抗体またはポリクローナル抗血清であることができる。

【0018】

キットは、サンドイッチ型イムノアッセイまたはE L I S Aアッセイを引き受けるために適当であることができる。

40

本発明の具体的態様及び図に関して記載する。

【図面の簡単な説明】

【0019】

【図1】予測スコアを構築するのに使用された22の遺伝子(「メラノーママーカー」)を示す。遺伝子はM a n n - W h i t n e y 検定を使用して選ばれた。

【図2】図2は、ディファレンシャルに発現された遺伝子の遺伝子オンロジーグループリングおよび関連した有意性を示す。最も有意なオンロジーは、カテゴリー間で重なる遺伝子の数、即ち、この多くの遺伝子が遺伝子リストおよびカテゴリーの両方にあることが偶然の一致(co-incidence)である尤度により決定される。

【図3】図3は、トレーニングセットおよびq P Sを使用するバリデーションセットAお

50

よび q P S を使用するセット B に適用された 2 つの独立したバリデーションセットを含む実験スキームである。トレーニングセットを使用して予測遺伝子を開発し、次いでこれを q P S を使用するバリデーションセット A および q P S を使用するセット B にアプリケーションした。

【図 4】図 4 は、アレイ実験および q P C R アッセイにおけるコンパレーターとしての両方に使用された参照 c D N A を創生するのに使用された R N A を示す。

【図 5】Universal Probe Library Probes を使用する q P C R のために使用されたアッセイを示す。

【図 6】検定セットおよびバリデーションセット A についての患者の特徴を示す。

【図 7】15 の遺伝子の悪い（空のボックス）予後グループから良好な（充填されたボックス）予後グループを分離する能力を証明する、すべての遺伝子（A）およびディファレンシャルに発現された遺伝子（B）を使用する主成分分析を示す。これらの遺伝子は、アレイおよび q P C R ベースの予測子を発生させるために使用された。

【図 8】トレーニングセットにおける a P S（a - b）および q P S（c - d）の適用を示し、T T P および全生存期間とのその相関を証明する。a P S は、アレイデータと q P C R データ間で最も強い相関を有する 15 の遺伝子のみを使用し、そして q P S は、2 つのグループを分離する最も強い能力を有する 5 つの遺伝子を使用した。

【図 9】図 9 は、トレーニングセットおよびバリデーションセット A に適用された q P S ロジスティック回帰アルゴリズムを示す。水平線は、平均値で引かれている。

【図 10】図 10 は、第 3 の独立したセットの良好な予後グループと悪い予後グループからの q P S スコアの分布を示す。

【0020】

詳細な説明

定義

本発明の態様を詳細に説明する前に、本明細書で使用される用語のいくつかの定義を与えることは有用であろう。

【0021】

「マーカー」という用語は、生物学的現象の存在と定量的または定性的に関連している分子を指す。「マーカー」の例は、現象の基礎をなす機構に直接または間接に関係しているようにいまいが、ポリヌクレオチド、例えば遺伝子もしくは遺伝子フラグメント、R N A もしくは R N A フラグメント；ポリペプチド、例えばペプチド、オリゴペプチド、タンパク質もしくはタンパク質フラグメントを含む遺伝子産物；または任意の関連した代謝物、副生物もしくは任意の他の同定分子、例えば抗体もしくは抗体フラグメントを含む。本発明のマーカーは、本明細書で開示されたヌクレオチド配列（例えば、Genbank 配列）、特に、完全長配列、任意のコード配列、その任意のフラグメントもしくは任意の相補体および上記したその任意の測定可能なマーカーを含む。

【0022】

用語「M P M」または「メラノーマ予後マーカー」または「M P M ファミリーメンバー」は、特定の予後、例えば本明細書に記載された、より進んだ段階への癌進行の高い尤度もしくはより低い尤度と関連している変化した発現を伴うマーカーを指すが、メラノーマの予後と関連していることが先行技術で知られている分子を除外することができる。用語 M P M は、マーカーがメラノーマに対してのみ特異的であることを必要としないと理解されるべきである。むしろ、M P M の発現は、悪性腫瘍を含む他のタイプの腫瘍においては変わることができる。

【0023】

用語「予後シグネチャー」「シグネチャー」等は、セットとして一緒に分析されるとき、事象、例えばメラノーマの予後アウトカムの決定または予測を可能とする 2 つ以上のマーカー例えば M P M のセットを指す。2 つ以上のマーカーを含むシグネチャーの使用は、個々の変化の効果を減少させそしてより強い予測 (robust prediction) を可能とする。M P M の非限定的例は X X で示される。本発明に関しては、任意の特定のセット（例えば任

10

20

30

40

50

意のシグネチャー)に列挙されたマーカーの「少なくとも1つ」、「少なくとも2つ」、「少なくとも5つ等という表現は、列挙されたマーカーの任意の1つまたは任意のおよびすべての組み合わせを意味する。

【0024】

用語「予測方法」は、予測モデルを特定するのに使用することができる統計学、機械学習、人口知能およびデータマイニングの分野からのより広い種類の方法を包含することが定義される。この用語は、アウトカムを予測するために適当な任意の方法も含み、そして、アウトカムを予測するために、多重マーカーの複雑な分析を使用するのみならず、コントロール組織の発現に対するまたは所定の閾値に対する単一マーカーもしくはシグネチャーの発現の直接の比較も使用する方法を含む。これらは、詳細な説明の節にさらに検討される。

10

【0025】

用語「予測モデル」は、予測方法をデータのコレクションに適用することにより得られた特定の数学的モデルを指す。本明細書に詳述された実施例において、このようなデータセットは、各サンプルのクラス(良好または悪い)が知られている、良好なもしくは悪い予後を有するメラノーマ患者から採取された組織サンプルにおける遺伝子活性の測定値からなる。このようなモデルは、(1)未知の予後状態のサンプルを良好または悪いの1つであるとして分類するためまたは(2)未知のサンプルにおける特定された遺伝子のコレクションの、mRNA発現レベルまたは発現産物の測定値に基づいて、未知のサンプルが良好な予後を有する尤度を表す確率的予測をする(即ち、例えば確率と解釈されるべき割合または百分率を作り出す)ために使用されうる。いかにしてこれらの遺伝子特異的測定値が組み合わされて分類および確率的予測をするかの正確な詳細は、モデルを構築するのに使用される予測方法の特定の機序に依存している。この用語は、アウトカムを予測するのに適当な任意のモデルも含み、そしてアウトカムを予測するために、多重マーカーの複雑な分析を使用するモデルのみならず、単一マーカーまたはシグネチャーの発現をコントロール組織の発現または所定の閾値と直接比較することを伴うモデルも含む。

20

【0026】

「感受性」、「特異性」(または「選択性」)、および「分類率」(classification rate)は、予測モデルの有効性を説明するのに適用されるとき、下記のことを意味する:

「感受性」は、ポジティブであることが(モデルにより)予測もされる真にポジティブなサンプルの割合を意味する。メラノーマの予後の検定においては、それは、良好であることがモデルにより予測された良好な予後を有する腫瘍の割合であろう。「特異性」または「選択性」は、ネガティブであることが(モデルによって)予測もされる真にネガティブなサンプルの割合を意味する。メラノーマの予後の検定においては、これは、モデルにより悪いことが予測される悪い予後を有するサンプルの割合に等しい。「分類率」は、予測モデルにより正確に分類されるすべてのサンプルの割合である(ポジティブまたはネガティブとしてのそれである)。

30

【0027】

本明細書で使用された「抗体」および同様な用語は、免疫グロブリン分子および免疫グロブリン(Ig)分子の免疫学的に活性な部分、即ち、抗原に特異的に結合する(と免疫反応する)抗原結合部位を含有する分子を指す。これらは、ポリクローナル、モノクローナル、キメラ、一本鎖、Fc、Fab、Fab'およびFab₂フラグメントおよびFab発現ライブラリーを含むが、それらに限定されない。抗体分子は、分子中に存在する重鎖の性質により互いに異なるクラスIgG、IgM、IgA、IgEおよびIgDのクラスのいずれかに関する。これらは、IgG1、IgG2およびその他などのサブクラスも含む。軽鎖は、カッパ鎖またはラムダ鎖であることができる。本明細書での抗体への言及は、すべてのクラス、サブクラスおよびタイプへの言及を含む。キメラ抗体、例えば、1つより多くのソース、例えばマウスまたはヒト配列に特異的なモノクローナル抗体またはそのフラグメントである。更にラクダ抗体、サメ抗体またはナノボディも含まれる。

40

「0000」

50

用語「癌」および「癌性」(cancerous)は、典型的には異常なまたは調節されていない細胞増殖により特徴付けられる哺乳動物における生理学的状態を指すかまたは説明する。癌および癌病理学は、例えば、転移、隣接細胞の正常な機能を果たすことの妨害、異常なレベルでのサイトカインまたは他の分泌産物の放出、炎症性または免疫学的応答の抑制または悪化、新形成、前悪性疾患、悪性疾患、周囲のまたは離れている組織または器官、たとえばリンパ節等の浸潤と関連していることができる。特にメラノーマが含まれる。

【0028】

用語「メラノーマ」は、皮膚に見出されるが、口および肛門性器粘膜表面、食道、髄膜および眼などの他の部位にも見出されるメラノサイトから生じる腫瘍を指す。これらの腫瘍は任意の器官に転移することができる。

10

【0029】

用語「ディファレンシャルに発現された」、「ディファレンシャルな発現」および同様なフレーズは、コントロール被検体(例えば参照サンプル)におけるその発現と比べて、状態(condition)、特に癌、例えばメラノーマを有する被検体(例えば、試験サンプル)においてより高いレベルまたはより低いレベルにその発現が活性化される遺伝子マーカーを指す。この用語は、同じ状態の異なる段階において;良好な予後または悪い予後を有する疾患において;またはより高いレベルもしくはより低いレベルの増殖を有する細胞において、より高いレベルまたはより低いレベルにその発現が活性化されるマーカーも含む。ディファレンシャルに発現されたマーカーは、ポリヌクレオチドレベルまたはポリペプチドレベルで活性化または抑制され得るか、または選択的スプライシングを受けて異なるポリペプチド産物をもたらすことができる。このような差は、例えば、mRNAレベルの変化、ポリペプチドの表面発現、分泌または他の分配の変化により証明されうる。

20

【0030】

「ディファレンシャルな発現」は、2つ以上のマーカー(例えば、遺伝子またはそれらの遺伝子産物)間の発現の比較;または2つ以上のマーカー(例えば、遺伝子またはそれらの遺伝子産物)間の発現の割合の比較;または同じマーカーの2つの異なってプロセシングされた産物(例えば転写物またはポリペプチド)の比較、を含むことができ、これらは、正常な被検体と疾患のある被検体間で異なるか、もしくは同じ疾患の種々の段階間で異なるか、もしくは良好な予後または悪い予後を有する疾患間で異なるか、もしくはより高いレベルの増殖とより低いレベルの増殖を有する細胞間で異なるか、もしくは正常な組織と罹病組織、例えば癌またはメラノーマとの間で異なる。ディファレンシャルな発現は、例えば正常細胞と罹病細胞間の、もしくは異なる疾患事象もしくは疾患段階を経た細胞間の、もしくは異なるレベルの増殖を有する細胞間の、遺伝子またはその発現産物における一過性または細胞性発現パターンの定量的差および定性的差の両方を含む。

30

【0031】

用語「発現」は、遺伝子または遺伝子の一部からのポリヌクレオチドおよびポリペプチドの産生、特にRNA(例えばmRNA)の産生を含み、そしてRNA、または遺伝子もしくは遺伝子の一部によりコードされたポリペプチドの産生、および発現と関連した検出可能な物質の出現を含む。例えば、ポリペプチド-ポリペプチド相互作用、ポリペプチド-ヌクレオチド相互作用または同様なものからの、例えば複合体の形成は、用語「発現」の範囲内に含まれる。他の例は、遺伝子または他のポリヌクレオチドまたはオリゴヌクレオチド、ポリペプチドまたはタンパク質フラグメントへの、ハイブリダイゼーションプロブもしくは抗体などの結合性リガンドの結合および結合性リガンドの可視化である。したがって、マイクロアレイ上のスポットの強度、ノーザンプロットなどのハイブリダイゼーションプロット上の強度、またはウエスタンプロットなどのイムノプロット上の強度、またはビーズアレイ上の強度またはPCR分析による強度は、基礎をなす生物学的分子の用語「発現」内に含まれる。

40

【0032】

用語「発現閾値」および「規定された発現閾値」は、相互に交換可能に使用され、そして問題のマーカーのレベルであって、そのレベルの外側ではポリヌクレオチドまたはポリ

50

ペプチドが患者生存の予測マーカーとして働く、マーカーのレベルを指す。閾値は、確立される予測モデルに依存するであろうし、下記の実施例に記載の臨床研究などの臨床研究から実験的に誘導される。使用される予測モデルに依存して、発現閾値は、最大感受性を達成するのに設定されることができ、または最大特異性のためまたは最小誤り(minimum error)(最大識別率(maximum classification rate))のために設定されうる。例えば、より高い閾値は、最小誤りを達成するために設定されうるが、これはより低い感受性をもたらすことがある。故に、任意の所与の予測モデルでは、臨床研究は、最小誤り率を有しながら最も高い感受性を一般に達成する発現閾値を設定するのに使用されるであろう。任意の状況についての発現閾値の決定は、十分当業者の知見の範囲内にある。

【0033】

用語「長期生存」は、本明細書では、手術または他の処置の後、少なくとも5年間、更に好ましくは少なくとも8年間、最も好ましくは少なくとも10年間の生存を指すのに使用される。

