

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第6366580号  
(P6366580)

(45) 発行日 平成30年8月1日(2018.8.1)

(24) 登録日 平成30年7月13日(2018.7.13)

(51) Int.Cl.		F I		
<b>C 1 2 Q</b>	<b>1/6886</b>	<b>(2018.01)</b>	C 1 2 Q	1/6886 Z N A Z
<b>C 1 2 N</b>	<b>15/09</b>	<b>(2006.01)</b>	C 1 2 N	15/09 2 0 0
<b>G O 1 N</b>	<b>33/53</b>	<b>(2006.01)</b>	G O 1 N	33/53 M

請求項の数 19 (全 105 頁)

(21) 出願番号 特願2015-518637 (P2015-518637)  
 (86) (22) 出願日 平成25年6月24日 (2013.6.24)  
 (65) 公表番号 特表2015-521483 (P2015-521483A)  
 (43) 公表日 平成27年7月30日 (2015.7.30)  
 (86) 国際出願番号 PCT/US2013/047354  
 (87) 国際公開番号 W02013/192616  
 (87) 国際公開日 平成25年12月27日 (2013.12.27)  
 審査請求日 平成28年5月27日 (2016.5.27)  
 (31) 優先権主張番号 61/663,428  
 (32) 優先日 平成24年6月22日 (2012.6.22)  
 (33) 優先権主張国 米国 (US)

(73) 特許権者 512112622  
 エイチティージー モレキュラー ダイア  
 グノスティクス, インコーポレイテッド  
 アメリカ合衆国 アリゾナ 85706,  
 ツーソン, イー. グローバル ルー  
 プ 3430  
 (73) 特許権者 507422150  
 ジョン ウェイン キャンサー インステ  
 イテュート  
 アメリカ合衆国 カリフォルニア州 サン  
 タ モニカ サンタ モニカ ブールバ  
 ド 2200  
 (74) 代理人 100078282  
 弁理士 山本 秀策

最終頁に続く

(54) 【発明の名称】メラニン細胞性病変における分子悪性度

(57) 【特許請求の範囲】

【請求項1】

メラニン細胞含有試料を特徴付ける方法であって、  
 被験体から得られたメラニン細胞含有試料中の、  
 ( i ) バイオマーカー MAGEA2、PRAME、PDIA4、NR4A1、PDLIM7、B4GALT1、SAT1、RUNX1、およびSOC3、ならびに  
 ( i i ) 少なくとも1つの正規化バイオマーカー  
 についての発現レベルを決定し、これにより、該バイオマーカーの各々および該少なくとも1つの正規化バイオマーカーについての生の発現値をもたらずステップと、  
該バイオマーカーの各々についての該生の発現値を、該少なくとも1つの正規化バイオマーカーについての該生の発現値に対して正規化して、該バイオマーカーの各々についての正規化された発現値をもたらずステップと、  
 該正規化された発現値を、回帰アルゴリズムまたは機械学習アルゴリズムにおいて使用して、出力値をもたらずステップと、  
 該出力値を、良性であるのか悪性であるのかがあらかじめ既知である複数のメラニン細胞含有試料中の、該バイオマーカーについての正規化された発現値から導出されたカットオフ値と比較するステップと、  
 該出力値が、該複数の既知の良性試料と、該カットオフ値の同じ側であれば、該試料を良性として特徴付けるか、または該出力値が、該複数の既知の悪性試料と、該カットオフ値の同じ側であれば、該試料を悪性として特徴付けるステップと

10

20

を含む方法。

【請求項 2】

前記少なくとも 1 つの正規化バイオマーカーが、母斑試料と原発性黒色腫試料との間で発現の統計学的有意差を有さない、請求項 1 に記載の方法。

【請求項 3】

前記少なくとも 1 つの正規化バイオマーカーが、表 3 のバイオマーカーのうちの 1 つ、2 つ、3 つ、4 つ、5 つ、6 つ、7 つ、8 つまたは 9 つ全てを含む、請求項 1 から 2 のいずれかに記載の方法。

【請求項 4】

前記メラニン細胞含有試料中の複数のバイオマーカーについての遺伝子発現値を測定するステップであって、該複数のバイオマーカーについての発現の範囲が、該試料のトランスクリプトームにおけるバイオマーカー発現の全範囲を表すステップと、

このような複数のバイオマーカーについての中心傾向発現値を計算するステップと、必要に応じて、外れ値を棄却し、該外れ値発現値を伴わずに再計算された複数の中心傾向発現値を計算するステップと

該中心傾向発現値、または、必要に応じて、該再計算された複数の中心傾向発現値を使用して、該バイオマーカーの各々についての前記生の発現値を正規化するステップとをさらに含む、請求項 1 または 2 に記載の方法。

【請求項 5】

被験体からのメラニン細胞含有試料中の遺伝子発現を決定する方法であって、  
該試料中の、バイオマーカー MAGEA2、PRAME、PDIA4、NR4A1、PDLIM7、B4GALT1、SAT1、RUNX1、および SOCS3を含む複数の遺伝子の発現レベルを決定するステップと、

該試料中の該複数の遺伝子の発現レベルの報告、または該試料の、該複数の遺伝子の該発現レベルに基づく、母斑または黒色腫としての特徴付けを提供するステップとを含む方法。

【請求項 6】

前記バイオマーカーがさらに、表 4 のバイオマーカーのうちの少なくとも 2 つを含み、前記出力値が、ロジスティック回帰アルゴリズムによりもたらされた、請求項 1 から 5 のいずれかに記載の方法。

【請求項 7】

前記アルゴリズムが、  
出力値 =  $0 + 1 \times 1 + 2 \times 2 + \dots + n \times n$   
[ 式中、 $X_n$  は、前記バイオマーカー MAGEA2、PRAME、PDIA4、NR4A1、PDLIM7、B4GALT1、SAT1、RUNX1、および SOCS3 についての発現値の対数であり、 $0$  は、 $-200$  を超え、かつ、 $200$  未満であり、 $n > 0$  に対する全ての  $n$  は、 $-1, 000$  を超え、かつ、 $1, 000$  未満である ]  
である、請求項 1 から 6 のいずれかに記載の方法。

【請求項 8】

前記バイオマーカーがさらに、表 11 および 13 のバイオマーカーのうちの少なくとも 2 つを含み、前記出力値が、機械学習アルゴリズムによりもたらされた、請求項 1 から 5 のいずれかに記載の方法。

【請求項 9】

前記バイオマーカーがさらに、表 11 のバイオマーカーのうちの少なくとも 2 つまたは表 13 のバイオマーカーのうちの少なくとも 2 つを含む、請求項 8 に記載の方法。

【請求項 10】

前記バイオマーカーがさらに、  
BAX、BIRC5、MET、MAGEC2、POLR2J3、ZFYVE16、および BEST1 のうちの少なくとも 2 つ、または  
POLR2J3、BEST1、BIRC5、MET、ZFYVE16、HIF1A、お

10

20

30

40

50

よび P I C A L M のうちの少なくとも 2 つ  
を含み、

前記出力値が、機械学習アルゴリズムによりもたらされた、  
請求項 1 から 9 のいずれかに記載の方法。

【請求項 1 1】

前記バイオマーカーがさらに、表 4、表 1 1、表 1 3、または表 4 および表 1 1 の両方の遺伝子のうちの少なくとも 5 0 %、少なくとも 7 5 %、少なくとも 8 0 %、少なくとも 9 0 %、少なくとも 9 5 %、または少なくとも 9 8 % を含む、請求項 1 から 1 0 のいずれかに記載の方法。

【請求項 1 2】

バイオマーカー B 4 G A L T 1、B A X、M A G E A 2、N R 4 A 1、P D I A 4、P R A M E、R U N X 1、S O C S 3、S A T 1、P D L I M 7、B I R C 5、M E T、M A G E C 2、P O L R 2 J 3、Z F Y V E 1 6、および B E S T 1 の発現レベルを、被験体から得られたメラニン細胞含有試料における悪性度の指標とする方法であって、該方法は、

該バイオマーカーの該発現レベルを入力として使用するアルゴリズムからの出力を計算するステップ

を含み、

該アルゴリズムの出力の、既知の悪性メラニン細胞含有試料に由来する参照標準との比較が、該試料が悪性であるのか悪性でないのかの指標となる、方法。

【請求項 1 3】

前記バイオマーカーの前記発現レベルを、

( a ) M F I 2、R A P 2 B、B M P 1、および / もしくは N C O R 2 のうちの少なくとも 1 つ、

( b ) M F I 2、N C O R 2、R A P 2 b、および B M P - 1、

( c ) R P S 6 K B 2 および / もしくは S D H A、または

( d ) 前記メラニン細胞含有試料中で発現した少なくとも 1 つの遺伝子であって、該バイオマーカーではなく、その発現が、代表的な複数のメラニン細胞含有試料において有意に異なる遺伝子

からなる群から選択される少なくとも 1 つの正規化バイオマーカーの発現レベルに対して正規化するステップをさらに含む、請求項 1 2 に記載の方法。

【請求項 1 4】

発現レベルを決定するステップが、核酸の発現を決定することを含む、請求項 1 から 1 3 のいずれかに記載の方法。

【請求項 1 5】

核酸の発現を決定することが、前記試料を、複数の核酸プローブまたは対合増幅プライマー対と接触させることを含み、各プローブまたは対合プライマーは、該複数の核酸プローブまたは対合プライマーがその / それらの相補的なバイオマーカーとハイブリダイズすることを可能とする条件下で、前記バイオマーカーのうちの 1 つに対して特異的、かつ、相補的である、請求項 1 4 に記載の方法。

【請求項 1 6】

前記試料を前記複数の核酸プローブと接触させた後で、該試料を、一本鎖核酸分子を消化するヌクレアーゼと接触させることをさらに含む、請求項 1 5 に記載の方法。

【請求項 1 7】

少なくとも 1 0 のアドレス可能な位置であって、各位置が、同じ特異性を有する固定化された捕捉プローブを含み、各位置が、他の各位置における捕捉プローブとは異なる特異性を有する捕捉プローブを含む位置

を含む、メラニン細胞含有試料を特徴付けるためのアレイであって、

該少なくとも 1 0 の位置のうちの 9 つにおける該捕捉プローブが、M A G E A 2、P R A M E、P D I A 4、N R 4 A 1、P D L I M 7、B 4 G A L T 1、S A T 1、R U N X

10

20

30

40

50

1、およびSOC3からなるバイオマーカーに、直接的または間接的に、特異的にハイブリダイズすることが可能であり、該少なくとも10の位置のうちの1つにおける該捕捉プローブが、表3に列挙される正規化バイオマーカーに、直接的または間接的に、特異的にハイブリダイズすることが可能であり、

各捕捉プローブの該特異性が、該アレイの該アドレス可能な位置により同定可能であるアレイ。

【請求項18】

前記少なくとも10のアドレス可能な位置の各々が、個別に同定可能なビーズまたはフローセル内のチャンネルである、請求項17に記載のメラニン細胞含有試料を特徴付けるためのアレイ。

10

【請求項19】

請求項17または18に記載のメラニン細胞含有試料を特徴付けるためのアレイ、ならびに

溶解緩衝液を含有する容器；

一本鎖核酸に特異的なヌクレアーゼを含有する容器；

複数の核酸プログラミングリンカーを含有する容器；

複数のNPPを含有する容器；

複数の二官能性検出リンカーを含有する容器；

該二官能性検出リンカーに特異的に結合する検出プローブを含有する容器；および

検出試薬を含有する容器

20

のうちの1または複数

を含むキット。

【発明の詳細な説明】

【技術分野】

【0001】

関連出願への相互参照

本願は、2012年6月22日に出願された米国仮特許出願第61/663,428号の利益を主張し、この米国仮特許出願の全体の内容は、本明細書中に参考として援用される。

【0002】

30

分野

本開示は、メラニン細胞性病変を良性または悪性として特徴付けるためのバイオマーカーに関する。特に、本開示は、母斑試料および原発性黒色腫試料において有意に差次的に発現するバイオマーカー(mRNAおよび/またはmiRNAを含む)の同定、このようなバイオマーカーの発現に基づく臨床的な予測アルゴリズム、ならびにそれらの使用法およびそれらの使用のための組成物に関する。

【0003】

共同研究契約の関係機関

HTG Molecular DiagnosticsおよびJohn Wayne Cancer Instituteが、本明細書で開示される発明を統括する共同研究契約の関係機関である。

40

【背景技術】

【0004】

背景

皮膚がんは、米国では、全てのがんのうちで最も一般的である。メラニン細胞に由来するがんである黒色腫が皮膚がんのうちで占める百分率は比較的小さい。しかし、黒色腫は、最も多くの皮膚がんによる死亡を引き起こすことから、最も危険な種類の皮膚がんのうちの1つとなっている。2012年には、黒色腫は、75,000例を超える皮膚がん症例を占めることになる。

【0005】

50

メラニン細胞はまた、眼（例えば、ブドウ膜内もしくはブドウ膜上、毛様体内もしくは毛様体上、結膜内もしくは結膜上、眼瞼内もしくは眼瞼上、虹彩内もしくは虹彩上、または眼窩内もしくは眼窩上）、内耳、髄膜、骨、および心臓を含む、皮膚以外の器官でも見出される。眼内黒色腫は、成人において最も一般的な眼腫瘍の種類であり、体内の原発性悪性黒色腫のうちで2番目に多い一般的な種類である。眼内黒色腫の発生率は、米国では、百万人当たり約5例である。

【0006】

黒色腫を診断するには、疑わしい組織を生検し、病理学者、好ましくは（しかし、そうでないことも多い）組織生検中の黒色腫を同定するように専門的に訓練された病理学者が顕微鏡下で調査する。病理学者が、黒色腫の所見について報告する場合、いくつかの因子（ミリメートル単位の腫瘍の深さ、潰瘍化の存在もしくは非存在、有糸分裂速度、および/または腫瘍が広がっているのかどうかを含む）を、患者の予後診断および処置（複数可）の経過を決定するのに使用する。腫瘍が広がっていない場合も、より広範な局所的切除を実施して、黒色腫周囲の正常組織の明確な縁に沿って病変の全体が摘出されたことを確認することが多い。より高度な処置（more extreme treatment）が適応である場合、患者にはまた、リンパ節切除、免疫療法、化学療法、または放射線療法も施すことができる。

【0007】

黒色腫は、そのごく早期に発見すれば、ほとんど常に治癒可能である。残念ながら、黒色腫の誤診は一般的である（Piepkornら、J. Am. Acad. Dermatol.、30巻：707頁、1994年；Farmerら、Hum. Pathol.、27巻：528頁、1996年；Coronaら、J. Clin. Oncol.、14巻：1218頁、1996年；Barnhillら、Hum. Pathol.、30巻：513頁、1990年；Brochezら、J. Pathol.、196巻：459頁、2002年）。診断の過誤には、良性メラニン細胞の母斑と早期黒色腫との鑑別、および異型母斑と異形成母斑との鑑別の困難を含む、いくつかの根本的原因がある（例えば、Ruiterら、Sem. Cutaneous Med. Surg.、22巻：33頁、2003年を参照されたい）。

【0008】

黒色腫診断の誤りは、患者、患者の家族、および社会一般に対して著しく有害な影響を及ぼしている。黒色腫であると誤って診断された患者は、不適切、かつ、潜在的に危険な療法（複数可）を受ける場合があり、再発の恐怖を常に感じながら生涯を過ごす可能性があり、生活または健康の確保（insurance）を得ることができない可能性がある。他方、黒色腫ではなくて母斑であると誤って診断された患者は、それらの悪性度に応じた適切な療法が与えられず、患者の生涯は早期に断ち切られる可能性がある。最後に、この問題の社会的損害は、黒色腫の誤診が、米国では、がんベースの医療過誤賠償請求の最も一般的な理由としての乳がんの誤診に次いで2番目であるという事実により裏付けられている（McDonaldら、Internet J. Fam. Practice、7巻（2号）、2009年；Troxel、Am. J. Surg. Pathol.、27巻：1278頁、2000）。

【先行技術文献】

【非特許文献】

【0009】

【非特許文献1】Piepkornら、J. Am. Acad. Dermatol.、30巻：707頁、1994年

【非特許文献2】Farmerら、Hum. Pathol.、27巻：528頁、1996年

【非特許文献3】Coronaら、J. Clin. Oncol.、14巻：1218頁、1996年

【非特許文献4】Barnhillら、Hum. Pathol.、30巻：513頁、

10

20

30

40

50

1990年

【非特許文献5】Brochezら、J. Pathol. 196巻：459頁、2002年

【非特許文献6】Ruiterら、Sem. Cutaneous Med. Surg.、22巻：33頁、2003年

【非特許文献7】McDonaldら、Internet J. Fam. Pract. ice、7巻(2号)、2009年

【非特許文献8】Troxel、Am. J. Surg. Pathol.、27巻：1278頁、200

【発明の概要】

10

【発明が解決しようとする課題】

【0010】

組織病理学的検査単独の限界を踏まえると、医科学では、黒色腫の適正な診断のためのさらなるツールを有することが極めて重要である。特に、いずれの生検(例えば、異形成母斑または判定保留の(indeterminate)母斑)が、実際は、黒色腫と誤診されており、かつ/またはいずれの生検(例えば、母斑)が、黒色腫の分子的特徴または黒色腫への進行を裏付け得るのかを決定するツールが必要とされている。

【課題を解決するための手段】

【0011】

要旨

20

メラニン細胞含有試料を特徴付ける、例えば、試料が、良性の母斑であるのか、悪性黒色腫であるのかを決定するための方法が開示される。いくつかの例では、これらの方法は、(i)MAGEA2、PRAME、PDIA4、NR4A1、PDLIM7、B4GALT1、SAT1、RUNX1、SOCS3、および表13のバイオマーカーから選択されるバイオマーカーのうち少なくとも2つ;ならびに(ii)被験体から得られるメラニン細胞含有試料(母斑試料など)中の、少なくとも1つの正規化バイオマーカーに(複数可)についての発現レベル(核酸レベルまたはタンパク質レベルなど)を決定し、これにより、少なくとも2つのバイオマーカーの各々および少なくとも1つの正規化バイオマーカー(複数可)についての生の発現値をもたらしことにより、メラニン細胞含有試料を特徴付けるステップを含む。少なくとも2つのバイオマーカーの各々についての生の発現値を、少なくとも1つの正規化バイオマーカー(複数可)についての生の発現値に対して正規化して、少なくとも2つのバイオマーカーの各々についての正規化された発現値をもたらし。正規化された発現値を、回帰アルゴリズムまたは機械学習アルゴリズムにおいて使用して、出力値をもたらし。結果として得られる出力値を、良性であるのか悪性であるのかがあらかじめ既知である複数のメラニン細胞含有試料中の、少なくとも2つのバイオマーカーについての正規化された発現値から導出されうるカットオフ値と比較する。次いで、被験体から得られたメラニン細胞含有試料を、例えば、出力値が、複数の既知の良性試料と、カットオフ値の同じ側であれば、良性として特徴付けるか、または出力値が、複数の既知の悪性試料と、カットオフ値の同じ側であれば、悪性として特徴付ける。

30

【0012】

40

また、メラニン細胞含有試料中の悪性度を決定するための方法も提示される。このような方法は、被験体から得られるメラニン細胞含有試料中の、B4GALT1、BAX、MAGEA2、NR4A1、PDIA4、PRAME、RUNX1、SOCS3、SAT1、PDLIM7、BIRC5、MET、MAGEC2、POLR2J3、ZFYVE16、およびBEST1から選択される少なくとも2つのバイオマーカーについての発現レベル(核酸の発現レベルなど)を決定するステップを含みうる。方法はまた、少なくとも2つのバイオマーカーの発現レベルを入力として使用するアルゴリズムからの出力を計算するステップと、アルゴリズムの出力から、出力を既知の悪性のメラニン細胞含有試料に由来する参照標準と比較することにより、試料が悪性であるのか悪性でないのかを決定するステップを含みうる。方法は、少なくとも2つの選択されたバイオマーカーの発現レベル

50

を、表3のバイオマーカーのうちの少なくとも1つなどの、少なくとも1つの正規化バイオマーカーの発現レベルに対して正規化するステップもさらに含む。

【0013】

また、生物学的試料（メラニン細胞含有試料など）を、良性の母斑または原発性黒色腫として診断するためのアレイおよびキットも開示される。例えば、アレイは、少なくとも3つのアドレス可能な位置であって、各位置が、同じ特異性を伴う固定化された捕捉プローブを有し、各位置が、各他の位置における捕捉プローブとは異なる特異性を伴う捕捉プローブを有する位置を含むことが可能であり、ここで、少なくとも3つの位置のうちの2つにおける捕捉プローブは、MAGEA2、PRAME、PDIA4、NR4A1、PDLIM7、B4GALT1、SAT1、RUNX1、SOCS3、および表13のバイオマーカーのうちの2つ以上を含むバイオマーカーに、直接的または間接的に、特異的にハイブリダイズすることが可能であり、少なくとも3つの位置のうちの1つにおける捕捉プローブは、表3に列挙される正規化バイオマーカーに、直接的または間接的に、特異的にハイブリダイズすることが可能であり、各捕捉プローブの特異性は、アレイのアドレス可能な位置により同定可能である。本明細書で提示される1または複数のアレイのほか、溶解緩衝液を含有する容器；一本鎖核酸に特異的なヌクレアーゼを含有する容器；複数の核酸プログラミングリンカーを含有する容器；複数のNPPを含有する容器；複数の二官能性検出リンカーを含有する容器；二官能性検出リンカーに特異的に結合する検出プローブを含有する容器；および検出試薬を含有する容器のうちの1または複数も含むキットが提示される。

本発明の実施形態において、例えば以下の項目が提供される。

(項目1)

メラニン細胞含有試料を特徴付ける方法であって、

被験体から得られるメラニン細胞含有試料中の、

(i) MAGEA2、PRAME、PDIA4、NR4A1、PDLIM7、B4GALT1、SAT1、RUNX1、SOCS3、および表13のバイオマーカーから選択されるバイオマーカーのうちの少なくとも2つ、ならびに

(ii) 少なくとも1つの正規化バイオマーカー

についての発現レベルを決定し、これにより、該少なくとも2つのバイオマーカーの各々および該少なくとも1つの正規化バイオマーカーについての生の発現値をもたらすステップと、

該少なくとも2つのバイオマーカーの各々についての該生の発現値を、該少なくとも1つの正規化バイオマーカーについての該生の発現値に対して正規化して、該少なくとも2つのバイオマーカーの各々についての正規化された発現値をもたらすステップと、

該正規化された発現値を、回帰アルゴリズムまたは機械学習アルゴリズムにおいて使用して、出力値をもたらすステップと、

該出力値を、良性であるのか悪性であるのかがあらかじめ既知である複数のメラニン細胞含有試料中の、該少なくとも2つのバイオマーカーについての正規化された発現値から導出されたカットオフ値と比較するステップと、

該出力値が、該複数の既知の良性試料と、該カットオフ値の同じ側であれば、該試料を良性として特徴付けるか、または該出力値が、該複数の既知の悪性試料と、該カットオフ値の同じ側であれば、該試料を悪性として特徴付けるステップとを含む方法。

(項目2)

前記少なくとも1つの正規化バイオマーカーが、母斑試料と原発性黒色腫試料との間で発現の統計学的有意差を有さない、項目1に記載の方法。

(項目3)

前記少なくとも1つの正規化バイオマーカーが、表3のバイオマーカーのうちの1つ、2つ、3つ、4つ、5つ、6つ、7つ、8つまたは9つ全てを含む、項目1から2のいずれかに記載の方法。

10

20

30

40

50

(項目4)

前記メラニン細胞含有試料中の複数のバイオマーカーについての遺伝子発現値を測定するステップであって、該複数のバイオマーカーについての発現の範囲が、該試料のトランスクリプトームにおけるバイオマーカー発現の全範囲を表すステップと、

このような複数のバイオマーカーについての中心傾向発現値を計算するステップと、  
必要に応じて、外れ値を棄却し、該外れ値発現値を伴わずに再計算された複数の中心傾向発現値を計算するステップと

該中心傾向発現値、または、必要に応じて、該再計算された複数の中心傾向発現値を使用して、該少なくとも2つのバイオマーカーの各々についての前記生の発現値を正規化するステップと

をさらに含む、項目1または2に記載の方法。

10

(項目5)

試料中の遺伝子発現を決定する方法であって、

被験体からメラニン細胞含有試料を得るステップと、

該試料中の、MAGEA2、PRAME、PDIA4、NR4A1、PDLIM7、B4GALT1、SAT1、RUNX1、SOCS3、および表13のバイオマーカーから選択されるバイオマーカーのうちの少なくとも2つを含む複数の遺伝子の発現レベルを決定するステップと、

該試料中の該複数の遺伝子の発現レベルの報告、または該試料の、該複数の遺伝子の該発現レベルに基づく、母斑または黒色腫としての特徴付けを提供するステップと  
を含む方法。

20

(項目6)

前記少なくとも2つのバイオマーカーが、

N4RA1およびB4GALT1、または

NR4A1およびSOCS3、または

NR4A1、SOCS3、およびB4GALT1、または

MAGEA2、PRAME、PDIA4、NR4A1、PDLIM7、B4GALT1、SAT1、RUNX1、およびSOCS3

を含む、項目1から5のいずれかに記載の方法。

30

(項目7)

前記少なくとも2つのバイオマーカーが、表4のバイオマーカーのうちの少なくとも2つを含み、前記出力値が、ロジスティック回帰アルゴリズムによりもたらされた、項目1から6のいずれかに記載の方法。

(項目8)

前記アルゴリズムが、

出力値 =  $0 + 1 \times 1 + 2 \times 2 + \dots + n \times n$

[式中、 $X_n$ は、前記少なくとも2つの(最大で $n$ )表4のバイオマーカーについての発現値の対数であり、 $0$ は、 $-200$ を超え、かつ、 $200$ 未満であり、 $n > 0$ に対する全ての  $1$ は、 $-1$ 、 $000$ を超え、かつ、 $1$ 、 $000$ 未満である]

である、項目1から7のいずれかに記載の方法。

40

(項目9)

前記少なくとも2つのバイオマーカーが、表11および13のバイオマーカーのうちの少なくとも2つを含み、前記出力値が、機械学習アルゴリズムによりもたらされた、項目1から6のいずれかに記載の方法。

(項目10)

前記少なくとも2つのバイオマーカーが、表11のバイオマーカーのうちの少なくとも2つまたは表13のバイオマーカーのうちの少なくとも2つを含む、項目9に記載の方法

(項目11)