【0034】

用語「マイクロアレイ」は、基材上の捕捉剤、好ましくはポリヌクレオチド(例えば、プローブ)またはポリペプチドの規則的配列または不規則的配列を指す。例えば、Microarray Analysis, M. Schena, John Wiley & Sons, 2002; Microarray Biochip technology, M. Schena, ed., Eaton Publishing, 2000; Guide to Analysis of DNA Microarray Data, S. Knudsen, John Wiley & Sons, 2004; およびProtein Microarray Technology, D. Kambhampati, ed., John Wiley & Sons, 2004参照。

【0035】

用語「オリゴヌクレオチド」は、限定することなく、一本鎖デオキシリボヌクレオチド、一本鎖もしくは二本鎖リボヌクレオチド、RNA:DNAハイブリッドおよび二本鎖DNAsを含むポリヌクレオチド、典型的にはプローブまたはプライマーを指す。オリゴヌクレオチド、例えば一本鎖DNAプローブオリゴヌクレオチドは、しばしば、例えば商業的に入手可能な自動化オリゴヌクレオチド合成機を使用して、または*in vitro*発現系、リコンビナント技術および細胞および生物における発現を含む種々の他の方法により合成される。

【0036】

用語「ポリヌクレオチド」は、単数または複数で使用されるとき、一般に、任意のポリリボヌクレオチドまたはポリデオキシリボヌクレオチドであって、改変されていないRNAもしくはDNAまたは改変されたRNAもしくはDNAであることができる、任意のポリリボヌクレオチドまたはポリデオキシリボヌクレオチドを指す。これは、限定することなく、一本鎖DNAおよび二本鎖DNA、一本鎖領域および二本鎖領域を含むDNA、一本鎖RNAおよび二本鎖RNA、ならびに一本鎖もしくは更に典型的には二本鎖であることができるかまたは一本鎖領域および二本鎖領域を含むDNAおよびRNAを含むハイブリッド分子を含む。RNAもしくはDNAまたはRNAとDNAの両方を含む三本鎖領域も含まれる。特にmRNA、cDNAおよびゲノムDNAおよびその任意のフラグメントが含まれる。この用語は、1つ以上の改変された塩基、例えばトリチウム化塩基または格別な塩基、例えばイノシンを含有するDNAおよびRNAを含む。本発明のポリヌクレオチドは、コード配列もしくは非コード配列またはセンス配列もしくはアンチセンス配列を包含することができる。本明細書で「ポリヌクレオチド」または同様な用語への各言及は、完全長配列およびその任意のフラグメント、誘導體もしくは変異体を含むことは理解されるであろう。

【0037】

本明細書で使用された「ポリペプチド」は、オリゴペプチド、ペプチドもしくはタンパク質配列またはそのフラグメントを指し、そして天然に存在する分子、リコンビナント分子、合成分子もしくは半合成分子を指す。天然に存在するタンパク質分子のアミノ酸配列に言及するために本明細書で「ポリペプチド」が詳述されるとき、「ポリペプチド」または同様な用語は、アミノ酸配列を、完全長分子のための完全な、ネイティブなアミノ酸配

10

20

30

40

50

列に限定することを意味しない。本明細書で「ポリペプチド」または同様な用語への各言及は、完全長配列およびその任意のフラグメント、誘導体もしくは変異体を含むことは理解されるであろう。

【0038】

用語「予後」は、医学的アウトカム、例えば、悪いアウトカム (poor outcome) もしくは良好なアウトカム (good outcome) (例えば、長期間生存の尤度) の予測を指し; ネガティブな予後または悪いアウトカムは、再発、疾患の進行 (例えば、腫瘍成長もしくは転移または薬物耐性) または死亡の予測を含み; ポジティブな予後または良好なアウトカムは、疾患の寛解 (例えば無病状態)、改善 (amelioration) (例えば腫瘍後退) または安定化を含む。

10

【0039】

用語「増殖」は、増加した細胞サイズまたは細胞数をもたらすプロセスを指し、そして腫瘍もしくは細胞成長、新脈管形成 (angiogenesis)、神経支配 (innervations) および転移の1つ以上を含むことができる。

【0040】

用語「qPCR」または「QPCR」は、例えば、PCR Technique: Quantitative PCR, J.W.Larrick, ed., Eaton Publishing, 1997, および A-Z of Quantitative PCR, S. Busti n, ed., IUL Press, 2004 に記載の定量的ポリメラーゼ連鎖反応を指す。

【0041】

用語「腫瘍」は、悪性であろうと良性であろうと、すべての新形成細胞成長および増殖、ならびにすべての前癌および癌性細胞および組織を指す。

20

【0042】

ハイブリダイゼーション反応の「ストリンジェンシー」は、当業者により容易に決定することができるが、そして一般に、プローブ長さ、洗浄温度および塩濃度に依存する経験的計算である。一般に、より長いプローブは適合的なアニーリングのためにより高い温度を必要とするが、より短いプローブはより低い温度を必要とする。ハイブリダイゼーションは、一般に、変性されたDNAが、相補性鎖がそれらの溶融温度以下で環境中に存在するとき、再アニーリングする能力に依存する。プローブとハイブリダイゼーション可能な配列との所望の相同性の程度が高ければ高い程、使用することができる相対的温度はより高い。結果として、より高い相対的な温度は反応条件をよりストリンジェントにする傾向があるが、より低い温度は、反応条件をより少なくストリンジェントにする傾向があるということになる。ハイブリダイゼーション反応のストリンジェンシーについての更なる詳細および説明は、例えば、Ausubel et al. Current Protocols in Molecular Biology, Wiley Interscience Publishers, (1995)に見出される。

30

【0043】

本明細書で使用された「ストリンジェントな条件」または「高いストリンジェンシー条件」は、典型的には、(1) 洗浄のための低いイオン強度および高い温度、例えば、50における0.015M塩化ナトリウム/0.0015Mクエン酸ナトリウム/0.1%ドデシル硫酸ナトリウムを使用し; (2) ハイブリダイゼーション期間中変性剤、例えばホルムアミド、50% (v/v)ホルムアミドを、0.1%ウシ血清アルブミン/0.1% F i c o l l / 0.1%ポリビニルピロリドン/750mM塩化ナトリウム、75mMクエン酸ナトリウムを伴うpH6.5の50mMリン酸ナトリウムバッファーとともに42で使用するか; または(3) 50%ホルムアミド、5XSSC (0.75M NaCl, 0.075Mクエン酸ナトリウム)、50mMリン酸ナトリウム (pH6.8)、0.1%ピロリン酸ナトリウム、5Xデンハート (Dewhardt) の溶液、音波処理されたサケ精子DNA (50 μg/ml)、0.1%SDSおよび10%デキストラン硫酸を42で使用し、更に0.2XSSC (塩化ナトリウム/クエン酸ナトリウム) 中42でおよび50%ホルムアミド中55で洗浄し、続いて55でEDTAを含有する0.1XSSCを含む高いストリンジェンシー洗浄を行う。

40

【0044】

50

「中程度にストリンジェントな条件」は、Sambrook et al., *Molecular Cloning: A Laboratory Manual*, New York: Cold Spring Harbor Press, 1989, により記載されたとおり同定され得、そして上記した洗浄溶液およびハイブリダイゼーション条件よりも少なくストリンジェントな洗浄溶液およびハイブリダイゼーション条件（例えば、温度、イオン強度および% SDS）の使用を含む。中程度にストリンジェントな条件の例は、20%ホルムアミド、5XSSC（150mM NaCl、15mMクエン酸三ナトリウム）、50mMリン酸ナトリウム（pH7.6）、5Xデンハートの溶液、10%デキストラン硫酸および20mg/mlの変性され剪断されたサケ精子DNA（denatured sheared salmon sperm DNA）を含む溶液中37℃で一夜のインキュベーション、続いて約37℃～50℃で1XSSC中でフィルターを洗浄することである。当業者は、プローブ長さおよび同様なものなどのファクターに適応するのに必要に応じて温度、イオン強度等をいかに調節するかを認識するであろう。

10

【0045】

本発明の実施は、特記しない限り、当技術分野のスキル内にある、分子生物学（リコンビナント技術を含む）、微生物学、差細胞生物学および生化学の慣用の技術を使用するであろう。このような技術は、文献、例えば*Molecular Cloning: A Laboratory Manual*, 2nd edition, Sambrook et al., 1989, *Oligonucleotide Synthesis*, MJ Gait, ed., 1984; *Animal Cell Culture*, R.L. Freshney, ed., 1987; *Methods in Enzymology*, Academic Press, Inc.; *Handbook of Experimental Immunology*, 4th edition, D. M. Weir & C.C. Blackwell, eds., Blackwell Science Inc., 1987; *Gene Transfer Vectors for Mammalian Cells*, J. M. Miller & M. P. Calos, eds., 1987; *Current Protocols in Molecular Biology*, F. M. Ausubel et al., eds., 1987; および*PCR: The Polymerase Chain Reaction*, Mullis et al., eds., 1994において完全に説明されている。

20

【0046】

発明の態様の説明

【0047】

本発明は、メラノーマにおける特定の予後マーカーおよびシグネチャーの特定の予後の役割を同定および決定するためのマイクロアレイの使用を開示する。本明細書で示されたマイクロアレイをベースとする研究は、メラノーマを有する患者について良好な予後または悪い予後を予測するのに使用されうるマーカーを確立する。特に、本明細書で示されたマイクロアレイをベースとする研究およびqPCR分析は、特定のディファレンシャルに発現された遺伝子が、特定の予後と関連している予後シグネチャーとして使用されうることを示す。したがって、本発明は、攻撃的な疾患(aggressive disease)を有しているような患者を同定するのに使用されうる。

30

【0048】

本発明は、疾患予後の決定のためのマーカーを提供する。本発明の方法を使用して、マーカーがメラノーマ予後と関連しておりそしてアウトカムを予測するのに使用されうることを見出された。種々の段階のメラノーマを有する患者から採取したサンプルのマイクロアレイ分析は、マーカー発現の特定のパターンが癌の予後と関連しているという驚くべき発見に導いた。したがって、本発明は、良好なアウトカムまたは悪いアウトカムを有するメラノーマにおいてディファレンシャルに発現される、表1に略述された遺伝子のセットを提供する。表1に略述された遺伝子は、メラノーマ予後マーカー(MPM)のセットを与える。

40

【0049】

例えば、あるメラノーマ予後マーカー(MPM)の減少は、特定の予後を示すことができる。反対に、他のMPMの増加は、特定の予後を示す。特定の予後は、消化進行の速度を含むことができる。発現の減少または増加は、例えば、試験サンプル、例えば患者の腫瘍サンプルを参照サンプル、例えば既知の予後と関連したサンプルと比較することにより決定されうる。特に、良好な予後を有する患者（1つまたは複数）からの1つ以上のサンプルを参照サンプルとして使用することができる。

50

【0050】

例えば、予後を得るために、患者のサンプル（例えば、腫瘍サンプル）中の発現レベルを、既知のアウトカムを有する患者からのサンプルと比較することができる。患者のサンプルが、悪いアウトカム（急速な疾患進行）を有するサンプルに匹敵する1つ以上のMPMの増加した発現もしくは減少した発現を示すならば、悪い予後が関係付けられる。患者のサンプルが、良好なアウトカム（遅い疾患進行）を有するサンプルに匹敵する1つ以上のMPMの発現を示すならば、ポジティブな予後または良好な予後が関係付けられる。

【0051】

更なる例として、予後を得るために、患者のサンプル（例えば、腫瘍サンプル）からの2つ以上のMPMを含む予後シグネチャーの発現レベルを、良好な予後または悪い予後を有することが知られている癌のサンプルと比較することができる。患者のサンプルが、良好な予後を有するサンプルとの比較によりMPMの増加した発現または減少した発現を示すならば、および/または悪い予後のサンプルに匹敵する発現を示すならば、ネガティブな予後が関係付けられる。患者のサンプルが、良好な予後のサンプルに匹敵するMPMの発現を示すならば、および/または、悪い予後を有するサンプルよりも低い発現もしくは高い発現を示すならば、ポジティブなまたは良好な予後が関係付けられる。

10

【0052】

1つのアプローチとして、予測モデルを発生させるために、予測方法をマーカーのパネル、例えば表1に略述されたMPMのパネルに適用することができる。これは、2つ以上のMPMを含む予後シグネチャーの発生を伴う。

20

【0053】

したがって、開示された表1のMPMは、癌の予後を決定しそしてその腫瘍に特異的な処置法(treatment regime)または処置様式(treatment modality)を確立するための予後シグネチャーを発生させるための有用なマーカーのセットを提供する。特に、ポジティブな予後は、特定の処置選択を推し進めることを決定するために患者により使用されうる。ネガティブな予後は、処置を終わらせることまたは高度に積極果敢な(aggressive)処置もしくは実験的な処置を推し進めることを決定するために患者により使用されうる。更に、患者は、予後マーカー（例えばMPMs）の発現から予測されるそれらの予後に基づいて処置を選ぶことができる。

【0054】

MPMのレベルは、腫瘍組織、腫瘍に近接した組織、リンパ節サンプル、血液サンプル、血清サンプル、尿サンプルまたは糞便サンプルにおいて、任意の適当な方法を使用して検出することができ、そしてオリゴヌクレオチドプローブ、定量的PCRまたはマーカーに対して生じた抗体を含むことができるが、それらに限定されない。予測シグネチャーの形態にある複数のMPMの存在および発現の量を分析し、そして予後シグネチャーを構築することにより、予後の感受性および正確度は増加するであろうことは認識されるであろう。したがって、本発明に従う多重マーカーは、癌の予後を決定するのに使用されうる。