前記少なくとも2つのバイオマーカーが、

50

B 4 G A L T 1、B A X、M A G E A 2、N R 4 A 1、P D I A 4、P R A M E、R U N X 1、S O C S 3、S A T 1、P D L I M 7、B I R C 5、M E T、M A G E C 2、P O L R 2 J 3、Z F Y V E 1 6、およびB E S T 1のうちの少なくとも2つ、または  
N R 4 A 1、S O C S 3、P R A M E、P O L R 2 J 3、B E S T 1、R U N X 1、B I R C 5、M E T、P D L I M 7、Z F Y V E 1 6、H I F 1 A、およびP I C A L Mのうちの少なくとも2つ

を含み、

前記出力値が、機械学習アルゴリズムによりもたらされた、  
項目1から10のいずれかに記載の方法。

(項目12)

前記少なくとも2つのバイオマーカーが、表4、表11、表13、または表4および表11の両方の遺伝子のうちの少なくとも50%、少なくとも75%、少なくとも80%、少なくとも90%、少なくとも95%、または少なくとも98%である、項目16または17に記載の方法。

(項目13)

メラニン細胞含有試料における悪性度を決定する方法であって、  
被験体から得られるメラニン細胞含有試料中の、B 4 G A L T 1、B A X、M A G E A 2、N R 4 A 1、P D I A 4、P R A M E、R U N X 1、S O C S 3、S A T 1、P D L I M 7、B I R C 5、M E T、M A G E C 2、P O L R 2 J 3、Z F Y V E 1 6、およびB E S T 1から選択される少なくとも2つのバイオマーカーの発現レベルを決定するステップと、

該少なくとも2つのバイオマーカーの該発現レベルを入力として使用するアルゴリズムからの出力を計算するステップと、

該アルゴリズムの出力から、該出力を既知の悪性メラニン細胞含有試料に由来する参照標準と比較することにより、該試料が悪性であるのか悪性でないのかを決定するステップと

を含む方法。

(項目14)

前記少なくとも2つの選択されたバイオマーカーの前記発現レベルを、

(a) M F I 2、R A P 2 B、B M P 1、および/もしくはN C O R 2のうちの少なくとも1つ、

(b) M F I 2、N C O R 2、R A P 2 b、およびB M P - 1、

(c) R P S 6 K B 2および/もしくはS D H A、または

(d) 前記メラニン細胞含有試料中で発現した少なくとも1つの遺伝子であって、該少なくとも2つのバイオマーカーではなく、その発現が、代表的な複数のメラニン細胞含有試料において有意に異なる遺伝子

からなる群から選択される少なくとも1つの正規化バイオマーカーの該発現レベルに対して正規化するステップをさらに含む、項目13に記載の方法。

(項目15)

発現レベルを決定するステップが、核酸の発現を決定することを含む、項目1から14のいずれかに記載の方法。

(項目16)

核酸の発現を決定することが、前記試料を、複数の核酸プローブまたは対合増幅プライマー対と接触させることを含み、各プローブまたは対合プライマーは、該複数の核酸プローブまたは対合プライマーがその/それらの相補的な少なくとも2つのバイオマーカーとハイブリダイズすることを可能とする条件下で、前記少なくとも2つのバイオマーカーのうちの1つに対して特異的、かつ、相補的である、項目15に記載の方法。

(項目17)

前記試料を前記複数の核酸プローブと接触させた後で、該試料を、一本鎖核酸分子を消化するヌクレアーゼと接触させることをさらに含む、項目16に記載の方法。

10

20

30

40

50

(項目18)

少なくとも3つのアドレス可能な位置であって、各位置が、同じ特異性を有する固定化された捕捉プローブを含み、各位置が、各他の位置における捕捉プローブとは異なる特異性を有する捕捉プローブを含む位置を含むアレイであって、

該少なくとも3つの位置のうちの2つにおける該捕捉プローブが、MAGEA2、PRAME、PDLIM7、NR4A1、PDIA4、NR4A1、PDLIM7、B4GALT1、SAT1、RUNX1、SOCS3、および表13のバイオマーカーのうちの2つ以上を含むバイオマーカーに、直接的または間接的に、特異的にハイブリダイズすることが可能であり、該少なくとも3つの位置のうちの1つにおける該捕捉プローブが、表3に列挙される正規化バイオマーカーに、直接的または間接的に、特異的にハイブリダイズすることが可能であり、各捕捉プローブの該特異性が、該アレイの該アドレス可能な位置により同定可能であるアレイ。

10

(項目19)

前記少なくとも3つのアドレス可能な位置の各々が、個別に同定可能なビーズまたはフローセル内のチャンネルである、項目18に記載のアレイ。

(項目20)

項目18または19に記載のアレイ、ならびに  
溶解緩衝液を含有する容器；  
一本鎖核酸に特異的なヌクレアーゼを含有する容器；  
複数の核酸プログラミングリンカーを含有する容器；  
複数のNPPを含有する容器；  
複数の二官能性検出リンカーを含有する容器；  
該二官能性検出リンカーに特異的に結合する検出プローブを含有する容器；および  
検出試薬を含有する容器  
のうちの1または複数  
を含むキット。

20

【0014】

本開示の上記の特徴および他の特徴は、付属の図面を参照しながら進められる、いくつかの実施形態についての以下の詳細な記載からより明らかとなる。

30

【図面の簡単な説明】【0015】

【図1】図1は、本明細書で開示された診断テストの実施形態(「生検」のポイントから下方に向かう矢印から生じているフローチャートの要素(灰色で網掛けした部分)によって示された通り)が、現行の国立総合がん情報ネットワーク(Nation Comprehensive Cancer Network、NCCN)の黒色腫診断に関するどの臨床推奨に当てはまるかを示す流れ図である。

【0016】

【図2A】図2Aおよび2Bは、それぞれのグラフの上に示された代表的な正規化遺伝子(すなわち、MFI2、RAP2B、BMP1、およびNCOR2)についてのボックスプロット(上)、平均プロット(中央)、およびSASディフォグラム(diffogram)(下)を示す図である。総合すると、これらの結果から、各正規子(normalizer)遺伝子で母斑試料と原発性黒色腫試料との間に統計学的有意差はなかったこと、およびこのような各遺伝子は一貫した結果を低い標準偏差でもたらしたことが示される。

40

【図2B】図2Aおよび2Bは、それぞれのグラフの上に示された代表的な正規化遺伝子(すなわち、MFI2、RAP2B、BMP1、およびNCOR2)についてのボックスプロット(上)、平均プロット(中央)、およびSASディフォグラム(diffogram)(下)を示す図である。総合すると、これらの結果から、各正規子(normalizer)遺伝子で母斑試料と原発性黒色腫試料との間に統計学的有意差はなかったこと

50

、およびこのような各遺伝子は一貫した結果を低い標準偏差でもたらしたことが示される。

【0017】

【図3】図3は、代表的なB4GALT1およびNR4A1(4-正規子)モデルの統計的有意性を実証するSASの出力を示す図である。総合すると、この出力から、モデルは解決に収束していることから、モデルの結果は信頼できるものであったことが実証される。全体帰無仮説のモデル当てはめおよび検定から、全モデルが統計学的に有意であるか、または観察された結果が単なる偶然で起こり得る確率よりもかなり低い確率でしか起こらないと予想され、ワルドのカイ二乗 = 15.856、2df、 $p = 0.0004$ であることが示される。ホスマーおよびレメシヨウ検定で、そのモデルへの不適合度がない；またはそのモデルは正確にデータを再現するという帰無仮説を検定する。モデルの値をさらに支持するホスマーおよびレメシヨウ検定を使用しても有意性は見出されなかった。有意なホスマーおよびレメシヨウのp値(例えば、0.05未満)から、そのモデルに対してある程度の不適合度があること、または提唱されているモデルは、いくつかの能力において、実験データを適切に適合させることができなかったことが示唆されることに留意されたい。

10

【0018】

【図4】図4は、代表的なB4GALT1およびNR4A1(4-正規子)モデルについてのROC曲線を示すグラフである。ROC曲線は、そのモデルについての極めて高い感度および特異度を例示する。感度は、真陽性率(すなわち、ヒトが疾患を有する場合、どれだけ頻繁に検定が陽性になるか；または感度 = (真陽性 / (真陽性 + 偽陰性))を示す。特異度は、真陰性率(すなわち、ヒトが疾患を有さない場合、どれだけ頻繁に検定が陰性になるか；または特異度 = (真陰性 / (真陰性 + 偽陽性))を示す。曲線下面積(AUC = 0.9892)は、2つの集団、すなわち母斑と原発性黒色腫とを極めて高い精度で鑑別するモデルの能力を例示する。

20

【0019】

【図5】図5は、代表的なB4GALT1およびNR4A1(4-正規子)モデルについての異なる閾値におけるSAS交差検証後の分類結果を示す図である。確率レベルは、検定試料を原発性黒色腫とコールする確率である。試料を原発性黒色腫とコールするために閾値(カットオフ値)を上げることにより、モデルは、極めて高い特異度および優れた感度を得た。これらの結果からさらに、多様な閾値にわたりこのモデルを使用して極めて高い特異度および優れた感度が得られることが実証された。

30

【0020】

【図6】図6は、図5の分類表の続きを示す図である。これらの続きの結果から、カットオフ閾値を低くすることにより、特異度のわずかなトレードオフを伴いより高い感度が結果として得られたが、それでもなお極めて高い全体的な分類精度を維持したことが示される。

【0021】

【図7】図7は、代表的なB4GALT1およびNR4A1(4-正規子)モデルは、複数の異なる推定ルーチン下でも極めて有意であったことを示す図である。回帰ベースのモデルにおける一般的な仮説の1つは、等分散である。不等分散は、特に試料サイズが不均一である場合、標準的な推定の実行が不正確な結果を生じる原因となる可能性がある。等分散性に関するブラウン-フォーサイスの検定は集団分散間で有意な差を示さなかったが(示さず)、不等分散または一般的な仮説の何かしら他の妨害が存在し得る場合に使用される実験的な共分散の「サンドイッチ」推定量検定を行った。サンドイッチ推定量検定(左のボックス)により、標準的なフィッシャーのスコア付け法によって得られた元の結果は、モデル仮定の妨害に起因するのではないことが確認された。同様に、ファースのバイアス減少のペナルティ付き尤度モデル(右のボックス)から、結果は推定手法に対して感度が低かったことを追加で確認した。

40

【0022】

50

【図8】図8は、ワルドのカイ二乗で示されるようなB4GALT1およびNR4A1(2-正規子)モデル当てはめも極めて有意であったことを示す図である。ROC曲線から、このモデルも極めて高い感度および特異度を有していたことが実証される。B4GALT1およびNR4A1(2-正規子)ならびにB4GALT1およびNR4A1(4-正規子)モデルについての曲線下面積の極めて小さい変化(すなわち、 $\Delta = 0.0125$ )は、2つのモデルは、母斑試料と原発性黒色腫試料との間を正確に鑑別するそれらの能力に関して極めて類似していることを示す。

【0023】

【図9】図9は、B4GALT1およびNR4A1(2-正規子)モデルについての確率の分類表を示す表である。これらの結果から、このモデルは極めて高い感度および特異度を維持したことが実証される。B4GALT1およびNR4A1(4-正規子)モデルと比較して、2-正規子モデルの全体的な特異度は、モデルの範囲全体にわたりいくらか低下した。しかしながら、常に感度と特異度との間がトレードオフになっている。0.34およびそれ未満の閾値についての全体的な感度から、このモデルは、優れた特異度を維持しつつ適度に高い感度を提供したことが示された。試料を誤診することにより臨床上的意義がさらに悪くなることを考えれば、感度と引き換えに多少特異度を犠牲にすることは、許容できる結果である。B4GALT1およびNR4A1(2-正規子)モデルは、およそ閾値の50%で88.9%またはそれより優れた全体的に正しい分類を有した。

【0024】

【図10A】図10Aは、3つの散布図を示すグラフであり、それぞれx軸に列挙した各遺伝子(mRNA発現によって測定)に関する単変量の統計的検定(AUC(上)、倍率変化(fch;中央)、およびFDR調整p値(下))の結果を示す。各散布図の点線は、統計的有意性について選択されたカットオフを示す。AUCカットオフを超える(四角で囲む)、倍率変化カットオフより下(四角で囲む)、またはFDR調整p値カットオフより下の場合、その結果は有意とみなされる。各遺伝子を示す記号は、そのArray Plate(AP)上で発現データが測定されたことを示す。

【0025】

【図10B】図10Bは、示された各miRNA(x軸)に関する図10Aと類似の結果を示すグラフであるが、倍率変化のカットオフが陽性1(陰性1に対して)であり、そのラインを超えたら倍率変化の結果が有意とみなされることを除く。各miRNAについての発現値は正規化されている(+ )か、または正規化されていない(\* )。

【0026】

【図11】図11は、Array Plate番号3からの発現データで構築された例示的な2(左下)から40(右上)の遺伝子の母斑/黒色腫分類器の(AUCに基づく)分類の精度を示すグラフである。いずれの場合も、AUCは0.9に等しいかまたはそれよりも高く、これは、分類器中の遺伝子の数に関係なく優れた精度を示し、およそ分類器中の遺伝子が18種までは分類器の精度が高くなり、それ以降、AUCはおおよそ0.95で比較的安定であることを示す。

【0027】

【図12】図12は、4つの線グラフを合わせたものであり、それぞれ提示された通りのArray Plate番号3~6から収集された発現データに基づく2から40の遺伝子(x軸)のAUC、t検定、ランダムフォレスト、またはLIMMA分類モデルの誤分類率(y軸)を示す。

【発明を実施するための形態】

【0028】

配列

本明細書で列挙される核酸配列は、37 C.F.R.1.822で規定される、ヌクレオチド塩基のための標準的な文字による略記法を使用して示す。各核酸配列の一方の鎖だけを示すが、相補鎖も、表示される鎖に対する任意の参照により含められるものと理解される。

10

20

30

40

50

## 【0029】

配列表は、「Sequence.txt」（約371kb）と名付けられたファイルの形式のASCIIテキストファイルであって、2013年6月24日に作成され、参照により本明細書に組み込まれるASCIIテキストファイルとして提出されている。

提示される配列のうち、

## 【0030】

配列番号1~36、123、および124は、代表的なヌクレアーゼ保護プローブ（NPP:nuclease protection probe）配列である。

## 【0031】

配列番号47~119は、母斑および原発性黒色腫において差次的に発現するとおりの開示された遺伝子についてのGenBank mRNA RefSeqである。 10

## 【0032】

配列番号37~46、120、および121は、開示される正規子についてのGenBank mRNA RefSeqである。

## 【0033】

配列番号122は、開示される陰性対照の植物遺伝子（ANT）についてのGenBank mRNA RefSeqである。

## 【0034】

配列番号125~144は、開示されるmRNA標的についての代表的なNPP配列である。 20

## 【0035】

配列番号145~164は、開示されるmiRNA標的についての代表的なNPP配列である。

## 【0036】

詳細な説明

別段に言及しない限り、技術用語は、従来の用法に従い使用する。分子生物学における一般用語の定義は、Benjamin Lewin、Genes IX、Jones and Bartlett刊、2008年（ISBN 0763752223）；Kendrewら（編）、The Encyclopedia of Molecular Biology、Blackwell Science Ltd.刊、1994年（ISBN 0632021829）；およびRobert A. Meyers（編）、Molecular Biology and Biotechnology: a Comprehensive Desk Reference、VCH Publishers, Inc.刊、1995年（ISBN 9780471185710）において見出すことができる。 30

## 【0037】

単数形用語「ある(a)」、「ある(an)」、および「その」は、文脈により別段であることが明らかに指示されない限り、複数の指示対象を含む。同様に、「または」という語も、文脈により別段であることが明らかに指示されない限り、「および」を含むことを意図する。用語「~を含む(comprises)」とは、「~を含む(includes)」を意味する。利益相反の場合は、用語の説明を含む本明細書により管理する。 40

## 【0038】

下記では、本開示を実施するかまたは試験するのに適する方法および材料が記載される。このような方法および材料は、例示的なものであるに過ぎず、制限することを意図するものではない。本明細書で記載される方法および材料と同様または同等の他の方法および材料を使用することができる。例えば、開示される発明が関する技術分野において周知である従来の方法は、例えば、Sambrookら、Molecular Cloning: A Laboratory Manual、2版、Cold Spring Harbor Laboratory Press、1989年；Sambrookら、Molecular Cloning: A Laboratory Manual、3版、C 50

old Spring Harbor Press、2001年；Ausubelら、Current Protocols in Molecular Biology、Greene Publishing Associates、1992年（および2000年版への増補）；Ausubelら、Short Protocols in Molecular Biology: A Compendium of Methods from Current Protocols in Molecular Biology、4版、Wiley & Sons、1999年；HarlowおよびLane、Antibodies: A Laboratory Manual、Cold Spring Harbor Laboratory Press、1990年；ならびにHarlowおよびLane、Using Antibodies: A Laboratory Manual、Cold Spring Harbor Laboratory Press、1999年を含む、多様な、一般的参考文献およびより特定の参考文献において記載されている。加えて、材料、方法、および例は、例示的なものであるに過ぎず、制限することを意図するものではない。

10

**【0039】**

本明細書で言及される全てのGenbank番号は、2012年6月22日現在で利用可能な配列について、参照により組み込まれる。

**【0040】**

本開示の多様な実施形態の概説を容易とするため、以下の特定の用語の説明が提示される。

20

**【0041】**

抗体：抗原またはその断片のエピトープ、例えば、表3、4、11、または13に示されるバイオマーカーのエピトープを特異的に認識し、これに特異的に結合する少なくとも軽鎖免疫グロブリン可変領域または重鎖免疫グロブリン可変領域を含むポリペプチドリガンド。抗体という用語は、無傷の免疫グロブリン、ならびにFab'断片、F(ab)'<sub>2</sub>断片、単鎖Fvタンパク質（「scFv」）、およびジスルフィド結合により安定化されたFvタンパク質（「dsFv」）など、当技術分野において周知である、それらの変異体および部分を含む。この用語はまた、キメラ抗体、ヘテロコンジュゲート抗体（二重特異性抗体など）など、遺伝子操作された形態も含む。抗体という用語は、ポリクローナル抗体およびモノクローナル抗体の両方を含む。ポリクローナル抗体およびモノクローナル抗体、分子的に操作された抗体、および抗体断片の調製は、当業者に周知である。（例えば、Greenら、「Production of Polyclonal Antisera」、Immunochemical Protocols、1～5頁、Manson編、Humana Press 1992年；およびHarlowら、Antibodies: a Laboratory Manual、726頁、Cold Spring Harbor Pub.、1988年を参照されたい）。

30

**【0042】**

結合または安定的結合（オリゴヌクレオチドの）：十分な量のオリゴヌクレオチドが、その標的核酸と塩基対を形成するかまたはハイブリダイズする、例えば、プローブまたはプライマーなどのオリゴヌクレオチドが、表3、4、11、または13に示される遺伝子の核酸配列に結合する場合、オリゴヌクレオチドは、標的核酸（表3、4、11、または13に示されるバイオマーカーなど）に結合するかまたは安定的に結合する。標的とオリゴヌクレオチドとの結合は、機能的（例えば、発現および/または活性の低減）アッセイおよび物理的結合アッセイの両方を含む、当業者に公知の任意の手順により検出することができる。

40

**【0043】**

接触させること：直接物理的に会合させることであり、固体形態および/または液体形態における会合を含み、例えば、試料（例えば、緩衝液中に懸濁させた試料）を、表3、4、11、または13に示されるバイオマーカーのうちの1つに特異的なプローブなどの核酸プローブと接触させることである。接触させることは、in vitroにおいて、

50

例えば、診断アッセイにおいて、または、他の例では、*ex situ*においてなされる。

【0044】

検出するのに十分な条件：所望の活性を可能とする任意の環境、例えば、抗体が抗原（表3、4、11、または13に示されるバイオマーカーなど）に結合することを可能とし、相互作用が検出されることを可能とする任意の環境。他の例では、それは、例えば、バイオマーカーの、核酸プローブとのハイブリダイゼーションを検出することによる、表3、4、11、または13に示されるバイオマーカーなどの核酸の検出である。

【0045】

縮重変異体：対象のタンパク質（表3、4、または11に示されるバイオマーカーなど）をコードするポリヌクレオチドであって、遺伝子コードの結果として縮重である配列を含むポリヌクレオチド。それらの大半が複数のコドンにより指定される20の天然アミノ酸が存在する。したがって、ヌクレオチド配列によりコードされるポリペプチドのアミノ酸配列が変化しない限りにおいて、全ての縮重ヌクレオチド配列が含まれる。

10

【0046】

検出する：作用物質（シグナルもしくは特定の核酸、核酸プローブ、またはタンパク質、例えば、表3、4、11、または13のそれうちの1つなど）が存在するのか、存在しないのかを決定すること。いくつかの例では、これは、定量化、例えば、遺伝子もしくはタンパク質、または組織内の特定の1もしくは複数の細胞など、試料の画分の量の定量化をさらに含む。

20

【0047】

診断：黒色腫など、がんなどであるがこれらに限定されない病理学的状態の存在または性質を同定すること。診断法は、それらの感度および特異度が異なる。診断アッセイの「感度」とは、罹患した個体が検査で陽性とされる百分率（真陽性のパーセント）である。診断アッセイの「特異度」とは、1 - 偽陽性率 [ここで、偽陽性率は、疾患を有しないヒトが検査で陽性とされる比率と定義される] である。特定の診断法では、状態の決定的な診断を下せない場合もあるが、方法により、診断の一助となる情報（例えば、陽性指標）がもたらされれば十分である。

【0048】

ハイブリダイゼーション：オリゴヌクレオチドおよびそれらの類似体は、相補的な塩基の間のワトソン-クリック型水素結合、フーグスティーン型水素結合、または逆フーグスティーン型水素結合を含む水素結合によりハイブリダイズする。一般に、核酸は、ピリミジン（シトシン（C）、ウラシル（U）、およびチミン（T））またはプリン（アデニン（A）およびグアニン（G））である窒素含有塩基からなる。これらの窒素含有塩基は、ピリミジンとプリンとの間に水素結合を形成し、ピリミジンのプリンへの結合を、「塩基対合」と称する。より具体的には、Aは、TまたはUに水素結合し、Gは、Cに結合する。「相補的な」とは、2つの異なる核酸配列または同じ核酸配列の2つの異なる領域の間で生じる塩基対合を指す。例えば、オリゴヌクレオチドは、表3、4、11、または13の遺伝子のうちの1つによりコードされるmRNA、DNA、またはdsDNAと相補的でありうる。

30

40

【0049】

「特異的にハイブリダイズ可能な」および「特異的に相補的な」とは、オリゴヌクレオチド（またはその類似体）とDNA標的またはRNA標的との間で、安定的で特異的な結合が生じるのに十分な程度の相補性を指し示す用語である。オリゴヌクレオチドまたはオリゴヌクレオチド類似体は、特異的にハイブリダイズ可能であるために、その標的配列と100%相補的である必要はない。オリゴヌクレオチドまたはオリゴヌクレオチド類似体と標的のDNA分子またはRNA分子（例えば、表3、4、11、または13のDNAまたはRNA）との間に、オリゴヌクレオチドまたはオリゴヌクレオチド類似体の、非標的配列への非特異的結合を回避するのに十分な程度の相補性が存在する場合に、特異的

50

にハイブリダイズ可能である。このような結合を、特異的ハイブリダイゼーションと称する。

【0050】

特定の程度の厳密性 (stringency) を結果としてもたらずハイブリダイゼーション条件は、最適なハイブリダイゼーション法の性質およびハイブリダイズさせる核酸配列の組成および長さに応じて変化する。一般に、ハイブリダイゼーションの温度およびハイブリダイゼーション緩衝液のイオン強度 (とりわけ、 $\text{Na}^+$  濃度) が、ハイブリダイゼーションの厳密性を決定するが、費やされる時間もまた、厳密性に影響を与える。オリゴヌクレオチド配列のハイブリダイゼーションは、ロックド核酸またはペプチド核酸を組み込むことによりなど、非天然塩基を配列へと組み込むことにより改変することができる。

10

【0051】

単離：「単離」された生物学的成分 (核酸分子、タンパク質、または細胞小器官など) は、成分が自然発生する生物体の細胞内の他の生物学的成分、例えば、他の染色体の、および染色体外の DNA および RNA、タンパク質、ならびに / または細胞小器官から実質的に分離または精製されている。「単離」された核酸およびタンパク質は、標準的な精製法により精製された核酸およびタンパク質を含む。この用語はまた、宿主細胞内の組換え発現により調製される核酸およびタンパク質のほか、プローブおよびプライマー、例えば、表 3、4、11、または 13 に示される核酸を検出および / または増幅するためのプローブおよびプライマーなど、化学合成された核酸も包含する。

20

【0052】

標識：抗体 (例えば、表 3、4、11、または 13 に示されるバイオマーカー (例えば、タンパク質) に特異的に結合する抗体) または核酸プローブ (例えば、表 3、4、11、または 13 の核酸に特異的に結合するかまたは間接的に結合する核酸プローブ) またはタンパク質などの別の分子に、その分子の検出を容易とするために、直接的または間接的にコンジュゲートさせうる検出可能な化合物または組成物。標識ならびに核酸およびタンパク質に標識する方法の特定の非限定的な例は、本開示を通して記載される。

【0053】

黒色腫：メラニン細胞の悪性腫瘍。メラニン細胞とは、皮膚の色の一因となる暗色の色素であるメラニンを生成する細胞である。メラニン細胞は、皮膚において主に生じるが、また、腸および眼を含む身体の他の部分でも見出される。したがって、原発性黒色腫は、皮膚以外の身体の領域でも生じうる (例えば、ブドウ膜黒色腫)。原発性黒色腫とは、起源の部位における新生物であり、原発性腫瘍が転移した場合であっても、起源の部位は依然として原発性であり、遠隔の部位は転移である。

30

【0054】

母斑 (nevus) (複数形：母斑 (nevi))：皮膚における、または腸もしくは眼などの身体の他の一部における、限局性の色素性斑点。母斑は一般に、母斑 (birth mark) または母斑 (mole) と称する場合がある。母斑は、母斑の色素性外観に寄与するメラニン細胞を含む。典型的に、母斑は、良性と考えられる。しかし、異形成母斑 (場合によってまた、非定型の母斑とも称する) は、異常な特徴を伴う母斑の種類である。異形成母斑は、非異形成母斑より大きい場合があり、その色、表面、および境界部は、非異形成母斑と異なりうる。皮膚表面上では、異形成母斑は、数種の色の混合 (例えば、ピンク ~ 濃褐色)、平滑またはやや落屑性または石目状の (pebbly) 表面、および周囲の皮膚へと溶け込みうる不規則な縁辺部を有する母斑として現れうる。異形成母斑は、黒色腫へと発症する可能性が「尋常性」母斑より高く、黒色腫のうちの約半数は、異形成母斑から発生する。しかし、大半の異形成母斑は、悪性となることはなく、したがって、いずれの母斑 (異形成性であれ、非異形成性であれ) が、実際のところ、原発性黒色腫へと誤ってまたは生物学的に変換しつつある (例えば、分子レベルで) のかを決定することが重要である。