30

【0055】

本発明は、セットにおけるマーカーのアッセイのための保管されたパラフィン包埋された生検物質の使用を含み、したがって、最も広く入手可能なタイプの生検物質と適合性である。本発明は、例えばコア生検または細針吸引による腫瘍組織採取のいくつかの異なる方法とも適合性である。ある局面では、RNAは、固定され、ワックス包埋された患者の癌組織標本から単離される。単離は、例えばコア生検組織または細針吸引細胞から、当技術分野で知られている任意の技術により行われうる。

40

【0056】

1つの局面では、本発明は、予後、例えば、処置後の癌患者の長期間の生存の尤度を予測する方法であって、患者から得られたサンプル中の他のRNA転写物もしくはそれらの産物の発現レベルまたはRNA転写物もしくはそれらの発現産物の参照セットの発現レベルに対して正規化された、患者から得られたサンプル中の1つ以上の予後マーカーまたはそれらの発現産物の発現レベルを決定することを含む方法に関する。

50

【0057】

特定の局面では、予後マーカーは、表1に列挙された1つ以上のマーカーであり、または表1に列挙されたマーカーに由来する予後シグネチャーの1つ以上として含まれる。

【0058】

更なる局面では、予後マーカーまたはそれらの発現産物の発現レベルは、例えば、表1に列挙されたマーカーについておよび表1に列挙されたマーカーに由来する予後シグネチャーについて決定される。他の局面では、本方法は、例えば表1に列挙されたマーカーについてまたは表1に列挙されたマーカーに由来する予後シグネチャーについて、予後マーカーまたはそれらの発現産物のフルセットの発現レベルの決定を含む。

【0059】

更なる局面では、本発明は、例えば、表1に列挙されたマーカーについて、または表1に列挙されたマーカーに由来する予後シグネチャーについて、2つ以上のマーカーにハイブリダイゼーションするポリヌクレオチドを含むアレイ（例えば、マイクロアレイ）に関する。特定の局面では、アレイは、表1に列挙されたマーカーに由来する予後シグネチャーにハイブリダイゼーションするポリヌクレオチドを含む。他の特定の局面では、例えば、表1に列挙されたマーカーについて、マーカーのフルセットにハイブリダイゼーションするポリヌクレオチドを含む。

10

【0060】

これらのアレイでは、ポリヌクレオチドは、cDNAsまたはオリゴヌクレオチドであることができ、そしてそれらがディスプレイされる固体表面は例えばガラスであることができる。ポリヌクレオチドは、本明細書に開示されたマーカーの1つ以上に、例えば完全長配列、任意のコード配列、その任意のフラグメント、またはその任意の相補体にハイブリダイゼーションすることができる。特定の局面では、1つ以上のMPMの発現レベルの増加または減少は、例えば癌再発による、長期間の生存の減少した尤度を示すが、1つ以上のMPMの発現レベルの増加または減少の欠如は、癌再発のない長期間の生存の増加した尤度を示す。

20

【0061】

【表 1】

表 1：メラノーマ予測マーカー

説明	P 値	普通名	Genbank
チオレドキシンドメイン含有 5	0.049	TXNDC5	NM_030810
ペアード免疫グロブリン様 2 型レセプター α	0.049	PILRA	NM_013439
主要組織適合性複合体, クラス I, E	0.049	HLA-E	NM_005516
kiaa1067; kiaa1067	0.049		XM_036173
イノシントリホスファターゼ (ヌクレオシド三リン酸ピロホスファターゼ)	0.049	ITPA	NM_033453
デスムスリン*	0.0482	DMN	NM_145728
GTP 結合タンパク質 2	0.0429	GTPBP2	NM_019096
乳脂肪球-EGF 因子 8 タンパク質	0.0429	MFGE8	NM_005928
イソクエン酸デヒドロゲナーゼ 1 (NADP+), 可溶性	0.0365	IDH1	NM_005896
ミトコンドリアリポソームタンパク質 S5	0.0365	MRPS5	NM_031902
レクチン, ガラクトシド結合性, 可溶性, 7 (ガレクチン 7)	0.0307	LGALS7	NM_002307
Kv チャネル相互作用タンパク質 2	0.0295	KCNIP2	AF347114
炭水化物(N-アセチルグルコサミン 6-O) スルホトランスフェラーゼ 4	0.0295	CHST4	NM_005769
アンサンブルゲンスキャンプレディクション	0.0295		AL451139.11.67295.95669.1
ヒトホスホチロシン非依存性リガンド	0.023	OSIL; A170; p62B	U46752
B細胞インヒビター β における κ 軽ポリペプチド遺伝子エンハンサーの核内因子	0.023	NFKBIB	NM_002503
ミトコンドリアキャリアーホモログ 2 (C. elegans)	0.023	MTCH2	NM_014342
ADP-リボシル化因子関連タンパク質 1	0.0136	ARFRP1	NM_003224
カバノキ花粉アレルゲン特異的免疫グロブリン γ 鎖**	0.0136	BABI-L	AJ131063
チューブリン α 1b***	0.0136	TUBA1B	NM_006082
部分的 n-myc エキソン 3	0.00371		AJ242956_2
プレキシシン B2	0.000756	PLXNB2	AB002313

* このマーカーは以前に kiaa0353 として知られていた; dmn (XM_031031).

** このマーカーは免疫グロブリン κ 可変 1-5 として知られていた (IGKC; AJ131063).

*** このマーカーは、以前にチューブリン α 6 に類似していることが知られていた; loc143712 (XM_084610).

【 0 0 6 2 】

予後マーカー検出への一般的アプローチ

下記のアプローチは、MPMファミリーメンバーを含む増殖マーカーを検出するのに使用することができる非限定的方法である：MPMについて選択的なオリゴヌクレオチドプローブを使用するマイクロアレイアプローチ；MPM特異的プライマーおよびプローブを使用する腫瘍サンプルに関するリアルタイム qPCR；MPM特異的プライマーおよびプローブを使用するリンパ節、血液、血清、糞便または尿サンプルに関するリアルタイム qPCR；酵素結合免疫学的アッセイ (ELISA)；抗マーカー抗体を使用する免疫組織化学；およびコンピュータを使用するアレイまたは qPCR データの分析。

【 0 0 6 3 】

他の有用な方法は、ノーザンブロットイングおよび in situ ハイブリダイゼーション (Parker and Barnes, Methods in Molecular Biology 106: 247-283(1999))；RNase 保護アッセイ (Hod, BioTechniques 13: 852-854(1992))；逆転写ポリメラーゼ連鎖反応 (RT-PCR; Weis et al., Trends in genetics 8: 263-264(1992))；遺伝子発現のシリアル分析 (SAGE; Velculescu et al., Science 270: 484-487(1995); and Velculescu et al., Cell 88: 243-51(1997)), MassARRAY 技術 (MassARRAY technology) (Sequenom, San Die

go,CA), および大規模平行シグネチャーシーケンシング(massively parallel signature sequencing)(MPSS; Brenner et al., Nature Biotechnology 18: 630-634(2000))による遺伝子発現分析を含む。あるいは、DNA二本鎖、RNA二本鎖およびDNA-RNAハイブリッド二本鎖またはDNA-ポリペプチド二本鎖を含む特定の複合体を認識することができる抗体を使用することができる。

【0064】

一次データを集めることができ、そして、例えば、腫瘍組織および非腫瘍組織におけるマーカー発現レベルの比較により；マーカー発現レベル対再発している腫瘍および非再発腫瘍において決定されたレベルの比較により；マーカー発現レベル対転移を有する腫瘍または転移なしの腫瘍において決定されたレベルの比較により；マーカー発現レベル対異なる段階の腫瘍において決定されたレベルの比較により；またはマーカー発現レベル対異なるレベルの増殖を有する細胞において決定されたレベルの比較により、倍率変化分析(fold change analysis)を行うことができる。ネガティブな予後またはポジティブな予後は、この分析に基づいて決定される。腫瘍マーカー発現の更なる分析は、増加した発現または減少した発現を示すこれらのマーカーを既知のメラノーマ腫瘍の発現プロファイルとマッチさせて予後を与えることを含む。

10

【0065】

発現が増加されることを結論するための閾値は、特定のマーカーおよびアプリケーションされるべき特定の予測モデルにも依存するであろう。閾値は、一般に、最も低い誤り率で最も高い感受性および選択性を達成するように設定されるが、特定の臨床状況については変更が望ましいことがある。所望の閾値は、任意の予測モデルの統計的変動性を考慮して十分なサイズの母集団を分析することにより決定されそして予測モデルを作成するのに使用されるサンプルのサイズから計算される。発現が減少していることを結論するための閾値の決定について同じことがあてはまる。増加した発現または減少した発現が起こったことを結論するための、他の閾値、または閾値を確立するための方法が、本発明の範囲から逸脱することなく選択されうることは認識されうる。

20

【0066】

予測モデルは、そのアウトプットとして数値、例えばスコア、尤度値または確率を作り出すことも可能である。これらの場合に、閾値を予測モデルにより作られた結果にアプリケーションすることが可能であり、そしてこれらの場合に、同様な原理が発現値のための閾値を設定するのに使用される原理として成り立つ。

30

【0067】

一旦、腫瘍サンプルにおける予後シグネチャーの発現レベルまたは予測モデルのアウトプットが得られると、次いで癌再発の尤度を決定することができる。

【0068】

同定されたマーカーから、1つ以上のMPMを含む予後シグネチャーを使用して、1つ以上のマーカーの発現レベルを開示された予後シグネチャーと比較することにより癌の予後を決定することができる。腫瘍サンプル中の1つ以上のMPMの発現を開示された予後シグネチャーと比較することにより、癌再発の尤度を決定することができる。予後を確立するための予後シグネチャーの発現レベルの比較は、先に述べた予測モデルをアプリケーションすることによりなされうる。

40

【0069】

癌再発の尤度を決定することは、医師にとって大きな価値がある。腫瘍が処置に応答しない高い尤度は、より長いもしくはより高い用量処置が考慮されるべきであるか、または、処置が全く与えられていないかもしれないことを意味する。正確な予後は患者にとっても利益である。それは、患者が彼らの配偶者、家族、および友人とともに処置に関する決定を下すことも可能とし、ならびに患者の将来および生活様式の変化に関する決定も可能とする。したがって、本発明は、腫瘍サンプルにおけるマーカーの発現をディファレンシャルな発現シグネチャーとマッチさせることにより確立された予後に基づく特定の癌に対する処置法を確立する方法も提供する。

50

【0070】

マーカー選択または予後シグネチャーの構築は、本明細書の表1に開示されたMPMに制限される必要はないが、開示されたシグネチャーからの1つ以上のMPMの使用を含むことができ、または開示されたマーカーリストから選ばれたMPMを使用して新規なシグネチャーが確立されうるとは認識されるであろう。任意のシグネチャーの要件は、それが、医師が処置法を確立するのを助けるのに十分な正確度で、急速な疾患進行の尤度を予測するという点である。

【0071】

逆転写PCR(RT-PCR)

上記した技術のなかで、最も感受性のそして最も柔軟性の定量的方法は、RT-PCRであり、これは、薬物処理してもしくは薬物処理しないで、正常な組織および腫瘍組織における異なるサンプル母集団においてRNAレベルを比較し、発現のパターンを特徴付け、密接に関連したRNAs間で識別しそしてRNA構造を分析するのに使用されう。

10

【0072】

RT-PCRでは、最初の工程は、ターゲットサンプルからのRNAの単離である。出発物質は、典型的には、ヒト腫瘍または腫瘍細胞系およびそれぞれ、対応する正常な組織または細胞系から単離されたトータルRNAである。RNAは、種々のサンプルから、例えば、胸、肺、結腸(例えば大腸もしくは小腸)、皮膚、結腸直腸、胃、食道、肛門、直腸、前立腺、脳、肝臓、腎臓、膵臓、脾臓、胸腺、睾丸、卵巣、子宮等、組織、からの腫瘍サンプルから単離されることができ、原発性腫瘍もしくは腫瘍細胞系から、および健康なドナーからのプールされたサンプルから単離されう。RNAのソースが腫瘍であるならば、RNAは、例えば、凍結されたもしくは保管された(archived)パラフィン包埋された組織サンプル、および固定された(例えばホルマリン固定された)組織サンプルから抽出されう。

20

【0073】

RT-PCRによる遺伝子発現プロファイリングの第1工程は、RNAテンプレートのcDNAへの逆転写、次いでPCR反応でのその幾何級数的な増幅である。2つの最もよく使用される逆転写酵素は、ニワトリ骨髄芽球症ウイルス逆転写酵素(avian myeloblastosis virus reverse transcriptase)(AMV-RT)およびモロニー Maus 白血病ウイルス逆転写酵素(Moloney murine leukemia virus reverse transcriptase)(MMLV-RT)である。逆転写工程は、典型的には、発現プロファイリングの状況および目標に依存して、特異的プライマー、ランダムヘキサマーまたはオリゴdTプライマーを使用してプライミングされる。例えば、抽出されたRNAを、Gene Amp RNA PCRキット(Perkin Elmer, CA, USA)を使用して、製造者の指示に従って逆転写させることができる。次いで誘導されたcDNAを、次のPCR反応におけるテンプレートとして使用することができる。

30

【0074】

PCR工程は、種々の熱安定性DNA依存性DNAポリメラーゼを使用することができるけれども、典型的には、5'-3'ヌクレアーゼ活性を有するが、3'-5'ブルーフリーディングエンドヌクレアーゼ活性を欠くTaq DNAポリメラーゼが使用される。かくして、TaqMan(q)PCRは、典型的にはTaqまたはTthポリメラーゼの5'ヌクレアーゼを利用して、そのターゲットアンプリコンに結合したハイブリダイゼーションプローブを加水分解するが、同等な5'ヌクレアーゼ活性を有する任意の酵素を使用することができる。