40

【0055】

50

ヌクレアーゼ：ホスホジエステル結合を切断する酵素。エンドヌクレアーゼとは、ヌクレオチド鎖内の内部のホスホジエステル結合を切断する（ヌクレオチド鎖の末端においてホスホジエステル結合を切断するエキソヌクレアーゼとは対照的に）酵素である。いくつかのヌクレアーゼは、エンドヌクレアーゼ活性およびエキソヌクレアーゼ活性の両方を有する。例示的なヌクレアーゼは、本開示を通して記載される。

【0056】

プライマー：DNAオリゴヌクレオチド、例えば、例えば、極高度な厳密性のハイブリダイゼーション条件下で、核酸ハイブリダイゼーションにより、相補的な標的核酸分子（表3、4、11、または13のバイオマーカーのうちの1つなど）とアニールして、プライマーと標的核酸鎖との間でハイブリッドを形成しうる、少なくとも15ヌクレオチドの配列などの短鎖核酸分子。

10

【0057】

プライマーは、ポリメラーゼ酵素により、標的核酸分子に沿って伸長しうる。したがって、プライマーを使用して、標的核酸分子（表3、4、11、または13に示される核酸分子の部分など）を増幅することができ、ここで、プライマーの配列は、例えば、プライマーが、極高度な厳密性のハイブリダイゼーション条件下で標的核酸分子とハイブリダイズするように、標的核酸分子に特異的である。

【0058】

プライマーの特異性は典型的に、その長さと共に増大する。したがって、例えば、30の連続ヌクレオチドを含むプライマーは、対応する15ヌクレオチドだけのプライマーより高度な特異性で標的配列とアニールする。したがって、より大きな特異性を得るためには、標的配列のうちの少なくとも15、20、25、30、35、40、45、50以上の連続ヌクレオチドを含むプローブおよびプライマーを選択することができる。

20

【0059】

特定の例では、プライマーは、標的核酸分子と相補的な少なくとも15連続ヌクレオチドなど、少なくとも10ヌクレオチドの長さである。本開示の方法を実施するのに使用される特定の長さのプライマー（例えば、表3、4、11、または13に示される核酸分子の領域を増幅するプライマー）は、10～60ヌクレオチド、10～50ヌクレオチド、または10～30ヌクレオチドのプライマーなど、増幅される標的核酸分子と相補的な少なくとも10、少なくとも11、少なくとも12、少なくとも13、少なくとも14、少なくとも15、少なくとも16、少なくとも17、少なくとも18、少なくとも19、少なくとも20、少なくとも21、少なくとも22、少なくとも23、少なくとも24、少なくとも25、少なくとも30、少なくとも35、少なくとも40、少なくとも45、少なくとも50以上の連続ヌクレオチドを有するプライマーを含む。

30

【0060】

例えば、当技術分野で公知であり、本開示の別の箇所で記載される、PCR、リアルタイムPCR、または他の核酸増幅法により核酸配列を増幅するために、プライマー対を使用することができる。「上流」または「フォワード」プライマーは、核酸配列上の参照点（reference point）に対して5'側のプライマーである。「下流」または「リバーズ」プライマーは、核酸配列上の参照点に対して3'側のプライマーである。

40

【0061】

プローブ：プローブは、標的核酸（表3、4、11、または13に示されるバイオマーカーの核酸配列など）とハイブリダイズすることが可能な単離核酸を含み、検出可能な標識またはレポーター分子を、核酸分子へと結合させることができる。例えば、標識は、プローブの5'端もしくは3'端、またはこれらの間の任意の位置において結合させることができる。具体例では、標識は、プローブの5'端における塩基、その3'端における塩基、その5'端におけるリン酸基、またはプローブに対して内部のTなどの修飾塩基へと結合させる。例示的な標識、標識するための方法、および多様な目的に適切な標識の選択における指針については、本開示の別の箇所で論じられる。

50

## 【0062】

プローブは一般に、20～500ヌクレオチド、100～250ヌクレオチド、20～50ヌクレオチド、または20～30ヌクレオチドなどの標的核酸分子(表3、4、11、または13の標的核酸分子など)と相補的な、少なくとも10、少なくとも15、少なくとも16、少なくとも17、少なくとも18、少なくとも19、少なくとも20、少なくとも21、少なくとも22、少なくとも23、少なくとも24、少なくとも25、少なくとも30、少なくとも35、少なくとも40、少なくとも45、少なくとも50、少なくとも55、少なくとも60、少なくとも70、少なくとも80、少なくとも90、少なくとも100、少なくとも120、少なくとも140、少なくとも160、少なくとも180、少なくとも200、少なくとも250、少なくとも300、少なくとも350、少なくとも400、少なくとも450、少なくとも500以上の連続ヌクレオチドなど、少なくとも15ヌクレオチドの長さである。

10

## 【0063】

配列同一性/配列類似性：2つ以上の核酸配列または2つ以上のアミノ酸配列の間の同一性/類似性は、配列の間の同一性または類似性との関係で表される。配列同一性は、同一性百分率との関係で測定することができ、百分率が大きいほど、配列の同一性は大きくなる。核酸配列またはアミノ酸配列の相同体またはオーソログは、標準的な方法を使用してアラインメントした場合、比較的高度の配列同一性/配列類似性を保有する。

## 【0064】

当技術分野では、比較のための配列アライメントの方法が周知であり、例えば、Altschulら、J. Mol. Biol.、215巻：403～10頁、1990年は、配列アライメント法および相同性の計算についての詳細な考察を提示している。NCBI Basic Local Alignment Search Tool (BLAST) (Altschulら、J. Mol. Biol.、215巻：403～10頁、1990年)は、National Center for Biological Information (NCBI、National Library of Medicine、Building 38A、Room 8N805、Bethesda、MD 20894)およびインターネット上を含む数種の供給源から入手可能である。

20

## 【0065】

本開示では、表4、11、または13に示される分子についての配列の相同体および変異体が包含され、典型的に、全長アライメントにわたりカウントされた、対象のアミノ酸配列または核酸配列との少なくとも約75%、例えば、少なくとも約80%、少なくとも85%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、または少なくとも99%の配列同一性を保有することを特徴とし、天然タンパク質または天然核酸の活性を保持しうる。当業者は、これらの配列同一性の範囲は、指針のためだけに提示されるものであり、提示される範囲の外側に外れる、非常に重要な相同体を得ることができることが完全に可能であることを十分に理解する。

30

## 【0066】

2つの核酸分子が近縁であることの1つの機能的な指標は、2つの分子が、ストリンジエント条件下で互いとハイブリダイズすることである。

40

## 【0067】

メラニン細胞含有試料を特徴付けるための方法および組成物

黒色腫を含む大半のがんでは、早期の検出が、生存に対して最大の影響を及ぼし、治療率の向上に寄与しうる。いくつかの場合には、古典的な方法(例えば、組織病理学的検査)だけに基いて良性病変と悪性病変とを鑑別することが困難である。したがって、良性の母斑を、黒色腫(例えば、原発性黒色腫)から鑑別することを可能とする方法が必要とされる。検査方法を発展させることは、例えば、このような悪性腫瘍が、顕微鏡レベルまたは生物体レベルにおいて信頼できる形で認識される前に、悪性腫瘍を分子レベルで同定する一助となりうる。分子的検査は、がんの表現型を、良性の母斑を悪性黒色腫(例えば

50

、原発性黒色腫）から鑑別するために本明細書で記載される、臨床的に関与性の遺伝子発現パターンに対して同定することを包含する。このような弁別により、良性の母斑だけを有する者に対する不要な治療を回避することができ、原発性黒色腫を有する者に、初期生検の後で適切な治療を施すことを確実にする一助とすることができる。

#### 【0068】

遺伝子発現データを収集するための準備

遺伝子発現とは、ゲノム（遺伝子）内でコードされた情報が、対応する遺伝子産物（例えば、RNA（mRNAおよびmiRNAなど）およびタンパク質）であって、一群の特徴（表現型としてもまた公知である）をもたらすように相互関連的に機能する遺伝子産物へと変換される（例えば、転写過程および翻訳過程を介して）過程である。本開示の目的では、遺伝子発現は、現在公知であるかまたは将来的に公知となる任意の技法により測定することができる。一般に、遺伝子発現は、目的とする被験体から回収された試料中で発現した遺伝子産物（例えば、mRNA、miRNA、および/またはタンパク質）を検出することにより測定する。

10

#### 【0069】

被験体および試料

本明細書で開示される方法における使用に適切な試料は、それらの遺伝子発現（例えば、mRNA発現、miRNA発現、またはタンパク質発現；表3、4、11、および/または13のmRNA発現、miRNA発現、またはタンパク質発現など）についての情報が所望される、メラニン細胞を含有する任意の従来の生物学的試料を含む。

20

#### 【0070】

試料は、被験体から得られる臨床試料（健常であるかまたは見かけ上健常なヒト被験体、または黒色腫などの、診断されるかまたは調査される状態または疾患に罹患するヒト患者に由来する試料を含む）など、被験体から得られる試料を含む。被験体とは、例えば、哺乳動物を含む範疇である、生存する多細胞脊椎動物の生物である。「哺乳動物」は、ヒト、およびイヌ、マウス、または他の獣医科被験体など、非ヒト哺乳動物の両方を含む。一例では、試料は、既往の黒色腫病歴を有さない被験体に由来するか、または黒色腫をかつて有したか、もしくは黒色腫を有するとかつて診断されたことがある被験体に由来する。いくつかの例では、被験体は、例えば、下記に記載される通り、皮膚がん（例えば、黒色腫）スクリーニング参加しているか、または黒色腫を有するかもしくは黒色腫を発症する危険性（または高い危険性）があると診断された患者などの患者である。いくつかの例では、試料は、既往の黒色腫病歴を有さない被験体に由来するか、または黒色腫を有するとかつて診断された被験体に由来する。

30

#### 【0071】

ヒトにおける最高の黒色腫率は、オーストラリアで報告されている（ニュージーランド、ノルウェー、スウェーデン、スイス、デンマーク、米国、オーストリア、アイスランド、オランダがこれに続く）。ヒト被験体が黒色腫を発症する危険因子は、（a）黒色腫の家族歴または患者病歴；（b）複数の母斑（例えば、50または100個を超える母斑）、（c）複数の異形成母斑（例えば、少なくとも3つ）、（d）日光に対する高い曝露（例えば、10歳以前における）、（e）蒼白色の白人種の皮膚、（f）赤毛または金髪、（g）少なくとも1カ所の水疱形成性日焼けの履歴、（h）高い社会経済階層、（i）サンベッド使用の履歴（とりわけ、30歳以前における）、（j）職業が航空機の乗務員であること、および（k）殺虫剤への曝露を含む（MacKieら、Annals of Oncology、20巻（増刊6号）、vi1~7頁、2009年）。

40

#### 【0072】

いくつかの例では、かつて使用された方法は、メラニン細胞含有試料が悪性であるのか、良性であるのかを信頼できる形で決定することが可能でなかった。したがって、本開示される方法は、別の方法により、例えば、組織病理学的検査により、悪性または良性と信頼できる形で診断できない、分析すべき試料を使用するステップおよび/または決定するステップを含みうる。このような任意選択のステップは、試料中の遺伝子発現レベルのレ

50

ベル（例えば、表4、11、および/または13の少なくとも2つの異なるバイオマーカー（表6、8、または14の遺伝子の組合せなど）、および/または少なくとも1つの正規化バイオマーカーの遺伝子発現）を決定する前に行うことができる。

#### 【0073】

例示的な試料は、限定なしに述べると、細胞、細胞溶解物、細胞遠心分離調製物、細胞学スミア、組織生検（例えば、母斑を含む皮膚生検などの皮膚生検または眼組織生検）、細針吸引物、および/または組織切片（例えば、クリオスタット組織切片および/またはパラフィン包埋組織切片）を含む。組織とは、複数の機能的に関連する細胞である。特定の例では、組織は、懸濁させた場合もあり、無傷の場合もある。一例では、メラニン細胞含有試料（組織試料など）は、母斑、異形成母斑、異型母斑、または疑黒色腫を含む。特定の例では、試料を直接使用する（例えば、新鮮試料または凍結試料）か、または、例えば、固定（例えば、ホルマリンを使用する）および/もしくは蠟中に包埋すること（ホルマリン固定パラフィン包埋（FFPE: formalin-fixed paraffin-embedded）組織試料など）により使用前に操作することができる。したがって、いくつかの例では、分析されるメラニン細胞含有試料を固定する。他の方法の実施形態は、試料（例えば、皮膚生検）を固定剤（例えば、ホルマリン）中で固定するステップと、試料を包埋する（例えば、パラフィンにより）ステップと、試料を細断するかもしくは切片化するステップまたはこれらの組合せを含む。