40

【0075】

2つのオリゴヌクレオチドプライマーを使用してPCR反応の典型的なアンプリコンを発生させる。第3オリゴヌクレオチドまたはプローブが、2つのPCRプライマー間に位置したヌクレオチド配列を検出するようにデザインされる。このプローブは、Taq DNAポリメラーゼ酵素によりエクステンション可能ではなく(non-extendible)、そしてレポーター蛍光染料およびクエンチャー蛍光染料で標識される。レポーター染料からの任意のレーザー誘導発光は、2つの染料がプローブ上にあるままで2つの染料が相互に近接して

50

位置するとき、クエンチング染料によりクエンチングされる。増幅反応期間中、Taq DNAポリメラーゼ酵素は、テンプレート依存性方式でプローブを開裂する。得られるプローブフラグメントは溶液中で解離し、そして放出されたレポーター染料からのシグナルは、第2フルオロフォアのクエンチング効果から免れる。レポーター染料の1つの分子が、合成された各新しい分子に対して遊離され、そしてクエンチングされなかったレポーター染料の検出は、データの定量的説明のための基礎を与える。

【0076】

TaqMan RT-PCRは、例えば、ABI PRISM 7700 Sequence Detection System(Perkin-Elmer-Applied Biosystems, Foster City, CA, USA)またはLightcycler(Roche Molecular Biotechnicals, Mannheim, Germany)などの市販の装置を使用することができる。好ましい態様では、5'ヌクレアーゼ処理は、ABI PRISM 7700 Sequence Detection Systemなどのリアルタイム定量的PCR装置で行われる。このシステムは、サーモサイクラー、レーザー、電荷結合デバイス(CCD)、カメラおよびコンピュータからなる。このシステムは、サーモサイクラー上で96ウエルフォーマットにおいてサンプルを増幅する。増幅期間中、レーザーで誘導された蛍光シグナルを、すべての96ウエルについて光学繊維ケーブルを通してリアルタイムで収集し、そしてCCDで検出する。システムは装置を動かすためおよびデータを分析するためのソフトウェアを含む。

10

【0077】

5'ヌクレアーゼアッセイデータは、最初にCtまたは閾値サイクルとして表現される。上記したとおり、蛍光値は、すべてのサイクル期間中記録されそして増幅反応においてその点に増幅された産物の量を表す。蛍光シグナルが統計的に有意であるとして最初に記録されるときは閾値サイクルである。

20

【0078】

誤りおよびサンプル間の変動の効果を最小にするために、RT-PCRは通常内部標準を使用して行われる。理想的な内部標準は異なる組織間で一定のレベルで発現されそして実験処置により影響されない。遺伝子発現のパターンを正規化するために最も頻繁に使用されるRNAsは、ハウスキーピング遺伝子グリセルアルデヒド-3-リン酸-デヒドロゲナーゼ(GAPDH)および-アクチンのためのmRNAsである。

【0079】

リアルタイム定量的PCR(qPCR)

30

RT-PCR技術のより最近の変化は、リアルタイム定量的PCRであり、これは、二重標識された蛍光発生プローブ(fluorigenic probe)(即ち、TaqManプローブ)によるPCR産物蓄積を測定する。リアルタイムPCRは定量的競合PCRおよび定量的比較PCRの両方と適合性である。前者は正規化のために各ターゲット配列のための内部競合体を使用し、後者は、サンプル内に含有される正規化遺伝子またはRT-PCRのためのハウスキーピング遺伝子を使用する。更なる詳細は、例えば、Held et al., Genome Research 6: 986-994(1996)により与えられる。

【0080】

発現レベルは、RNAソースとして固定され、パラフィン包埋された組織を使用して決定することができる。本発明の1つの局面に従えば、PCRプライマーおよびプローブは、増幅されるべき遺伝子に存在するイントロン配列に基づいてデザインされる。この態様においては、プライマー/プローブデザインにおける最初の工程は、遺伝子内のイントロン配列の脱線状化である。これは、Kent, W. J., Genome Res. 12(4): 656-64(2002)により開発されたDNA BLASTソフトウェアなどの公に入手可能なソフトウェアにより、またはそのバリエーションを含むBLASTソフトウェアにより行うことができる。次の工程は、PCRプライマーおよびプローブデザインの十分に確立された方法に従う。

40

【0081】

非特異的シグナルを回避するために、プライマーとプローブをデザインするとき、イントロン内の反復配列をマスクすることは有用である。これは、反復エレメントのライブラリーに対してDNA配列をスクリーニングしそして反復エレメントがマスクされている問

50

い合わせ配列(query sequence)に戻す、Bayer College of Medicineを通じてオンラインで入手可能なRepeat maskerプログラムを使用することにより容易に達成されうる。次いでマスクされた配列を使用して任意の市販のまたはそうでなければ公に入手可能なプライマー/プローブデザインパッケージ、例えば、Primer Express(Applied Biosystems); MGB assay-by-design(Applied Biosystems); Primer3(Steve Rozen and Helen J. Skaletsky(2000) Primer3 on the WWW for general users and for biologist programmers in: Krawetz S, Misener S(eds) Bioinformatics Methods and Protocols: Methods in Molecular Biology. Humana Press, Totowa, NJ, pp 365-386)、を使用してプライマーおよびプローブ配列をデザインするのに使用することができる。

【0082】

PCRプライマーデザインにおいて考慮される最も重要な因子が、プライマー長、溶融温度(T_m)およびG/C含有率、特異性、相補性プライマー配列および3'末端配列を含む。一般に、最適なPCRプライマーは、一般的に長さは17~30塩基であり、そして約20~80%のG+C塩基、例えば約50~60%のG+C塩基を含有する。50~80、例えば約50~70の溶融温度が典型的に好ましい。PCRプライマーおよびプローブデザインについての更なるガイドラインは、例えば、Diefenbach, C. W. et al., General Concepts for PCR Primer Design in: PCR Primer, A Laboratory Manual, Cold Spring Harbor Laboratory Press, New York, 1995, pp. 133-155; Innis and Gelfand, Optimization of PCTs in: PCR Products, A Guide to Methods and Applications, CRC Press, London, 1994, pp. 5-11; and Plasterer, T. N. primerselect: Primer and probe design. Methods Mol. Biol. 70: 520-527(1997)を見られたい。その全体の開示は参照により明白に組み込まれる。

【0083】

マイクロアレイ分析

ディファレンシャルな発現は、マイクロアレイ技術によって同定または確認することもできる。したがって、MPMの発現プロファイルは、マイクロアレイ技術を使用して新しい腫瘍組織またはパラフィン包埋された腫瘍組織のいずれかにおいて測定することができる。この方法では、関心のあるポリヌクレオチド配列(cDNAsおよびオリゴヌクレオチドを含む)をマイクロチップ基材上にプレートまたはアレイする。次いでアレイされた配列(即ち捕捉プローブ)を関心のある細胞または組織からの特異的ポリヌクレオチド(即ち、ターゲット)とハイブリダイゼーションさせる。丁度RT-PCR法におけると同じく、RNAのソースは、ヒト腫瘍もしくは腫瘍細胞系および対応する正常な組織もしくは細胞系から単離されたトータルRNAである。したがって、RNAは、種々の原発性腫瘍または腫瘍細胞系から単離することができる。RNAのソースが原発性腫瘍であるならば、RNAは、例えば、毎日の臨床実施においてルーチンに調製されそして保存されている、凍結されまたは保管された、ホルマリン固定され、パラフィン包埋された(FFPE)組織サンプルおよび固定された(例えば、ホルマリン固定された)組織サンプルから抽出されうる。

【0084】

マイクロアレイ技術の特定の態様では、cDNAクローンのPCR増幅されたインサートを基剤に適用する、基材は、1、2、5、10、15、20、25、30、35、40、45、50または75までのヌクレオチド配列を含むことができる。他の局面では、基材は、少なくとも10,000のヌクレオチド配列を含むことができる。マイクロチップ上に固定化されたマイクロアレイされた配列は、ストリンジентな条件下のハイブリダイゼーションのために適当である。他の態様として、マイクロアレイのためのターゲットは、長さが少なくとも50、100、200、400、500、1000または2000塩基、または長さが50~100、100~200、100~500、100~1000、100~2000もしくは500~5000塩基であることができる。更なる態様として、マイクロアレイのための捕捉プローブは、長さが少なくとも10、15、20、25、50、75、80または100塩基、または長さが10~15、10~20、10~2

10

20

30

40

50

5、10～50、10～75、10～80、または20～80塩基であることができる。

【0085】

蛍光的に標識されたcDNAプローブは、関心のある組織から抽出されたRNAの逆転写により蛍光性ヌクレオチドの組み込みにより発生させることができる。チップに適用された標識されたcDNAプローブは、アレイ上のDNAの各スポットに特異性をもってハイブリダイゼーションする。非特異的に結合したプローブを除去するためのストリンジেন্টな洗浄後に、共焦点レーザー顕微鏡によりまたはCCDカメラなどの他の検出方法によりスキャンされる。各アレイされたエレメントのハイブリダイゼーションの定量は、対応するmRNA量(abundance)の評価を可能とする。二色蛍光により、2つのRNAのソースから発生させた別々に標識されたcDNAプローブをペアー方式でアレイにハイブリダイゼーションさせる。かくして、各特定された遺伝子に対応する2つのソースからの転写物の相対的量を同時に決定する。

10

【0086】

ハイブリダイゼーションのミニチュア化されたスケールは、多数の遺伝子の発現パターンの便利で迅速な評価を与える。このような方法は、細胞当たり少数のコピーで発現される希少転写物を検出するのに必要な感受性および発現レベルにおいて少なくとも約2倍の差を再現可能に検出するのに必要な感受性を有することが示された(Schena et al., Proc. Natl. Acad. Sci. USA 93(2): 106-149(1996))。マイクロアレイ分析は、市販の装置、例えば、Affimetrix GenChip技術、イルミナマイクロアレイ技術またはIncyteのマイクロアレイ技術により、製造者のプロトコルに従って、行うことができる。遺伝子発現の大規模分析のためのマイクロアレイ法の開発は、種々の腫瘍タイプにおける癌分類およびアウトカム予測の分子マーカーを系統的に研究することを可能とする。

20

【0087】

RNA単離、精製および増幅

mRNA抽出のための一般的な方法は、当技術分野で周知でありそしてAusubel et al., Current Protocols of Molecular Biology, John Wiley and Sons(1997)を含む分子生物学の標準方法に開示されている。パラフィン包埋された組織からのRNA抽出のための方法は、例えば、Rupp and Locker, Lab Invest. 56: A67(1987)およびDe Sanders et al., BioTechniques 18: 42044(1995)に開示されている。特に、RNA単離は、Quiagenなどの商業的製造者からの精製キット、バッファセットおよびプロテアーゼを使用して、製造者のインストラクションに従って行うことができる。例えば、培養物中の細胞からのトータルRNAは、Quiagen RNeasyミニカラムを使用して単離されうる。他の市販のRNA単離キットは、MasterPure Complete DNA and RNA Purification Kit(EPICENTRE(D, Madison, WI)およびParaffin Block RNA Isolation Kit(Ambion, Inc)を含む。組織サンプルからのトータルRNAは、RNA Stat-60(Tel-Test)を使用して単離されうる。腫瘍から調製されたRNAは、例えば、塩化セシウム密度勾配遠心により単離されうる。

30

【0088】

mRNA単離、精製、プライマーエクステンションおよび増幅を含む、RNAソースとして固定された、パラフィン包埋された組織を使用する遺伝子発現をプロファイリングするための代表的プロトコルの工程は、種々の公表された雑誌文献(例えば、T. E. Godfrey et al. J. Molec. Diagnostics 2: 84-91(2000); K. Specht et al., Am. J. Pathol. 158: 419-29(2001))に記載されている。簡単にいえば、代表的な方法は、パラフィン包埋された腫瘍組織サンプルの10µm厚さの切片を切断することで出発する。次いでRNAを抽出し、そしてタンパク質およびDNAを除去する。RNA濃度の分析の後に、RNA修復および/または増幅工程を必要に応じて含ませることができ、そしてRNAを遺伝子特異的プロモーターを使用して逆転写し、次いでRT-PCRを行う。最後に、データを分析して、検査された腫瘍サンプルにおいて同定された特徴的遺伝子発現パターンに基づいて患者に利用可能な最善の処置選択(1つまたは複数)を同定する。

40

【0089】

免疫組織化学およびプロテオミクス

50

免疫組織化学法は、本発明の増殖マーカーの発現レベルを検出するのにも適当である。したがって、各マーカーに対して特異的な抗体もしくは抗血清、好ましくはポリクローナル抗血清、最も好ましくはモノクローナル抗体が発現を検出するのに使用される。抗体は、例えば、放射性標識、蛍光標識、ハプテン標識、例えばビオチンまたは酵素、例えばセイヨウワサビペルオキシダーゼもしくはアルカリホスファターゼによる、抗体自体の直接標識化により検出されうる。あるいは、標識されていない第一抗体を、第1抗体に対して特異的な、抗血清、ポリクローナル抗血清またはモノクローナル抗体を含む標識された第二抗体と共に使用する。免疫組織化学プロトコールおよびキットは、当技術分野で周知でありそして市販されている。

【0090】

プロテオミクスを使用して、ある時点でのサンプル（例えば、組織、器官、または細胞培養物）中に存在するポリペプチドを分析することができる。特に、プロテオーム技術を使用して、サンプル中のポリペプチド発現の全体的変化を評価することができる（発現プロテオミクスとも呼ばれる）。プロテオーム分析は、典型的には、（1）二次元ゲル電気泳動(2-D PAGE)によるサンプル中の個々のポリペプチドの分離；（2）例えば、質量分析法またはN末端配列決定による、ゲルから回収された個々のポリペプチドの同定および（3）バイオインフォマティクスを使用するデータの分析を含む。プロテオミクス法は、遺伝子発現プロファイリングの他の方法に対する価値ある補足であり、そして単独でまたは他の方法と組み合わせて、本発明の増殖マーカーの産物を検出するのに使用することができる。

【0091】

一旦腫瘍サンプル中の1つ以上の予後マーカーの発現レベルが評価されると、処置に回答する癌の尤度が決定されうる。本発明者等は、患者データセットにおいて、処置に回答しない（悪い予後）メラノーマに比べて処置に回答する（良好な予後）メラノーマにおいてディファレンシャルに発現される多数のマーカーを同定した。マーカーは表1に示されそして下記の実施例に示される。

【0092】

ディファレンシャルに発現された遺伝子の選択

有意と思われる遺伝子の選択への早期のアプローチは、関心のある2つのグループ間での所与の遺伝子の「倍率変化」(fold change)を単に調べることを含んでいた。このアプローチは、最も劇的に変化するとと思われる遺伝子に関して効果があるが、基本的統計学の考察は、分散（またはノイズレベル）が極めて高い（しばしばマイクロアレイ実験で見られるとおり）ならば、みかけは大きい倍率変化がただ偶然によるだけでしばしば起こりうることを認識させる。

【0093】

マイクロアレイ実験、例えば本明細書で説明されたそれらは、典型的には、数千の遺伝子の同時測定を伴う。2つのグループ（例えば良好な予後の腫瘍と悪い予後マーカーの腫瘍）間での特定の遺伝子の発現レベルを比較している場合に、有意についての典型的な検定（t検定などの）は十分ではない。何故ならば、これは、数千の実験（この状況では、各遺伝子は「実験」を構成する）の群において、少なくとも1つの実験が偶然によるだけで通常の判断基準を通過する確率は本質的に一つ(unity)であるからである。有意についての検定において、典型的には、「帰無仮説」が正しいという確率を計算する。2つのグループを比較する場合に、帰無仮説は、2つのグループ間に差はないということである。統計的な検定がある閾値（通常0.05または0.01）以下で帰無仮説について確率を生じるならば、我々は帰無仮説を拒絶することができそして2つのグループは有意に異なるという仮説を受け入れることができるといわれる。明らかに、このような検定では、偶然のみによる帰無仮説の拒絶が20回に1回（または100回に1回）期待されうる。有意についてのt検定のまたは他の同様な統計的検定の使用は、マイクロアレイに関しては、あまりにも多くの偽陽性（または第1種の過誤）を生じることになる。

【0094】

10

20

30

40

50

このタイプの状況においては、同時に多数の仮説を検定する場合には、Bonferroni Method¹²などの典型的な多重比較法が適用される。しかしながら、このような検定は、大抵のマイクロアレイ実験のためにはあまりにも保守的であり、あまりにも多くの偽陰性（第二種の）誤りをもたらす。

【0095】

より最近のアプローチは、所与の検定が有意であるための確率を適用する試みを捨て、そして予期される割合の第一種の過誤（または偽発見率(false discovery rate)¹³）がコントロールされるような実験のサブセットを選ぶための手段を確立することである。種々の実現(implementations)により；即ち、BRB Array Tools¹⁴およびBioconductorのlimma^{15, 16} package（統計的環境 R (R statistical environment)を使用する；^{17, 18}）を備えた方法によりこの研究で使用されたのはこのアプローチである。

10

【0096】

データマイニングのための一般的方法：予後シグネチャーの発生

データマイニングは、「知識」の抽出、換言すれば（通常）大容量のデータ（データセット）からの「ノーハウ」または予測能力を説明するために使用される用語である。これは、予後シグネチャーを発生させるのにこの研究で使用されたアプローチである。この研究の場合に、「ノーハウ」は、遺伝子発現測定またはシグネチャー（この節で一般的に述べられそして実施例の節で更に詳細に述べられた）の所与のセットから予後を正確に予測する能力である。

【0097】

この研究で使用される方法のために使用される特定の詳細は、実施例17 - 20に記載されている。しかしながら、データマイニング法（実施例に記載の方法およびここで記載の方法の両方）のいずれの適用もこの一般的プロトコルに従うことができる。

20

【0098】

データマイニング¹⁹および関連した話題の機械学習²⁰は、1つ以上の適切なコンピュータソフトウェアパッケージ（下記参照）の使用を含む複雑な反復的な数学的タスクである。ソフトウェアの使用は、一方では、正しい方法に忠実であるとの条件下に、データマイニング技術を成功的に使用するために各技術の背景の理論の複雑さに完全に習熟する必要はないという点で有利である。欠点は、データマイニングの適用がしばしば「ブラックボックス」：人がデータを挿入しそして答えを受け取る、と考えられうるということである。これがいかにして達成されるかは、しばしばエンドユーザーからマスクされており（これは説明した技術の多くについてそうである）そしてしばしばデータマイニングのために選ばれる統計的方法に影響を与えることがありうる。例えば、ニューラルネットワークおよびサポートベクターマシンは、エンドユーザーが決定をするのに使用される「ルール」を引き出すことを非常に困難にする特に複雑な実装を有する。他方、k最近傍(k-nearest neighbours)および線形判別分析(linear discriminant analysis)は、ユーザーから隠されていない、決定をするための非常に透明なプロセスを有する。

30

【0099】

データマイニングで使用される2つのタイプのアプローチがある：教師ありアプローチ(supervised approach)および教師なしアプローチ(unsupervised approach)。教師ありアプローチでは、カテゴリーデータ（例えば良好な予後対悪い予後）などのデータに連結されている情報は知られている。必要とされるのは、インプット変数に観測された応答（例えば、良好な予後対悪い予後）を連結する能力である。教師なしアプローチでは、データセット内のクラスは、前以て知られておらず、そしてデータマイニング技術は、データセット内のクラスまたは構造を発見することを試みるために使用される。

40

【0100】

本実施例では、教師ありアプローチが使用され、そしてここで詳細に論議されるが、他の技術のいずれかを使用することも認識されるであろう。

【0101】

全体のプロトコルは、下記の工程を含む。

50

データ表示

これは、選ばれたデータマイニング技術で最も成功的に働きそうな形態にデータを変換することを含む。調査されるべきデータが遺伝子発現の相対的レベルを表すこの研究におけるように、データが数値である場合には、これは非常に簡単である。データが大きなダイナミックレンジを含むならば（即ち、多くの桁の大きさ）、しばしばデータの対数が採用される。別々の研究者による、別々の日の別々のサンプルの多くの測定値を含むならば、系統誤差が最小になることを確実にするよう、特に注意しなければならない。系統誤差（即ち、プロトコール差、機械差、オペレーター差および他の定量可能な因子から生じる誤差）の最小化は、ここでは「正規化」と呼ばれるプロセスである。

【0102】

特徴選択

典型的には、データセットは、日々の基準で測定するのに実際的であるよりも多くのデータエレメントを含有し、そして更に予測モデルを作成するのに必要な情報を与えない多くのエレメントを含有する。予測モデルのデータセットを説明する実際の能力は、データセットの全次元(full dimensionality)のいくつかのサブセットから誘導される。これらの次元はデータセットの最も重要な成分（または特徴）である。マイクロアレイデータに関しては、データセットの次元は個々の遺伝子であることに留意されたい。ここで述べた文脈において、特徴選択は、最も「ディファレンシャルに発現される」これらの遺伝子を発見することを含む。より一般的な意味においては、それは、有意についてのいくつかの統計的検定に通るこれらのグループを含む、即ち、調査されているグループの一方もしくは他方において一貫してより高いかまたはより低い特定の変数のレベルである。時には、特徴は、最も大きい分散を示すこれらの変数（または次元）である。

【0103】

特徴選択の適用は、予測モデルを作り出すのに使用される方法とは完全に独立であり、そして所望の結果を達成するための多量の実験を含む。本発明内では、有意な遺伝子の選択は、特徴選択を伴う。更に、データ縮小(data reduction)の方法（主成分分析などの）をデータセットにアプリケーションすることができる。

【0104】

トレーニング

一旦クラス例えば、良好な予後 / 悪い予後）および選択の特徴が確立されそしてデータがデータマイニングのためのインプットとして受け入れ可能な形態で表示されると、縮小したデータセット（特徴で述べられたとおりの）を選択された予測モデルにアプリケーションする。このモデルのためのインプットは、関連したアウトプット（クラス標識または応答）とともに、通常多次元数値インプット（ベクトルとして知られた）の形態にある。トレーニングプロセスにおいては、選択されたデータは、順次に（ニューラルネットワークなどの技術において）またはまとめて（線形モデル、線形判別分析、サポートベクターマシンなどのいくつかの回帰を適用する技術において）予測モデルにインプットされる。ある場合には（例えばk最近傍）、データセット（または特徴選択の後に得られたデータセットのサブセット）は、それ自体モデルである。論議されたとおり、有効なモデルは、成功的結果を最ももたらしそうであるとエキスパートアナリストによりモデルのパラメータが予め決定されている種々のソフトウェアパッケージの使用により、詳細な数学の最小の理解で確立されうる。

【0105】

バリデーション

これは、データマイニングプロトコールのキー成分であり、そしてこれの誤ったアプリケーションは、しばしば誤りに導く。データセットの一部は、予測モデルの成功を検定するために、特徴選択およびトレーニングとは別に、取っておくべきである。更に、バリデーションの結果がモデルの特徴選択およびトレーニングを行うのに使用されるならば、モデルが実生活の状況にアプリケーションされる前に、モデルを検定するための更なるバリデーションセットが得られる。このプロセスが厳密にモデルに忠実でない(not adhered)

10

20

30

40

50

ならば、モデルは、実世界の状況において失敗する可能性が高い。バリデーションの方法は以下に更に詳細に述べられる。

【0106】

アプリケーション

モデルが構築されそしてバリデーションされると、それはエンドユーザーに受け入れられるような方法でパッケージされなければならない。これは、しばしば、モデルが埋め込まれているスプレッドシートアプリケーションのある形態の実装、統計的ソフトウェアパッケージのスク립ティング、または情報技術スタッフによるハードコードドアプリケーション(hard-coded application)へのモデルのリファクタリングを含む。

【0107】

しばしば使用されるソフトウェアパッケージの例は、多数のベンダーから得られるスプレッドシートプラグイン(spreadsheet plugins)、統計的環境 R (R statistical environment)、市販のパッケージ Matlab、S-Plus、SAS、SPSS、STATA、フリーオープンソースソフトウェア(free open-source software)、例えば Octave (MatLab クローン)

市販のクローズドソースセッティング(closed-source setting)における予測モデルを実装するのに使用されうる多くのおよび多様な C++ ライブラリー。

【0108】

データマイニング法の例

本発明の方法は、最初にデータマイニング(上記)の工程を引き受け、次いで適切な既知のソフトウェアパッケージをアプリケーションすることにより行うことができる。データマイニングのプロセスの更なる説明は多くの極めてよく書かれているテキストに詳細に記載されている¹⁹。

【0109】

線形モデル^{19, 21}

データは、線形回帰モデルのインプットとして処理され、そのクラス標識または応答変数はアウトプットである。クラス標識または他のカテゴリーデータは、数値(通常整数)に変換されなければならない。一般化された線形モデルにおいては、クラス標識または応答変数は、それ自体はインプットデータに線形で関係していないが、「連結関数」の使用により変換される。ロジスティック回帰は、一般化された線形モデルの最も普通の形態である。

【0110】

線形判別分析(linear discriminant analysis)^{19, 22, 23}

データが線形に分離可能である(即ち、データのグループまたはクラスが、閾値の n 次元拡張である超平面により分離されうる)との条件下で、この技術は適用されうる。変数の組み合わせは、グループ間分散が最大化されそしてグループ内分散が最小になるように、クラスを分離するのに使用される。この副産物は、分類ルールの形成である。未知のクラスのサンプルへのこのルールの適用は、クラス所属関係の予測または分類がそのサンプルについてなされることを可能とする。マイクロアレイ分析のために普通に使用されるニアレストシュランケンセントロイド(nearest shrunken centroids)などの線形判別分析の変形もある。

【0111】

サポートベクターマシン²⁴

変数のコレクションを重量のコレクションとともに使用してこれらの重り付き変数によってクラス間の分離を最大化するモデルを決定する。次いでこのモデルのサンプルへの適用は、そのサンプルの分類またはクラス所属関係の予測を生じさせる。

【0112】

ニューラルネットワーク²³

データは、ノードのネットワークへのインプットとして処理され、ノードのネットワー

10

20

30

40

50

クは、表面的には生物学的ニューロンに似ており、インプットをすべてのノードからそれらが連結されているノードにアプリケーションし、そしてインプットをアウトプットに変換する。普通は、ニューラルネットワークは、「乗算および和算アルゴリズム」(multiply and sum algorithm)を使用して、多数の接続されたインプットノードからのインプットをシングルアウトプットに変換する。ノードは、そのノードへのインプットがある閾値を超えないならば、必ずしもアウトプットを生成しなくてもよい。各ノードは、そのインプットとしていくつかの他のノードからのアウトプットを有し、最終アウトプットノードは通常カテゴリー変数に連結されている。ノードの数およびノードのトポロジーは、殆ど無限に変わることができ、他の方法では分類することができないことがある非常にノイズのあるデータを分類する能力を与える。ニューラルネットワークの最も普通の実装は、多層パーセプトロンである。

10

【 0 1 1 3 】

分類および回帰木^{2 5}

これらにおいては、変数は、段階的方式でたどられる(followed)ことができるルールの階層制(hierarchy of rules)を規定して、サンプルのクラスを決定するのに使用される。典型的なプロセスは、特定のクラスアウトプットまたは識別できないことの特定の説明をもたらすルールのセットを創り出す。例としての分類木は、下記のようなアルゴリズムの実装である。

```

if gene A> x and gene Y> x and gene Z=z
then
    class A
else if gene A=q
    then
    class B