10

#### 【0074】

本開示において有用な、試料を収集するための標準的な技法が利用可能である（例えば、Schlugerら、J. Exp. Med.、176巻：1327～33頁（1992年）；Bigbyら、Am. Rev. Respir. Dis.、133巻：515～18頁（1986年）；Kovacsら、NEJM、318巻：589～93頁（1988年）；およびOgnibeneら、Am. Rev. Respir. Dis.、129巻：929～32頁（1984年）を参照されたい）。いくつかの例では、試料は、切除生検、切開生検、パンチ生検、杯形成生検（saucerization biopsy）、または細針吸引生検により得られる皮膚試料または眼組織である。切除生検では、成長部分の全体を、正常な周囲の皮膚または眼組織の縁辺部と共に切除するかまたは切り出す。一般に、生検が陽性である場合は、さらに広範な、正常な周囲の皮膚の局所的切除も要請されるであろう。縁辺部の幅は、がんの厚みに依存する。切開生検またはコア生検では、成長部分の試料だけを摘出する。パンチ生検では、皮膚または眼組織の小型で円筒形の試料を摘出する。パンチ生検は、表皮、真皮、および下層組織の一部を含む。杯形成生検では、病変の下方を「スコップ様」様式で切り出すことにより、病変の全体を摘出し、生検従事者に、腫瘍の構造をより十分に解析するための完全な検体をもたらす。細針吸引生検は、極細型注射針およびシリンジによりなされる。細針吸引生検では、極少量の組織試料を摘出する。この種類の生検は、疑わしい母斑または皮膚もしくは眼内の成長部分に対して行うことができる。加えて、この種類の生検は、黒色腫が広がっているのかどうかを確かめるために、近傍のリンパ節または内臓など、他の深部組織に対しても行うことができる。被験体から組織を得る任意の方法を活用することができ、使用される方法の選択は、組織の種類、被験体の年齢、または生検従事者に利用可能な手順などの多様な因子に依存することが十分理解される。

20

30

40

#### 【0075】

いくつかの実施形態では、メラニン細胞を含有する試料は、細胞溶解物および/または組織溶解物である。細胞溶解物は、細胞内に含有されるタンパク質および核酸のうちの多くを含有し、例えば、表3、4、11、または13に示されるバイオマーカーを含む。当技術分野では、細胞溶解物を得るための方法が周知であり、例えば、Ausubelら（Current Protocols in Molecular Biology、John Wiley & Sons、New York、1998年）において見出すことができる。いくつかの例では、試料中の細胞を、水溶液中で（例えば、溶解緩衝液を使用して）溶解させるかまたは透過処理する。水溶液または溶解緩衝液は、洗浄剤（ドデシ

50

ル硫酸ナトリウムなど)、および1または複数のカオトロピック剤(ホルムアミド、グアニジンHCl、グアニジウムイソチオシアネート、または尿素など)を含む。溶液はまた、緩衝液(例えば、SSC)も含有しうる。いくつかの例では、溶解緩衝液は、約8%~60%のホルムアミド(v/v)、約0.01%~0.5%のSDS、および約0.5~6倍濃度のSSC(例えば、約3倍濃度のSSC)を含む。緩衝液は必要に応じて、約0.001~約2.0mg/mlのtRNAまたはリボヌクレアーゼも含みうる。溶解緩衝液はまた、Phenol RedなどのpH指示薬も含みうる。細胞を、水溶液中で、細胞を溶解させるかまたは透過処理するのに十分な期間(約1分間~約60分間、例えば、約5分間~約20分間、または約10分間など)にわたり、十分な温度(約22~約115、例えば、約37~約105、または約90~約100など)でインキュベートする。いくつかの例では、例えば、検出される核酸がRNAであれば、溶解は、約95で実施する。他の例では、例えば、検出される核酸がDNAであれば、溶解は、約105で実施する。いくつかの例では、溶解条件は、ゲノムDNAはプローブにアクセス可能でないのに対し、RNA(例えば、mRNA)はプローブにアクセス可能であるか、またはRNAは破壊され、DNAだけがプローブのハイブリダイゼーションにアクセス可能であるような溶解条件でありうる。いくつかの例では、さらなる精製を伴わない、粗細胞溶解物を直接使用する。

#### 【0076】

##### 参照標準

参照標準はまた、「対照」とも称する場合がある。対照は、組織内もしくは細胞内またはそれらの集団内(正常非がん性皮膚組織内または細胞内など)に存在する発現(表4、11、または13に示されるバイオマーカーの発現など)の基礎レベルまたは量を示す既知の値または値の範囲でありうる。対照はまた、細胞対照または組織対照でもありうる。

#### 【0077】

対照試料は、表3に示される正規化マーカーなど、それに対して表4、11、または13に示される黒色腫バイオマーカーの発現を比較する、任意の適切な試料(例えば、細胞対照試料、組織対照試料、または器官対照試料)を含む。いくつかの実施形態では、対照試料は、複数の非腫瘍組織試料などの非腫瘍組織である。一例では、非腫瘍組織は、良性の母斑など、良性であることが既知の組織である。いくつかの例では、非腫瘍組織は、正常の外見を呈する、すなわち、母斑、良性病変、または黒色腫が存在しない皮膚試料を含む。いくつかの例では、非腫瘍組織は、悪性黒色腫に隣接するかなまたは悪性黒色腫から遠隔の非腫瘍組織など、同じ被験体から得られる。他の例では、非腫瘍組織は、健全な対照被験体またはいくらかの健全な対照被験体から得られる。例えば、非腫瘍組織は、複数の健全な対照被験体(例えば、複数のこのような被験体に由来する良性の母斑を含有する試料など、黒色腫を含む任意のがんを有さない対照被験体)から得ることができる。

#### 【0078】

いくつかの実施形態では、対照試料は、黒色腫(例えば、原発性黒色腫)試料のトレーニングセットなど、複数の既知の黒色腫試料など、既知の腫瘍組織である。他の実施形態は、母斑試料のトレーニングセットなど、良性母斑であることが既知である組織の対照を包含する。いくつかの実施形態では、試料(例えば、母斑および黒色腫)のトレーニングセットは、このような試料型の間を識別するアルゴリズム(例えば、機械学習アルゴリズム)を開発または「トレーニングする」のに有用である。

#### 【0079】

被験試料と対照との間の差違は、増大の場合もあり、逆に減少の場合もあり、例えば、表4、11、または13に示されるバイオマーカーの発現の減少または増大でありうる。差違は、定性的差違の場合もあり、定量的差違の場合もあり、例えば、統計学的有意差でありうる。いくつかの例では、差違は、対照と比べた、少なくとも約10%、少なくとも約20%、少なくとも約30%、少なくとも約40%、少なくとも約50%、少なくとも約60%、少なくとも約70%、少なくとも約80%、少なくとも約90%、少なくとも約100%、少なくとも約150%、少なくとも約200%、少なくとも約250%、少

10

20

30

40

50

なくとも約300%、少なくとも約350%、少なくとも約400%、少なくとも約500%、または500%を超えるなど、少なくとも約1%の量の増大または減少である。いくつかの実施形態では、対照は、原発性黒色腫（複数可）または良性の母斑（複数可）であることが既知である試料（複数可）の表4、11、または13に示されるバイオマーカーについて予測される発現レベルなど、参照値または値の範囲である。他の実施形態では、対照試料から得られる参照値は、集団の中心傾向（central tendency）（「CT」）（平均（例えば、算術平均または幾何平均）、中央値、最頻値、または平均値）、または集団のCTを0.5、1.0、1.5、または2.0標準偏差上回る範囲および/もしくはこれを下回る範囲など、値の参照範囲でありうる。例えば、1またはそれより多い参照値は、健全な対照の被験体群から（例えば、複数の既知の良性の母斑から）得られる平均発現値から導出することもでき、黒色腫（例えば、複数の既知の悪性母斑から）を有するがん患者の群から得られる平均発現値から導出することもできる。

10

#### 【0080】

##### 試料分析の選択肢

特定の例では、メラニン細胞含有試料（例えば、皮膚生検）など、分析される試料は、固定されるかまたは固定されている。固定技法は、施設間、国の間、研究者間などにより異なる可能性があり（Dissecting the Molecular Anatomy of Tissue、Emmert-Buck、Gillespie、およびChuaqui編、New York、Springer-Verlag、244頁（2010年））、検出される遺伝子産物（複数可）の完全性および/または検出される遺伝子産物（複数可）へのアクセス可能性に影響を及ぼしうる。したがって、固定された試料を伴う、いくつかの開示される方法（例えば、PCRまたは核酸シーケンシングなど、遺伝子発現産物（複数可）を単離するためのステップを伴う方法の実施形態）では、RNAの回収（例えば、可逆性架橋形成剤、エタノールベースの固定剤、および/またはRNAの抽出もしくは精製を使用する（全体的にまたは部分的に））が有利な場合がある。とりわけ、他の代表的な方法（例えば、qNPAを含む）では、RNAの回収は、任意選択であるか、またはRNAの回収は、明確に不要である。同様に、いくつかの方法の実施形態では、組織のコンディショニングを使用して、タンパク質の遺伝子産物を固定された組織から回収し、これにより、このようなタンパク質産物の検出の一助とすることができる。

20

#### 【0081】

生物学的試料中の腫瘍または疑腫瘍（例えば、黒色腫）の百分率は、変化しうる。したがって、いくつかの開示される実施形態では、試料面積（または試料体積）または試料中の全細胞のうちの少なくとも5%、少なくとも10%、少なくとも25%、少なくとも50%、少なくとも75%、少なくとも80%、または少なくとも90%は、腫瘍または疑腫瘍（例えば、黒色腫）である。他の例では、例えば、異常（例えば、異形成）であるかまたは異常（例えば、異形成）であると考えられる領域または細胞を試料からマクロダイセクトすること（macrodissection）により、試料を、腫瘍（または疑腫瘍）細胞について濃縮することができる。必要に応じて、病理学者または他の適切に訓練された従事者は、試料（例えば、H&E染色組織切片）を精査して、マクロダイセクトされる領域を調べ、かつ/または印をつけるのに十分な異常性（例えば、疑腫瘍）が、試料中に存在するかどうかを決定することができる。具体例では、被験試料のマクロダイセクションは、可能な限り多くの壊死性領域および/または出血性領域を回避する。いくつかの開示される方法において有用な試料が有する壊死は、試料体積もしくは試料面積または全細胞の25%、15%、10%、5%、2%、または1%未満である。

30

40

#### 【0082】

試料投入量は、検出のために利用可能な遺伝子産物（例えば、表3、4、11、または13のバイオマーカーのうちの1または複数）の量および/または濃度に影響を与える。特定の実施形態では、少なくとも1ng、10ng、100ng、1ug、10ug、100ug、500ug、1mgの全RNA（例えば、mRNAまたはmiRNA）、少なくとも1ng、10ng、100ng、1ug、10ug、100ug、500ug、1

50

mgの全DNA、または少なくとも0.01ng、0.1ng、1ng、10ng、100ng、1ug、10ug、100ug、500ug、または1mgの全タンパク質が、試料（試料溶解物など）から単離され、かつ/または試料中に存在する。いくつかの実施形態では、少なくとも3、5、8、または10 $\mu$ m（例えば、約3～約10 $\mu$ m）の厚さおよび/または少なくとも0.15、0.2、0.5、1、1.5、2、5、または10cm<sup>2</sup>の面積である組織試料（例えば、FFPEにより切片化された皮膚生検）を使用する。いくつかの方法の実施形態では、緩衝液中に懸濁させる試料の濃度は、少なくとも0.006cm<sup>2</sup>/ul（例えば、25ulの緩衝液（例えば、溶解緩衝液）当たり0.15cm<sup>2</sup>のFFPE組織）である。

#### 【0083】

10

#### 遺伝子および遺伝子セット

本明細書で開示される技術革新には、それらの発現（例えば、mRNA発現、miRNA発現、またはタンパク質発現により測定される）が、良性メラニン細胞含有試料（例えば、母斑）と悪性メラニン細胞含有試料（例えば、原発性黒色腫）とを識別するための、開示される方法、アレイ、およびキットにおいて有用な遺伝子（また、バイオマーカーとも称する）および遺伝子のセットがある。また、母斑試料および黒色腫（例えば、原発性黒色腫）試料のための正規子（例えば、試料-試料間の対照）として有用な遺伝子および遺伝子セットも開示される。

#### 【0084】

いくつかの例では、例えば、少なくとも1つの正規化マーカー（表3の正規化マーカーのうちの1または複数など）に対して正規化された、表4、11、および/または13のうちのいずれかまたは全てからの少なくとも2つの異なるバイオマーカー（限定なしに述べると、表6、8、または14の遺伝子の組合せを含む）の発現の変化（上方調節または下方調節など）を、母斑または黒色腫の特異的マーカーとして使用することもでき、良性の母斑と原発性黒色腫との間の移行についてのマーカーとして使用することもできる。このようなマーカーは、本開示においてより詳細に記載される様々な方法および組成物に有用であり、例えば、表4、11、および/または13のうちのいずれかまたは全てからの2つ以上の異なるバイオマーカー（例えば、表6、8、または14の遺伝子の組合せを含む）の発現レベルを測定または検出することにより、ヒト被験体などの被験体を、良性の母斑を有する、または黒色腫を有すると診断するための方法を含む。一例では、ヒト被験