```

20

【 0 1 1 4 】

最近傍法(nearest neighbor methods)^{2 2 , 2 3}

予測または分類は、サンプル(未知のクラスの)をその周囲のサンプル(既知のクラスの)と比較することによりなされ、近さ(closeness)は距離関数により定義される。多くの異なる距離関数を定義することが可能である。普通で使用される距離関数は、ユークリッド距離(三角測量におけるようにピタゴラスの距離のn次元への拡張、種々の形態の相関(ピアソン相関係数を含む)である。普通は意味のある距離メトリック(meaningful distance metric)により相互連結されないデータ点を、ユークリッド距離が適用されうるように、ユークリッド空間に変換する変換関数もある(例えばマハラノビス距離)。距離メトリックは極めて複雑でありうるけれども、k最近傍の基本的前提は極めて単純であり、本質的に未知のインプットに最も類似したkデータベクトルを「発見することの」言い換えであり、それらがどのクラスに対応するかを発見し、そして未知のインプットがどのクラスであるかに関して票決する(vote)。

30

【 0 1 1 5 】

他の方法

ベイジアンネットワーク

40

有向非環式グラフ(directed cyclic graph)を使用して変数のコレクションをそれらの同時確率分布(joint probability distribution)と共に表示し、次いでこれを使用してサンプルのクラス所属関係の確率を決定する。

独立したシグナル(例えばクラス所属関係)が変数のコレクションから分離される(成分に)独立成分分析。次いでこれらの成分を使用して、サンプルのクラス所属関係の分類または予測を作成することができる。

予測方法のコレクションを組み合わせてサンプルのクラス所属関係の同時分類または予測(joint classification or prediction)を作成するアンサンブル学習法。

【 0 1 1 6 】

探索されうるこれらの方法の多くの変法があり^{1 9}、そして多くの新しい方法が絶えず

50

定義されそして開発されている。許容しうる結果を得るために、これらの方法のいずれかを適用することができることは当然である。過学習(overfitting)を回避するために、包括的バリデーションスキーム(comprehensive validation scheme)によりすべての結果を検定することを確実にすることによって、特別な注意を払わなければならない。

【0117】

バリデーション

説明した予測方法のいずれかのアプリケーションは、この方法を新しいデータセット(例えば臨床試験からのデータ)にアプリケーションすることができる前に、トレーニングとクロスバリデーション^{1 2, 2 6}の両方を伴う。トレーニングは、関心のあるデータセット(この場合にメラノーマからの遺伝子発現測定値)のサブセットを、それが検定されるべきクラス(この場合に急速な進行の良好な尤度または悪い尤度を有する腫瘍)を横切って階層化されるように、採用することを含む。このトレーニングセットを使用して予測モデル(上記した)を発生させ、これはデータの残り検定セット)に関して検定される。

10

【0118】

検定セットのより良好な性能を得るように予測モデルのパラメータを変えることが可能であるが、これは過学習として知られた状況をもたらすことがあり、過学習においては、予測モデルはトレーニングデータセットには働くが、任意の外部のデータセットには働かない。これを回避するために、バリデーションのプロセスが続く。典型的に適用される2つの主要なタイプのバリデーションがあり、第1(ホールドアウトバリデーション)は、データセットを、3つのグループ:検定、トレーニングおよびバリデーションに分配することを含む。バリデーションセットは、なんであれ、トレーニングプロセスへのインプットを持たず、それにより検定セット(しかしバリデーションセットではない)へのアプリケーション期間中、パラメータの任意の調節または他のリファインメントを行わなければならない。第2の主要なタイプは、クロスバリデーションであり、これは下記したいくつかの異なる方法でアプリケーションされうる。

20

【0119】

2つの主要なタイプのクロスバリデーションがある: K分割クロスバリデーション(K-fold cross-validation)およびリーブワンアウトクロスバリデーション(leave-one-out cross-validation)。

【0120】

K - 分割クロスバリデーション: データセットをK個のサブサンプルに分割し、各サブサンプルはオリジナルとほぼ同じ割合のクラスグループを含有する。

30

【0121】

バリデーションの各ラウンドにおいて、K個のサブサンプルの1つを取って置き、データセットの残りを使用してトレーニングを達成する。そのラウンドのトレーニングの有効性は、残して置いたグループの分類がいかにより正確であるかにより判定される。この手順をK回繰り返し、そして全体の有効性は、既知のクラスとの予測されたレベルの比較により確認される。

【0122】

リーブワンアウトクロスバリデーションは、 $K = n$ (ここでnはサンプルの数である)である、K分割クロスバリデーションのよく使用される変法。

40

【0123】

M P M Sの組み合わせ, 例えば表1に上記したM P M Sの組み合わせを使用して予後のための予測モデルを構築することができる。

【0124】

予後シグネチャー

これらのマーカーの1つ以上を含む予後シグネチャーを使用して、シグネチャー由来の1つ以上の予測モデルのアプリケーションにより患者のアウトカムを決定することができる。特に、臨床医または研究者は、シグネチャーにおける1つ以上のマーカーのディファレンシャルな発現(例えば、増加した発現または減少した発現)を決定し、予測モデルを

50

アプリケーションし、それによりネガティブな予後、例えば患者の疾患再発の尤度、またはポジティブな予後（連続した寛解）の尤度を決定することができる。

【0125】

予後シグネチャーを開発した。下記実施例で述べられるとおり、22の遺伝子を含む予後シグネチャーをメラノーマを有する患者のセットから確立した（表1）。患者サンプル（例えば腫瘍サンプル）を得、そしてサンプルにおける1つ以上のマーカーの発現レベルをディファレンシャルな発現プロファイルにマッチングさせることにより、癌進行の尤度を迅速に決定することができる。

【0126】

薬物試験

本発明は、特定の薬物試験のための個体を選択するのに使用することもできる。メラノーマを有する個体の予後を確立することにより、患者が応答しそうな慣用の処置を受けるべきかどうか、または患者が特定の薬物試験に参加するべきかどうか、即ち、特定の腫瘍タイプもしくは腫瘍段階を目標とするべきかどうかに関してより良い決定をすることができる。

【0127】

疾患進行への短い予測時間を有する患者の選択は、薬物試験の期間を短縮することも可能とし、そして記録されるべきより少数の患者が統計的に有意な薬物応答データを達成することを可能とする。

【0128】

実施例

ここに述べられた実施例は、本発明の態様を説明する目的のためである。他の態様、方法および分析のタイプは、分子診断技術分野における当業者の範囲内にあり、ここで詳細に述べる必要はない。当技術分野の範囲内の他の態様は、本発明の一部であると考えられる。

【0129】

段階IIIメラノーマにおける臨床アウトカムに影響を与えることがある腫瘍内の生物学的メカニズムを調査するために、段階IIIBおよびIIICメラノーマのリンパ節切除後の様々な臨床的アウトカムを有する患者からの29のメラノーマ標本の最初の検定セットに関して遺伝子発現プロファイル化を行った。次いでこれを使用して、10人の患者および14人の患者を含む2つの独立したバリデーションセットにおける分子プロファイルに基づいて、臨床的アウトカムを将来的に予測した。この分子情報を使用して、2つの患者グループ間でディファレンシャルにレギュレーションされることができるそして治療介入のための可能なターゲットである、細胞経路およびネットワークも同定された。

【0130】

物質および方法

マイクロアレイ分析のための標本採集および選択

行われた実験の全体的なスキームを図3に示す。Austin Healthで1997~2004年の間に臨床的に触診可能な節の外科リンパ節切除を受けた29人の患者からのex vivoメラノーマ組織をマイクロアレイ分析のために選択した。すべての標本は、Austin Healthヒト調査倫理委員会により認可された組織入手プロトコールの下にそして各患者の書面のインフォームドコンセントを伴って採集された。急に凍結した標本を最適切削温度化合物（OCT）中に包埋しそしてLudwig/Austin組織バンク貯蔵室内に-80で組織ブロックとして貯蔵した。すべての場合に診断は病理学者により確認された。

【0131】

患者サンプルは、段階IIIから段階IV疾患への腫瘍進行に要した時間（TTP）に基づいてマイクロアレイ分析のために選択され、そして16人の「悪い」（平均TTP 4か月）予後の患者および13人の「良好な」（平均TTP 42か月）予後の患者を含んでいた。専用のメラノーマユニットにおける手術後の検査を、リンパ節切除後の最初の12

10

20

30

40

50

カ月間毎月行い、次いで臨床的必要性に従ってその後3か月および6か月検査を4年まで行い、その後年に1回の検査を行った。段階調査は、臨床的疑いに応じてまたはルーチンに3~6か月ごとに行った。

【0132】

組織は、最小の壊死が存在しそして腫瘍細胞が全細胞集団の少なくとも60%を含むならばこの研究のために許容されうると考えられた。RNA抽出時に、2つの5 μ m切片を切断しそしてヘマトキシリンおよびエオシンで染色して、抽出された組織のインテグリティを確実にした。

【0133】

RNA抽出およびcDNA合成

cDNA合成および共通の参照デザインとのハイブリダイゼーションを29人の選択された患者について二重に(in duplication)行った。トータルRNAを、OCT包埋した組織から、組織切片をTri試薬(Molecular Research Center, Cincinnati, OH)中に浸漬および均質化することにより抽出し、クロロホルム1.5mLを均質化物に加え、サンプルを遠心しそして上部層を取り出し、そして100%エタノールと混合した。RNeasyカラムを使用する精製を、製造者のインストラクション(Quiagen, Valencia, CA)に従って行った。RNA品質を260:280吸光度の比に基づいて確認し、そしてインテグリティをrRNA標準マーカに対してホルムアルデヒド-アガロースゲル上で検査した。オリゴ(dT)およびアミノアリルデオキシヌクレオチドの存在下にRNA20 μ gからcDNAを合成した。Cy染料(Amersham Biosciences, Buckinghamshire, UK)を腫瘍cDNAおよび並行して産生された参照cDNAにカップリングさせた。参照cDNAをメラノーマを含む種々の腫瘍および細胞系からのプールされたRNAおよび正常な組織からのプールされたRNAから合成した(図4)。

【0134】

オリゴヌクレオチドアレイおよびデータ分析

個々の遺伝子および内部コントロールを表す30,888のオリゴヌクレオチドプローブを、MWG Biotech(Erbesberg, Germany)から入手し、そしてOmniGridロボット(Gene Machines, San Carlos, CA)を使用して高密度アレイとしてスポットした。標識された腫瘍/参照cDNAを共ハイブリダイゼーションし、そしてGenePix 4000Aマイクロアレイスキャナー(Axon Instruments, Union City, CA)を使用してスキャンした。マトリックスオーバーレイをスキャンされたイメージにアラインさせ、そして特徴抽出を、GenePix v.6.0ソフトウェア(Axon Instruments, Foster City, CA)を使用して行った。生のデータをGeneSpring v7.2(Silicon Genetics, Redwood City, CA)を使用して分析した。データをプリントチップグループ(print-tip group)に対して正規化し、次いでメディアン正規化した(median normalized)。簡単に言えば、lowess曲線を対数密度プロット対対数比プロットに適合させた。データの20%を使用して各点におけるlowess適合を計算した。この曲線を使用して各測定値についてコントロール値を調節した。次いで各遺伝子をすべてのサンプルにおけるその測定値のメディアンにより除した。

【0135】

EORTCメラノーマ研究²⁷からの独立したバリデーションセットについてのデータは、Array Express公的データ貯蔵所:<http://www.ebi.ac.uk/arrayexpress/>を通じて入手可能となった。データをGenespring v7.2にアップロードし、そしてスポットにつき、チップにつきおよび遺伝子につき正規化した。簡単に言えば、各遺伝子の測定された強度を、各サンプルにおけるそのコントロールチャンネル値により割り、次いでそのサンプルにおけるすべての測定値の第50百分位数(50th percentile)により除した。最後に各遺伝子を、すべてのサンプルにおけるその測定値のメディアンで除した。ディファレンシャルに発現された遺伝子についての発現値を使用して、下記したとおりの予測スコアを計算した。

【0136】

統計的方法

遺伝子発現データを、最初に、すべてのサンプルにおいて存在しなかったプローブを排除するフィルターに供した。考慮された最初の30,888のプローブの内、18,807はこのフィルターに通しそして分散分析、階層的クラスタリング(hierarchical clustering)および主成分分析のために使用された。ディファレンシャルに発現された遺伝子は、0.05のp値カットオフに基づく多重検定補正(multiple testing correction)のために補正するのに使用されるBenjamini and Hochbergの偽発見率制御方法(false discovery rate controlling method)²⁸と共に、ウイルコクソン-マン-ホイットニー検定(Wilcoxon-Mann-Whitney test)を行うことにより発見された。サンプルの階層的クラスタリングは、距離関数としてスピアマン相関(Spearman correlation) および平均リンケージ(average link)を使用して行われた。

10

【0137】

定量的リアルタイムPCR(qPCR)

アレイ結果を確かめるためにディファレンシャルに発現された遺伝子に関してqPCRを行い、次いでバリデーションセットAを使用して予測子をバリデーションするのに行った。第1鎖cDNAを、ランダムヘキサマープライマー(Promega, Madison, WI)を使用して、アレイ実験のために抽出されたトータルRNA 2 μgから合成した。逆転写を省くことによりネガティブコントロールを得た。Universal Probe Library アッセイデザインセンター<https://www. Roche-applied-science.com> (Roche, Mannheim, Germany) を使用して、イントロンスパニングマルチプレックスアッセイをqPCRのためにデザインした(アッセイデザインについては図5参照)。すべての反応は、ABI 7700配列検出器(Applied Biosystems, Foster City, CA)を使用して二重に行った。サーマルサイクラー条件は、下記のとおりであった: 50 2分間、95 10分間、続いて40サイクルの94 で20秒および60 で45秒。すべての結果を18S増幅に対して正規化した(Applied Biosystems, Foster City, CA)。本発明者らは、本発明者らのコンパレーターとして参照のためのターゲット閾値(C_T)の値を使用して相対的発現を計算した²⁹。

20

【0138】

次いで個々の遺伝子についての相対的発現値を、正規化されたlog₂比アレイ値および計算された相関係数を一緒にプロットした。

【0139】

結果

検定セットおよびバリデーションセットAに含まれた患者の臨床的特徴および病理学的特徴を列挙する(図6)。全ての患者は、最初の診断時の年齢、性およびポジティブなリンパ節転移の数および位置に関する情報を有していた。すべてではない患者が本発明者らの病院で初期診断をされ、それ故いくらかの症例では原発性メラノーマにおいて潰瘍化が存在していたかどうかを確かめることはできなかった。原発性メラノーマにおける潰瘍化は、もし存在するならば疾患をIIBからIICに段階を上げる独立した予後因子である³⁰。

30

【0140】

「良好な」予後グループの平均TTPは、「悪い」グループの4カ月に比べて40カ月であった。「良好な」グループはより若く見えそしてより多くの女性を含んでいたけれども、グループ間でメディアン年齢および性における統計的に有意な差はなかった。サンプルサイズの限界はあったけれども、AJCC病期分類、アジュバントインターフェロンの使用および腫瘍浸潤リンパ球の存在を含む他の既知の予後特徴において統計的に有意な差はなかった。

40

【0141】

1人の患者は、切除された脾臓に限られた孤立した段階IV疾患を有していたが、彼らが無病であることを考えて、このサンプルは含められた。このサンプルの排除は遺伝子発現プロファイルを変えなかった。

【0142】

50

ディファレンシャルに発現された遺伝子は、2つの予後グループを分離する。

教師なしの階層的クラスタリングは、予後とも他の臨床的情報とも相関しないメラノーマのサブグループを示さなかった。これはサンプル間の類似性を考えると予期されたことであった。予後グループを有効に分離することができた遺伝子を調査するために、ディファレンシャルな遺伝子発現を調べた。2,140の遺伝子が2つのグループ間でディファレンシャルに発現されたが、しかしながら、多重検定補正の厳格な適用は、これを高度に有意なディファレンシャルな発現を有する22の遺伝子に減少させた(図1)。22の遺伝子を、qPCRを使用してトレーニングセットにおいて更にバリデーションし、そして2つのプラットフォーム間で最も高い相関係数を有する遺伝子($r > 0.5$ 、 $p < 0.05$)が更なる分析のために選ばれた(データは示されていない)。最初の22の内、15の遺伝子は高いクロスプラットフォーム相関(cross-platform correlation)を示し、そしてこれらは予測スコアの開発のために使用された。主成分分析は、15の遺伝子の予後グループを分離する能力を証明した(図7)。

10

【0143】

予測スコアの開発

最初の検定セットを使用して2つの独立したバリデーションセットに関して検定された予測子を開発した。2つの予測アルゴリズムは、アレイデータおよびqPCRデータに基づいて開発された。

【0144】

1. アレイデータ(aPS)についての予測スコアを計算するために、アレイとqPCRとの間の最も有意な相関を有する15の遺伝子を使用した。正規化された \log_2 発現比は、その値を2乗することにより変換された。「良好な」予後グループにおいてダウンレギュレーションされた遺伝子はネガティブな値に帰された。次いですべての15の遺伝子についての値の和により最終スコアを計算した。ポジティブなスコアは改善されたアウトカムと関連していた。

20

【0145】

2. qPCRデータ(qPS)について、15の最も相関した遺伝子についての C_T 値をAkaike Information Criterionを利用するロジスティック回帰アルゴリズムにアプリケーションしてクラス識別に寄与する遺伝子のみを選択した。これは5つの有意な遺伝子を選択し、これらは次いで下記式において使用された：

30

$$qPS = [1328.15 - 187.42(\text{IDH}) + 137.10(\text{MFG8}) + 73.61(\text{PILRA}) + 211.22(\text{HLA-E}) + 143.94(\text{TXNDC5})] \times -1$$

aPSに同じく、ポジティブスコアは、この方法を使用して改善されたアウトカムと関連していた。

【0146】

予測スコアはTTPおよび全生存期間と相関する

予期されたとおり、検定セットにアプリケーションされたaPSおよびqPSの両方が、2つの予後グループを識別することができた。個々のスコアの大きさ(aPSによる高いスコアおよびqPSでのネガティブスコア)がqPSおよびaPSの両方について改善されたアウトカムと相関するような(図8、スピアマン順位相関、 $r = 0.7908$ 、 $p < 0.0001$)、個々のスコアとTTPおよび全生存期間との強い相関は明白であった。これは、これらのディファレンシャルに発現された遺伝子の発現レベルが、臨床的アウトカムに直接影響を与える基礎をなす生物学的機序に関係することを示唆しており、それらの予後の妥当性を強調している。

40

【0147】

3つの独立したセットへの予測スコアのアプリケーション

次いで結果を、独立に発生させたデータにアプリケーションした。我々自身に類似した患者のサブグループに関する1つの公表されたデータセットが同定された。この研究²⁷においてプロフィール化された83人の患者の内、14人は長期追跡により段階III疾患を有していた。このサブグループでは、我々の検定セットにおいて適用された同様な基準

50

を使用して、10人の患者は「悪い」として分類され（平均TTP10カ月）、そして4人は「良好」として分類された（平均TTP6.2カ月）。aPSアルゴリズムをこれらのサンプルにアプリケーションすると、すべての10人の「悪い」患者および4人の「良好な」患者の内2人は正しく予測され、85%の総合的な正しい分類率(overall correct classification rate)を生じる。

【0148】

次いで、我々は、Ludwig/Austrian組織バンクからの10の腫瘍の独立したセットにqPSアルゴリズムをアプリケーションし、これらに対するqPCRアッセイを5つの最も強力な予測性の遺伝子を使用して行った。予測子は「良好な」予後の腫瘍の5つのすべてを正しく分類したが、5つの「悪い」サンプルの1つを間違えて分類した（図9）。間違えて分類された「悪い」サンプルは、TTPは短いが転移疾患を伴って6年という長期の全生存期間を有した患者を表した。

10

【0149】

5つの遺伝子qPSも、段階3メラノーマサンプルの第3の独立したセットにアプリケーションした。これらのサンプルは、段階3疾患の診断の後18か月以下の生存期間を有する19人の患者および段階3診断から4年より長く生存した更なる18人の患者からなっていた。これらの良好な予後グループからのqPSスコアの分布と悪い予後グループからのqPSスコアの分布は、有意に異なっており（ $p = 0.02$ ）そして図10に示される。

20

【0150】

検討

この実施例は、マイクロアレイ遺伝子発現データおよびqPCRから誘導された発現プロフィールを使用して段階IIIメラノーマ患者の他の方法では識別できないグループにおける臨床的アウトカムの成功した予測を示す。2つの独立したセットにおいて、15のディファレンシャルに発現された遺伝子に基づく2つの開発された予測スコアルゴリズムを、マイクロアレイおよびqPCRデータにアプリケーションして、段階IIIB/Cメラノーマを有する患者における臨床的アウトカムを将来的に予測することができることが確立された。

【0151】

これらの患者は、同様な段階の疾患について選ばれ、そしていくつかの研究は、同様な段階の疾患を有する患者間よりも異なる段階で採取された自己由来のサンプル間の遺伝子発現のより高い類似性を証明した^{27、31、32}。92%以下の正確度でアウトカムを将来的に予測するのに使用されうるグループ間でディファレンシャルに発現された遺伝子があるという観察は、それらの重要性を強調する。更に、TTPおよび全生存期間の両方との予測子の相関は、スコアの差の大きさが臨床的アウトカムと直接相関するような予測子の有用性も強調する。

30

【0152】

説明において、既知の均等物を有する整数または成分に関して言及された際、このような均等物は、あたかも個々に説明されたかのように、本明細書に組み込まれる。本発明は、実施例としておよびその可能な態様に関して説明されているけれども、本発明の範囲から逸脱することなく改善および/または改変がなされうることは認識されるべきである。

40

【0153】

【表 2】

参考文献

1. Australian Institute of Health and Welfare (AIHW): Cancer in Australia 2001. Canberra, Australian Institute of Health and Welfare
Australasian Association of Cancer Registries (AACR), 2004
2. Florez A, Cruces M: Melanoma epidemic: true or false? *Int J Dermatol* 43:405-7, 2004
3. Thursfield V, Farrugia H, Giles G: Cancer in Victoria 2004, Canstat. Victoria, Cancer Epidemiology Centre, 2006, pp 32
4. Thompson JF, Scolyer RA, Kefford RF: Cutaneous melanoma. *Lancet* 365:687-701, 2005
5. Verma S, Quirt I, McCreedy D, et al: Systematic review of systemic adjuvant therapy for patients at high risk for recurrent melanoma. *Cancer* 106:1431-42, 2006
6. Hersey P: Adjuvant therapy for high-risk primary and resected metastatic melanoma. *Intern Med J* 33:33-43, 2003
7. Kirkwood JM, Manola J, Ibrahim J, et al: A pooled analysis of eastern cooperative oncology group and intergroup trials of adjuvant high-dose interferon for melanoma. *Clin Cancer Res* 10:1670-7, 2004
8. Sondak VK, Sabel MS, Mule JJ: Allogeneic and autologous melanoma vaccines: where have we been and where are we going? *Clin Cancer Res* 12:2337s-2341s, 2006
9. Balch CM, Sober AJ, Soong SJ, et al: The new melanoma staging system. *Semin Cutan Med Surg* 22:42-54, 2003
10. Kirkwood JM, Strawderman MH, Ernstoff MS, et al: Interferon alfa-2b adjuvant therapy of high-risk resected cutaneous melanoma: the Eastern Cooperative Oncology Group Trial EST 1684. *J Clin Oncol* 14:7-17, 1996
11. Kirkwood JM, Ibrahim JG, Sondak VK, et al: High- and low-dose interferon alfa-2b in high-risk melanoma: first analysis of intergroup trial E1690/S9111/C9190. *J Clin Oncol* 18:2444-58, 2000
12. Efron, B. and Tibshirani, R. *An Introduction to the Bootstrap*. Chapman & Hall. 2005
13. McLaughlan GJ, Do K, Ambrose C *Analyzing Microarray Gene Expression Data* (Wiley Series in Probability and Statistics) 2004
14. Wright GW, Simon RM A random variance model for detection of differential gene expression in small microarray experiments. *Bioinformatics* 2003;19:2448-2455.
15. Smyth GK. Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. *Statistical Applications in Genetics and Molecular Biology* 2004;3:Article 3.

10

20

30

40

16. Lönnstedt I. and Speed TP. Replicated microarray data. *Statistica Sinica* 2002;12:31-46.
17. Ihaka R, Gentleman R. R: A language for data analysis and graphics. *Journal of Computational and Graphical Statistics* 1996;5:299-314.
18. Becker RA, Chambers, JM and Wilks AR *The New S Language*. Wadsworth & Brooks/Cole 1988.
19. Hastie T, Tibshirani R, Friedman J *The Elements of Statistical Learning Data Mining, Inference and Prediction* Springer 2003 10
20. Gentleman R., Carey VJ, Huber W., Irizarry RA, Dudoit S. *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*. Springer 2005.
21. Neter J, Kutner MH, Wasserman W, Nachtsheim CJ, *Applied Linear Statistical Models* McGraw-Hill/Irwin 1996
22. Venables, WN, Ripley, BD *Modern Applied Statistics with S*. 4th ed.. Springer 2002.
23. Ripley, B. D. *Pattern Recognition and Neural Networks* Cambridge University Press 1996
24. Cristianini N, Shawe-Taylor J *An Introduction to Support Vector Machines (and other kernel-based learning methods)* Cambridge University Press 2000 20
25. Breiman L, Friedman J, Stone CJ, Olshen RA *Classification and Regression Trees* Chapman & Hall/CRC 1984
26. Good, PI *Resampling Methods: A Practical Guide to Data Analysis* Birkhauser 1999
27. Winnepeninckx V, Lazar V, Michiels S, et al: Gene expression profiling of primary cutaneous melanoma and clinical outcome. *J Natl Cancer Inst* 98:472-82, 2006
28. Benjamini Y, Hochberg Y: Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society* 57:289-300, 1995 30
29. Livak KJ, Schmittgen TD: Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C(T)}$ Method. *Methods* 25:402-8, 2001
30. Balch CM, Sober AJ, Soong SJ, et al: The new melanoma staging system. *Semin Cutan Med Surg* 22:42-54, 2003
31. Wang E, Miller LD, Ohnmacht GA, et al: Prospective molecular profiling of melanoma metastases suggests classifiers of immune responsiveness. *Cancer Res* 62:3581-6, 2002 40
32. Ramaswamy S, Ross KN, Lander ES, et al: A molecular signature of metastasis in primary solid tumors. *Nat Genet* 33:49-54, 2003