20

30

#### 【0085】

本開示では、対象のメラニン細胞含有試料（集団）（例えば、黒色腫試料と対比した母斑試料）中で有意に差次的に発現する（SDE）遺伝子を同定し、同定されたSDE遺伝子の例示的な組合せを解析して、良性の母斑または原発性黒色腫としてのメラニン細胞含有試料の特徴付けを可能とする予測値を有する、SDE遺伝子の組合せを同定した（例えば、実施例2、3、および4を参照されたい）。本明細書では、同定されたSDE遺伝子の特定の組合せについて記載するが、当業者は、本開示が今や、表4、11、および/または13に示されるSDE遺伝子の他の組合せであって、母斑または黒色腫としての試料を頑健に特徴付ける他の組合せの同定も可能とすることを十分に理解する。例えば、表4、11、および/または13のうちのいずれかまたは全てにおける、バイオマーカーの任意の重複しない組合せであって、全ての予測子の変数X<sub>n</sub>（選択されたバイオマーカーの発現値）の分散拡大係数（VIF）が10未満である組合せは、良性の母斑に由来する試料を、原発性黒色腫に由来する試料と対比して鑑別するのに有用な予測値を有することが期待され、したがって、本開示で想定される。加えて、本明細書で開示される方法（例えば、AUC）または当技術分野で一般に公知の方法のうちのいずれかを使用して、表4、11、および/または13の遺伝子の任意の組合せによる母斑-黒色腫間の分類子を、許容可能な分類性能（例えば、誤分類が試料のうちの1%、2%、3%、4%、5%、6%、7%、8%、または10%未満であるか、または分類精度が75%、80%、85%、90%、92%、93%、94%、95%、96%、97%、98%、または99%以

40

50

上である)について調べることができる。

【0086】

本開示を通して記載される特定の方法の実施形態は、被験体から得られる試料(例えば、皮膚試料)中で、下記の(a)~(r)のうちの任意の1または複数から選択される少なくとも2つの異なる(すなわち、重複しない)バイオマーカーの発現レベル(核酸レベルまたはタンパク質レベルなど)を決定し、いくつかの場合には、少なくとも1つの正規化バイオマーカー(表3において列挙されるものなど)の発現レベルを決定するステップを含む。同様に、本開示を通して記載される特定の組成物実施形態は、下記の(a)~(r)のうちの任意の1または複数から選択される少なくとも2つの異なる(すなわち、重複しない)バイオマーカーの発現レベル(核酸レベルまたはタンパク質レベルなど)を特異的に測定し、いくつかの場合には、少なくとも1つの正規化バイオマーカー(表3において列挙されるものなど)の発現レベルを特異的に測定するのに使用されうる、特異的結合剤(例えば、プローブ、プライマー、アプタマー、抗体など)を含みうる。いくつかの例では、当てはまる場合、(a)~(r)のうちのいずれか1つにおいて列挙されるバイオマーカーのうちの少なくとも3つ、少なくとも4つ、少なくとも5つ、少なくとも6つ、少なくとも7つ、少なくとも8つ、少なくとも9つ、少なくとも10、少なくとも11、少なくとも12、少なくとも13、少なくとも14、少なくとも15、少なくとも16、少なくとも17、少なくとも18、少なくとも19、少なくとも20、少なくとも21、少なくとも22、少なくとも23、少なくとも24、少なくとも25、少なくとも26、少なくとも27、少なくとも28、少なくとも29、少なくとも30、少なくとも31、または全て(列挙されるバイオマーカーのうちの2つ~20、2つ~10、4つ~10、4つ~15、または2つ~5つなど)についての発現レベル(核酸レベルまたはタンパク質レベルなど)を試料中で決定するか、または開示される組成物(例えば、アレイまたはキット)を使用して特異的に検出することができる。他の例では、下記の(a)~(r)のうちの任意の1または複数から選択される少なくとも2つの異なる(すなわち、重複しない)バイオマーカーについての発現レベル(核酸レベルまたはタンパク質レベルなど)は、バイオマーカーが選択される特定の群(例えば、表4、11、および/または13)内で列挙される複数の遺伝子の少なくとも50%、少なくとも75%、少なくとも80%、少なくとも90%、少なくとも95%、または少なくとも98%である。

(a)表4において記載される遺伝子(すなわち、NR4A1、B4GALT1、SAT1、TP53、TADA3、BRAF、TFRC、RUNX1、SOCS3、PDLIM7、SP100、PIP4K2A、SOX4、PDIA4、MCM6、CTNNB1、RPL37A、GNAS、TGFB1、PIIA、PTEN、MAGED2、1PRAME、GALNTL1、MAGEA2、TEX13A、CREBBP、TPSAB1、CDK2、STAT2、SQSTM1、およびB2M);ならびに/または

(b)表11において記載される遺伝子(すなわち、B4GALT1、BAX、MAGEA2、NR4A1、PDIA4、PRAME、RUNX1、SOCS3、SAT1、PDLIM7、BIRC5、HIF1A、MET、MAGEC2、ERCC1、POLR2J3、LDHA、PICALM、ZFYVE16、およびBEST1);ならびに/または

(c)表13において記載される遺伝子(すなわち、産物hsa.mir.122、hsa.mir.1291、hsa.mir.191、hsa.mir.19b、hsa.mir.200a、hsa.mir.200c、hsa.mir.203、hsa.mir.205、hsa.mir.21、hsa.mir.23b、hsa.mir.29c、hsa.mir.342.3p、hsa.mir.375、hsa.mir.665、hsa.mir.1304、hsa.mir.142.5p、hsa.mir.1254、hsa.let.7a、hsa.mir.140.5p、およびhsa.mir.183)を発現させる遺伝子);ならびに/または

(d)NR4A1、B4GALT1、SOX4、SQSTM1、B2M、TFRC、TP53、GALNTL1、CREBBP、SOCS3、およびCTNNB1;ならびに/または

10

20

30

40

50

(e) NR4A1、B4GALT1、SOX4、SQSTM1、B2M、TFRC、TP53、CREBBP、SOCS3、RPL37A、SAT1、BRAF、およびTPSAB1；ならびに/または

(f) NR4A1、B4GALT1、SOX4、SQSTM1、B2M、TFRC、TP53、CREBBP、およびSOCS3；ならびに/または

(g) NR4A1、B4GALT1、SOX4、SQSTM1、B2M、TFRC、TP53、SOCS3、およびBRAF；ならびに/または

(h) NR4A1、B4GALT1、SOX4、SQSTM1、B2M、TFRC、TP53、CREBBP、SOCS3、およびBFAF；ならびに/または

(i) MAGEA2、PRAME、PDIA4、NR4A1、PDLIM7、B4GALT1、SAT1、RUNX1、およびSOCS3；ならびに/または

(j) 表6において記載される任意の遺伝子セット；ならびに/または

(k) 表8において記載される任意の遺伝子セット；ならびに/または

(l) 表14において記載される任意の遺伝子セット；ならびに/または

(m) 下記：

[NR4A1、B4GALT1]、[NR4A1、SOX4]、[NR4A1、SQSTM1]、[NR4A1、B2M]、[NR4A1、TFRC]、[NR4A1、TP53]、[NR4A1、CREBBP]、[NR4A1、SOCS3]、[NR4A1、BRAF]、[B4GALT1、SOX4]、[B4GALT1、SQSTM1]、[B4GALT1、B2M]、[B4GALT1、TFRC]、[B4GALT1、TP53]、[B4GALT1、CREBBP]、[B4GALT1、SOCS3]、[B4GALT1、BRAF]、[SOX4、SQSTM1]、[SOX4、B2M]、[SOX4、TFRC]、[SOX4、TP53]、[SOX4、CREBBP]、[SOX4、SOCS3]、[SOX4、BRAF]、[SQSTM1、B2M]、[SQSTM1、TFRC]、[SQSTM1、TP53]、[SQSTM1、CREBBP]、[SQSTM1、SOCS3]、[SQSTM1、BRAF]、[B2M、TFRC]、[B2M、TP53]、[B2M、CREBBP]、[B2M、SOCS3]、[B2M、BRAF]、[TFRC、TP53]、[TFRC、CREBBP]、[TFRC、SOCS3]、[TFRC、BRAF]、[TP53、CREBBP]、[TP53、SOCS3]、[TP53、BRAF]、[CREBBP、SOCS3]、[CREBBP、BRAF]、および[SOCS3、BRAF]の角カッコ([ . . . ])内で対をなす特異的組合せのうちのいずれか；ならびに/または

(n) NR4A1、B4GALT1、SOX4、SQSTM1、B2M、TFRC、TP53、CREBBP、SOCS3、およびBRAFのリストに由来する1つ(または2つ)の他の非重複遺伝子と組み合わされた、(m)内の対のうちのいずれかにより記載される3つ(または4つ)の組合せ；ならびに/または

(o) 下記：

[MAGEA2、PRAME]、[MAGEA2、PDIA4]、[MAGEA2、NR4A1]、[MAGEA2、PDLIM7]、[MAGEA2、B4GALT1]、[MAGEA2、SAT1]、[MAGEA2、RUNX1]、[MAGEA2、SOCS3]、[PRAME、PDIA4]、[PRAME、NR4A1]、[PRAME、PDLIM7]、[PRAME、B4GALT1]、[PRAME、SAT1]、[PRAME、RUNX1]、[PRAME、SOCS3]、[PDIA4、NR4A1]、[PDIA4、PDLIM7]、[PDIA4、B4GALT1]、[PDIA4、SAT1]、[PDIA4、RUNX1]、[PDIA4、SOCS3]、[NR4A1、PDLIM7]、[NR4A1、B4GALT1]、[NR4A1、SAT1]、[NR4A1、RUNX1]、[NR4A1、SOCS3]、[PDLIM7、B4GALT1]、[PDLIM7、SAT1]、[PDLIM7、RUNX1]、[PDLIM7、SOCS3]、[B4GALT1、SAT1]、[B4GALT1、RUNX1]、[B4GALT1、SOCS3]、[SAT1、RUNX1]、[SAT1、SOCS3]、または[RU

10

20

30

40

50

NX1、SOCS3]の角カッコ([・・・])内で対をなす特異的組合せのうちいずれか；ならびに/または

(p)MAGEA2、PRAME、PDIA4、NR4A1、PDLIM7、B4GALT1、SAT1、RUNX1、およびSOCS3のリストに由来する1つ(または2つ)の他の非重複遺伝子(複数可)と組み合わせられた、(o)内の対のうちいずれかにより記載される3つ(または4つ)の組合せ；ならびに/または

(q)下記：

[miR.122、miR.1291]、[miR.122、miR.191]、[miR.122、miR.19b]、[miR.122、miR.200a]、[miR.122、miR.200c]、[miR.122、miR.203]、[miR.122、miR.205]、[miR.122、miR.21]、[miR.122、miR.23b]、[miR.122、miR.29c]、[miR.122、miR.342.3p]、[miR.122、miR.375]、[miR.122、miR.665]、[miR.122、miR.1304]、[miR.122、miR.142.5p]、[miR.122、miR.1254]、[miR.122、let.7a]、[miR.122、miR.140.5p]、[miR.122、miR.183]、[miR.1291、miR.191]、[miR.1291、miR.19b]、[miR.1291、miR.200a]、[miR.1291、miR.200c]、[miR.1291、miR.203]、[miR.1291、miR.205]、[miR.1291、miR.21]、[miR.1291、miR.23b]、[miR.1291、miR.29c]、[miR.1291、miR.342.3p]、[miR.1291、miR.375]、[miR.1291、miR.665]、[miR.1291、miR.1304]、[miR.1291、miR.142.5p]、[miR.1291、miR.1254]、[miR.1291、let.7a]、[miR.1291、miR.140.5p]、[miR.1291、miR.183]、[miR.191、miR.19b]、[miR.191、miR.200a]、[miR.191、miR.200c]、[miR.191、miR.203]、[miR.191、miR.205]、[miR.191、miR.21]、[miR.191、miR.23b]、[miR.191、miR.29c]、[miR.191、miR.342.3p]、[miR.191、miR.375]、[miR.191、miR.665]、[miR.191、miR.1304]、[miR.191、miR.142.5p]、[miR.191、miR.1254]、[miR.191、let.7a]、[miR.191、miR.140.5p]、[miR.191、miR.183]、[miR.19b、miR.200a]、[miR.19b、miR.200c]、[miR.19b、miR.203]、[miR.19b、miR.205]、[miR.19b、miR.21]、[miR.19b、miR.23b]、[miR.19b、miR.29c]、[miR.19b、miR.342.3p]、[miR.19b、miR.375]、[miR.19b、miR.665]、[miR.19b、miR.1304]、[miR.19b、miR.142.5p]、[miR.19b、miR.1254]、[miR.19b、let.7a]、[miR.19b、miR.140.5p]、[miR.19b、miR.183]、[miR.200a、miR.200c]、[miR.200a、miR.203]、[miR.200a、miR.205]、[miR.200a、miR.21]、[miR.200a、miR.23b]、[miR.200a、miR.29c]、[miR.200a、miR.342.3p]、[miR.200a、miR.375]、[miR.200a、miR.665]、[miR.200a、miR.1304]、[miR.200a、miR.142.5p]、[miR.200a、miR.1254]、[miR.200a、let.7a]、[miR.200a、miR.140.5p]、[miR.200c、miR.203]、[miR.200c、miR.205]、[miR.200c、miR.21]、[miR.200c、miR.23b]、[miR.200c、miR.29c]、[miR.200