【 0 1 5 4 】

工業的用途

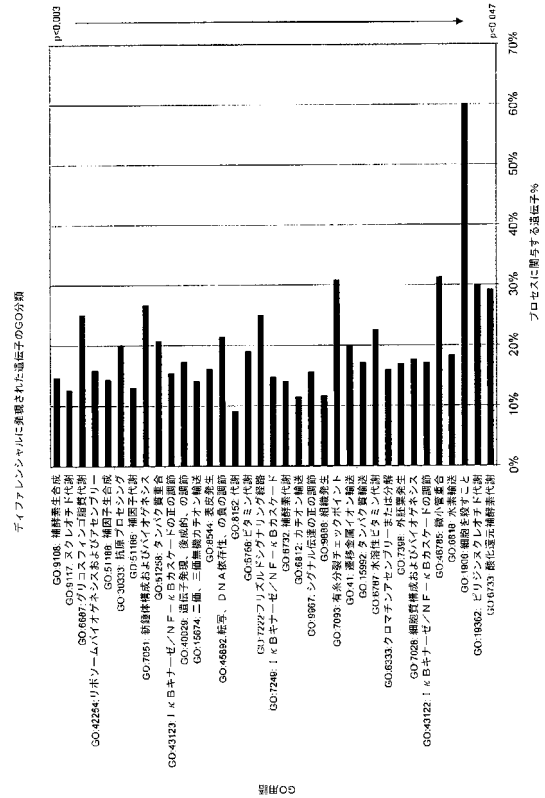
予後癌マーカー、特にメラノーマ予後マーカーに基づく本発明の方法、成分、キットおよび装置は、癌、特にメラノーマの予後および処置のために有用である。

【 図 1 】

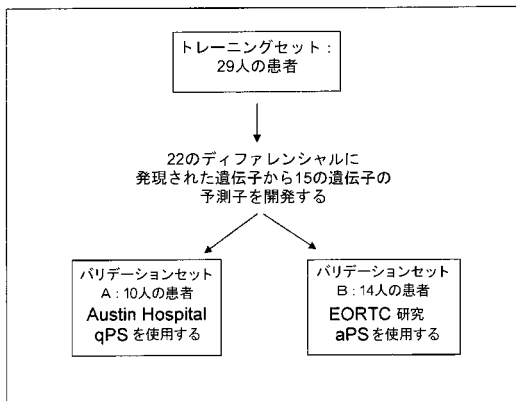
プローブ名称	P値	登録名	Genebank	説明
megihuman30k4A.3667	0.049	TXNDG5	NM_030810	チナレドキノンドメイン含有 5
megihuman30k4A.00694	0.049	PILRA	NM_013439	ベナード免疫グロブリン様 2 型セプター α
megihuman30k4A.06347	0.049	HLAE	NM_005516	主要組織適合性複合体、クラス II、E
megihuman30k4E5.2443	0.048	TPA	NM_009163	Human T cell leukemia virus type 1
megihuman30k4E5.1847	0.048	DMN	NZ_146728	チヌムスリン、スプアファターゼ (ヌクレオチド三リン酸デヒドロスファターゼ)
megihuman30k4A.07569	0.042	CTPBP2	NM_019098	GTP結合タンパク質 2
megihuman30k4A.09053	0.0429	MFSFR	NM_005928	乳糖酸様 -EGF 因子タンパク質
megihuman30k4A.06363	0.0365	DH1	NM_005886	イソクエン酸チロシドドクターゼ 1 (NADPH)、可溶性
megihuman30k4E5.9466	0.0305	MRFPS6	NM_031902	ミトコンドリアリボソームタンパク質 S5
megihuman30k4A.04768	0.0307	LGALS7	NM_002307	レクチン、ガラクトシド結合性、可溶性、7 (ガラクチン?)
megihuman30k4E5.04955	0.0295	KCNIP2	AF347114	Kvチャネル相互作用タンパク質 2
megihuman30k4C.4262	0.0295	CHS1A	NM_005769	加水化糖 (N-アセチルグルコサミン 6-O) スルホトランスフェラーゼ 4
megihuman30k4E5.03441	0.023	OSL1/170_6CB	AL451391.1.67295.96699.1	アラングリン結合タンパク質 1 (アラングリン)
megihuman30k4E5.03441	0.023	MFE1B	NM_002903	両極性細胞分裂調節因子 1
megihuman30k4A.00949	0.023	MTCF2	NM_014942	ミトコンドリアリボソームタンパク質 2
megihuman30k4A.04349	0.0138	ARFRP1	NM_003224	MED12 結合タンパク質
megihuman30k4C.0720	0.0138	BAB1L	AJ131063	ADP-リボソイル化因子関連タンパク質 1
megihuman30k4E5.4873	0.0138	TUBA1B	NM_005862	カバノキ花物アレルゲン-特異的免疫グロブリン 1
megihuman30k4E5.00371	0.00371	PLXNB2	AJ426560.2	チヌムリン α 1b...
megihuman30k4E5.00331	0.000756	PLXNB2	AE002313	部分的 n-myc エキソン 3

* このマーカーは以前に Kyo0353 として知られていた。dmn (KM_031031)
 ** このマーカーは免疫グロブリン様 1-α として知られていた。(GKC_AJ131063)
 *** このマーカーは、以前にチヌムリン α 1b と呼ばれていた。loc143712 (XM_084610)

【 図 2 】



【 図 3 】



【 図 4 】

細胞系または 組織系	ソース	重量 (g)	容積 (mL)	濃度 (mg/mL)	合計 (mg)	ローディング された容積
786	腎臓癌	-	7.25	0.71	5.2	2.8
Ovcar 3	卵巣癌	-	5	0.82	4.1	2.4
Ovcar	卵巣癌	-	4.5	2.21	10	0.9
J82	膀胱癌	-	5	1.11	5.6	1.8
AGS	胃癌	-	4.5	1.2	5.4	1.7
CCRF-CEM	T細胞白血病	-	2.5	0.34	0.9	5
NCI-H460	肺癌	-	7.5	1.2	9	1.7
U251	脳癌	-	4	2.27	9.1	0.8
COLO	結腸癌	-	4	1.55	6.2	1.3
M14	メラノーマ	-	8	0.15	1.2	5
T24	膀胱癌	-	7	1.15	8	1.7
MCF7	乳癌	-	4.5	2.8	12.6	0.7
PC3	前立腺癌	-	9.5	1	9.5	2
HL60	急性骨髄性白血病	-	5	0.58	2.9	3.4
CACO6	結腸癌	-	4.5	0.9	4.1	2.2
KATO III	胃癌	-	4	0.58	2.3	3.5
C33A	頸部癌	-	4.5	1.41	6.3	1.4
組織	正常な胃	1.5	2	0.36	0.7	5
Wilms Rhabdo	腎臓/肉腫	9	4.5	0.3	1.4	5

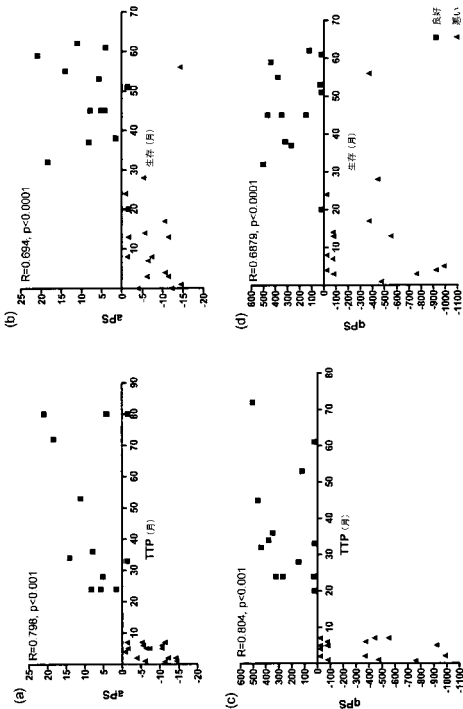
【 図 5 】

遺伝子名	Genbank	説明	起病年齢	若年型	高齢型	別	年齢	年齢
TANC5	NM_030810	チロシンリンゲイムタンパク質	1	actgagcactcaatg	2	14	64	64
PILRA	NM_013439	ヘマトポエチン様因子2超家族メンバー4	3	agggagcaaacctccac	4	64		
HLA-E	NM_005516	主要組織適合性複合体、クラスI、E	5	ggaagcagcagcagggg	6	89		
ITPA	NM_003453	インソリン結合型アミンオキシド酵素	7	agagcactgcccaccac	8	17		
GTF2IP2	NM_019096	GTF2I結合タンパク質2	9	ggggcagcagcactcca	10	47		
MFCB3	NM_005528	乳癌特異的GTF2I結合タンパク質	11	ctctcaggagccacacac	12	21		
IDH1	NM_003896	イソクエン酸脱水素ドナルドドナーゼ1 (HADP+)	13	gagcaaacctctcaatgact	14	77		
MIR55	NM_011969	ミトコンドリア7Sリボソームタンパク質55	15	gcatccaccacagcagc	16	17		
KCNIP2	AF341194	カリウムチャンネル補助タンパク質2	17	ttatagcagcagcagc	18	65		
CHST4	NM_005769	キチンキチナシアンヒドロラーゼ4	19	ttatagcagcagcagc	20	19		
NFKB1B	U45752	核因子カプシムB1 (NF- κ B2)	21	aggcagcagcagcagc	22	72		
MITCH2	NM_014342	ミトコンドリアATPaseサブユニット2 (F ₁ F ₀)	23	aggcagcagcagcagc	24	76		
ARFRP1	NM_003224	ADP-リポソームATPase補助タンパク質1	25	gagcagcagcagcagc	26	5		
RLXN1	NM_004510	ニューロリン酸1 (Nucleolin)	27	gctcagcagcagcagc	28	4		
ADPS2	AF05522	アデノスインゼン2	30	atctctgagcagcagc	30	64		
ADPS2	AF05522	アデノスインゼン2	31	ccagcagcagcagcagc	32	16		
ADPS2	AF05522	アデノスインゼン2	33	ccagcagcagcagcagc	34	30		

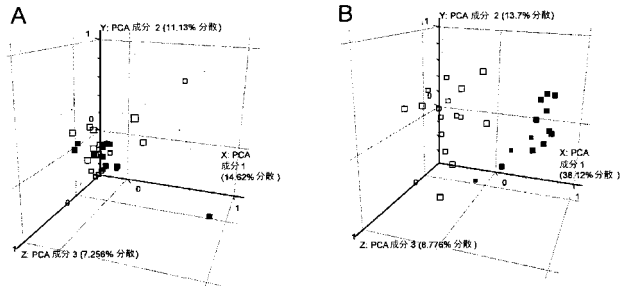
【 図 6 】

メジアン年齢 (年)	トレーニングセット n=29		P値	バリデーションセット A n=10		P値
	良好な (n=13)	悪い (n=16)		良好な (n=5)	悪い (n=5)	
35 (19-89)	48.5 (28-74)	0.2275†	35 (24-73)	49 (43-72)	0.8125†	
AJCC 段階						
IIIB	4	8	0.394†	3	4	0.5647†
IIIC	7	7		1	1	
IIIU	1	1		1	0	
IV	1	0		0	0	
性						
男性	5	8	0.4495†	4	5	0.2723†
女性	8	6		1	0	
平均 TTP (月)	40.3 (24-72)	4.1 (0-6)	<0.001§	52.2 (23-108)	6.6 (1-15)	<0.0256§
腫瘍浸潤リンパ球						
活発	3	0	1‡	0	0	0.7523†
活発でない	4	8		3	3	
存在しない	1	1		1	0	
得られない	5	7		1	2	
アジュバント						
インターフェロン	1	1		1	0	
IIIU=少なくともIIIB						
における不完全なデータ						
†ワイルコソン順位相検定						
‡フィッシャーの正確検定						
§コックス比例ハザード						

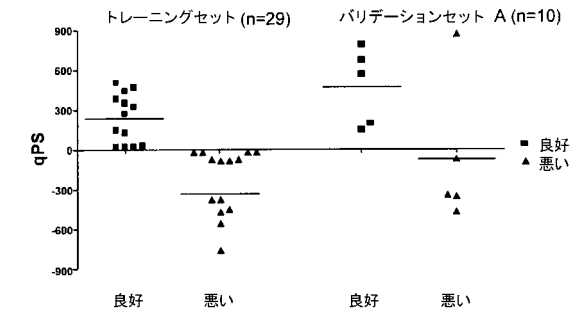
【 図 8 】



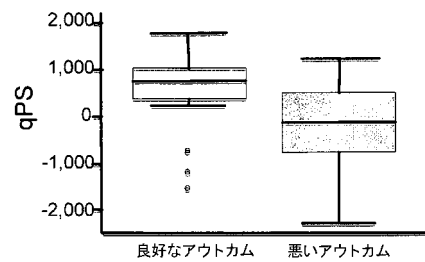
【 図 7 】



【 図 9 】



【 図 10 】



【配列表】

2020150949000001.app

【手続補正書】

【提出日】令和2年7月3日(2020.7.3)

【手続補正1】

【補正対象書類名】特許請求の範囲

【補正対象項目名】全文

【補正方法】変更

【補正の内容】

【特許請求の範囲】

【請求項1】

本願明細書に記載される発明。

フロントページの続き

- (72)発明者 ブラック, マイケル・アラン
ニュージーランド国、ダニディン、ロズリン、ファルコン・ストリート 17
- (72)発明者 セボン, ジョナサン
オーストラリア国、ビクトリア 3068、クリフトン・ヒル、グラント・ストリート 6
- (72)発明者 ギルフォード, パリー・ジョン
ニュージーランド国、ダニディン、イースト・タイエリ、リカトン・ロード 38
- (72)発明者 ジョン, トーマス
オーストラリア国、ビクトリア 3102、イースト・キュー、ベーカー・アベニュー 44
- Fターム(参考) 2G045 AA26 CB02 DA14
4B029 AA07 AA23 BB11 BB20
4B063 QA01 QA19 QQ02 QQ53 QR32 QR55 QR62 QS25 QS34 QS39

【外国語明細書】
2020150949000001.pdf