c、miR.342.3p]、[miR.200c、miR.375]、[miR.200c、miR.665]、[miR.200c、miR.1304]、[miR.200c、miR.142.5p]、[miR.200c、miR.1254]、[miR.200c、let.7a]、[miR.200c、miR.140.5p]、[miR.200c、miR.183]、[miR.203、miR.205]、[miR.203、miR.21]、[miR.203、miR.23b]、[miR.203、miR.29c]、[miR.203、miR.342.3p]、[miR.203、miR.375]、[miR.203、miR.665]、[miR.203、miR.1304]、[miR.203、miR.142.5p]、[miR.203、miR.1254]、[miR.203、let.7a]、[miR.203、miR.140.5p]、[miR.203、miR.183]、[miR.205、miR.21]、[miR.205、miR.23b]、[miR.205、miR.29c]、[miR.205、miR.342.3p]、[miR.205、miR.375]、[miR.205、miR.665]、[miR.205、miR.1304]、[miR.205、miR.142.5p]、[miR.205、miR.1254]、[miR.205、let.7a]、[miR.205、miR.140.5p]、[miR.205、miR.183]、[miR.21、miR.23b]、[miR.21、miR.29c]、[miR.21、miR.342.3p]、[miR.21、miR.375]、[miR.21、miR.665]、[miR.21、miR.1304]、[miR.21、miR.142.5p]、[miR.21、miR.1254]、[miR.21、let.7a]、[miR.21、miR.140.5p]、[miR.21、miR.183]、[miR.23b、miR.29c]、[miR.23b、miR.342.3p]、[miR.23b、miR.375]、[miR.23b、miR.665]、[miR.23b、miR.1304]、[miR.23b、miR.142.5p]、[miR.23b、miR.1254]、[miR.23b、let.7a]、[miR.23b、miR.140.5p]、[miR.23b、miR.183]、[miR.29c、miR.342.3p]、[miR.29c、miR.375]、[miR.29c、miR.665]、[miR.29c、miR.1304]、[miR.29c、miR.142.5p]、[miR.29c、miR.1254]、[miR.29c、let.7a]、[miR.29c、miR.140.5p]、[miR.29c、miR.183]、[miR.342.3p、miR.375]、[miR.342.3p、miR.665]、[miR.342.3p、miR.1304]、[miR.342.3p、miR.142.5p]、[miR.342.3p、miR.1254]、[miR.342.3p、let.7a]、[miR.342.3p、miR.140.5p]、[miR.342.3p、miR.183]、[miR.375、miR.665]、[miR.375、miR.1304]、[miR.375、miR.142.5p]、[miR.375、miR.1254]、[miR.375、let.7a]、[miR.375、miR.140.5p]、[miR.375、miR.183]、[miR.665、miR.1304]、[miR.665、miR.142.5p]、[miR.665、miR.1254]、[miR.665、let.7a]、[miR.665、miR.140.5p]、[miR.665、miR.183]、[miR.1304、miR.142.5p]、[miR.1304、miR.1254]、[miR.1304、let.7a]、[miR.1304、miR.140.5p]、[miR.1304、miR.183]、[miR.142.5p、miR.1254]、[miR.142.5p、let.7a]、[miR.142.5p、miR.140.5p]、[miR.142.5p、miR.183]、[miR.1254、let.7a]、[miR.1254、miR.140.5p]、[miR.1254、miR.183]、[let.7a、miR.140.5p]、[let.7a、miR.183]、または[miR.140.5p、miR.183]の角カッコ( [ · · · ] )内で対をなす特異的组合せ(各場合において、「hsa」は除外されるが、識別子の一部として意図される)のうちのいずれか；

10

20

30

40

50

ならびに/または

( r ) h s a . m i R . 1 2 2、 h s a . m i R . 1 2 9 1、 h s a . m i R . 1 9 1、  
h s a . m i R . 1 9 b、 h s a . m i R . 2 0 0 a、 h s a . m i R . 2 0 0 c、 h s  
a . m i R . 2 0 3、 h s a . m i R . 2 0 5、 h s a . m i R . 2 1、 h s a . m i R  
. 2 3 b、 h s a . m i R . 2 9 c、 h s a . m i R . 3 4 2 . 3 p、 h s a . m i R .  
3 7 5、 h s a . m i R . 6 6 5、 h s a . m i R . 1 3 0 4、 h s a . m i R . 1 4 2  
. 5 p、 h s a . m i R . 1 2 5 4、 h s a . l e t . 7 a、 h s a . m i R . 1 4 0 .  
5 p、または h s a . m i R . 1 8 3 のリストに由来する 1 つ ( または 2 つ ) の他の非重  
複 m i R N A ( 複数可 ) と組み合わされた、 ( q ) 内の対のうちのいずれかにより記載され  
る 3 つ ( または 4 つ ) の組合せ。

10

【 0 0 8 7 】

特定の方法の実施形態は、疾患特異的バイオマーカー ( 複数可 ) ( 例えば、上記の ( a ) ~ ( r ) を参照されたい ) の発現を、少なくとも 1 つの正規化バイオマーカーに対して正規化するステップを含む。本開示の別の箇所ですらに詳細に論じられる通り、正規化は、いくつかの方法の実施形態に包含されるステップであって、遺伝子発現値のある種の交絡変動 ( c o n f o u n d i n g v a r i a b i l i t y ) を制御するのに有用なステップである。全ての疾患特異的変数の値を、指定された正規化バイオマーカーの発現に対して調整すること ( 例えば、除算または減算により ) は、このような疾患特異的変数を正規化する、 1 つの非限定的な方途である。一般的な規則として、指定された正規化バイオマーカーは、被験体の試料型の間 ( 母斑試料型と原発性黒色腫試料型との間など ) で発現の統計学的有意差を有していない。母斑試料および黒色腫試料のための例示的な正規化バイオマーカーを、表 3 に列挙する。いくつかの開示される方法は、疾患特異的バイオマーカー ( 例えば、表 4、 1 1、および/または 1 3 を参照されたい ) の発現を、表 3 で列挙されるか、または ( i ) M F I 2、 R A P 2 B、 B M P 1、 N C O R 2、 R P S 6 K B 2、および S D H A ; ( i i ) B M P - 1、 M F I 2、 N C O R 2、および R A P 2 b ; もしくは ( i i i ) R P S 6 K B 2 および S D H A からなる群から選択される、正規化バイオマーカー ( 複数可 ) のうちの少なくとも 2 つ、少なくとも 3 つ、少なくとも 4 つ、少なくとも 5 つ、少なくとも 6 つ、少なくとも 7 つ、少なくとも 8 つ、または全てについての発現レベルに対して正規化するステップを想定する。本明細書では、例示的な正規化バイオマーカーを指定するが、下記では、本開示される方法において有用な、他の正規化法についても論じられる。

20

30

【 0 0 8 8 】

遺伝子発現の検出

開示される方法は、本明細書で発見される遺伝子 ( 表 4、 1 1、および/または 1 3 を参照されたい ) であって、良性 ( 例えば、母斑 ) のメラニン細胞含有試料を、悪性 ( 例えば、原発性黒色腫 ) のメラニン細胞含有試料から鑑別するか、またはこのような試料型 ( 表 3 を参照されたい ) の発現レベルを正規化するのに適する遺伝子の発現を検出することもさらに包含する。様々な技法が、対象の試料中の遺伝子発現を測定するのに利用可能である ( または利用可能になりうる ) 。しかし、本開示は、遺伝子発現を得るか、測定するか、または検出する特定の方法に限定されない。多くのこのような技法は、このような試料中で発現した遺伝子産物 ( 例えば、核酸 ( m R N A または m i R N A など ) および/またはタンパク質 ) を検出することを包含する。その結果として得られる遺伝子産物を測定することとは独立に、遺伝子または染色体 D N A の活性 ( 例えば、転写速度 ) を直接検出することもまた可能な ( または可能となりうる ) 場合があり、このような技法もまた、本明細書で開示される方法において有用である。

40

【 0 0 8 9 】

開示される方法では、遺伝子発現レベルは、 P C R またはヌクレアーゼ保護アッセイまたは核酸シーケンシングなど、溶液ベースの ( すなわち、 e x s i t u ) アッセイを使用して決定することができる。他の例では、発現レベルは、インサイチュ ( i n s i t u ) アッセイを使用して、例えば、免疫組織化学検査またはインサイチュハイブリダイゼ

50

ーションを使用して決定または検出する。

【0090】

核酸遺伝子産物の検出

核酸遺伝子産物とは、名称が示唆する通り、遺伝子発現の産物であって、核酸である産物である。それらの発現を検出する例示的な核酸は、cDNA、タンパク質をコードするRNA（例えば、mRNA）またはタンパク質をコードしないRNA（例えば、miRNAまたはlncRNA）などのDNAまたはRNAを含む。特定の例では、方法は、mRNAの発現、miRNAの発現、またはこれらの両方を検出することを含む。RNAまたはDNAの相補鎖の間の塩基対合（すなわち、核酸ハイブリダイゼーション）は、核酸遺伝子産物を検出するための技法の大きな代表的なクラスの基盤の全部または一部を形成する。他の代表的な検出法は、核酸シーケンシングを包含し、これは、ハイブリダイゼーションステップおよび/またはバイオインフォマティクスステップ（例えば、核酸配列情報を、その対応する遺伝子と関連させるための）を伴う場合もあり、これらを伴わない場合もある。本明細書では、代表的な技法について記載するが、当技術分野では、核酸を検出する、これらの方法および他の方法が公知であり、本開示は、核酸検出の特定の方法に限定されることを意図しない。

10

【0091】

開示される方法のいくつかの実施形態では、メラニン細胞含有試料中の遺伝子発現のレベルを決定するステップは、例えば、試料中のそのような核酸の相対量または実際量を決定することにより、表4、11、および/または13に示される2つ以上の核酸を検出することを含む（かつ、いくつかの例ではまた、表3に示される1または複数の核酸を検出することを含む）。例示的な核酸は、cDNA、miRNA、またはmRNAなど、DNAまたはRNAを含む。

20

【0092】

核酸分子の発現レベルは、例えば、*in vitro*における核酸増幅および/または核酸ハイブリダイゼーションを使用して、検出または測定することができる。このような検出法の結果は、例えば、ハイブリダイゼーションの量または増幅の量を決定することにより、定量化することができる。したがって、いくつかの例では、本明細書で提示される方法においてバイオマーカーの発現レベルを決定する（表3、4、11、および/または13のバイオマーカーの発現レベルを、個別にまたは表6、8、または14の組合せを含む任意の組合せで決定するなどの）ステップは、試料を、複数の核酸プローブ（ヌクレアーゼ保護プローブ（NPP）など）または対合増幅プライマーと接触させることを含むことが可能であり、ここで、各プローブまたは対合プライマーは、複数の核酸プローブまたは対合プライマーが、表4、11、および/または13の、その/それらの相補的な少なくとも2つのバイオマーカーとハイブリダイズすることを可能とする条件下で、少なくとも2つの、表4、11、および/または13の非重複バイオマーカーのうちの1つに対して特異的、かつ、相補的である。一例では、方法はまた、試料を複数の核酸プローブ（NPPなど）と接触させた後で、試料を、一本鎖核酸分子を消化するヌクレアーゼと接触させることも含む。

30

【0093】

任意選択の核酸の単離

いくつかの例では、核酸は、試料中のこのような核酸を、相補的な核酸プローブまたはプライマーと接触させ、かつ/または試料中のこのような核酸を他の形で検出する前に、メラニン細胞含有試料から単離または抽出する。核酸（RNA（例えば、mRNAまたはmiRNA）またはDNAなど）は、いくつかの方法のうちのいずれかに従い、試料から単離することができる。核酸を単離および精製する代表的な方法は、Laboratory Techniques in Biochemistry and Molecular Biology、3章：Hybridization With Nucleic Acid Probes、I部Theory and Nucleic Acid Preparation、P. Tijssen編、Elsevier、N.Y.（199

40

50

3年)において詳細に記載されている。当技術分野では、RNA(例えば、mRNAまたはmiRNA)を抽出するための代表的な方法も同様に周知であり、Ausubelら、Current Protocols of Molecular Biology、John Wiley and Sons(1997年)を含む、分子生物学の標準的な教科書において開示されている。

【0094】

具体的な方法は、例えば、酸性グアニジンフェノールクロロホルム抽出法を使用して、全核酸を試料から単離するステップ、および/またはオリゴdTカラムクロマトグラフィーもしくは(dT)n磁気ビーズにより、ポリA+mRNAを単離するステップを包含しうる(例えば、Sambrookら、Molecular Cloning: A Laboratory Manual(2版)、1~3巻、Cold Spring Harbor Laboratory(1989年);またはCurrent Protocols in Molecular Biology、F. Ausubelら編、Greene Publishing and Wiley-Interscience、N.Y.(1987年)を参照されたい)。他の例では、核酸の単離は、製造元の指示に従い、QIAGEN(登録商標)(Valencia、CA)など、商業的製造元からの精製キット、緩衝液セット、およびプロテアーゼを使用して実施することができる。例えば、細胞(被験体から得られる細胞など)に由来する全RNAは、QIAGEN(登録商標)RNeasy mini-columnを使用して単離することができる。他の市販される核酸単離キットとしては、MASTERPURE(登録商標) Complete DNA and RNA Purification Kit(EPICENTRE(登録商標)、Madison、Wis)、およびParaffin Block RNA Isolation Kit(Ambion、Inc)が挙げられる。組織試料に由来する全RNAは、RNA Stat-60(Tel-Test)を使用して単離することができる。腫瘍または他の生物学的試料から調製されるRNAは、例えば、塩化セシウム密度勾配遠心分離により単離することができる。パラフィン包埋組織からRNAを抽出するための方法は、例えば、Rupp and Locker、Biotechniques、6巻:56~60頁(1988年);およびDe Andresら、Biotechniques、18巻:42~44頁(1995年)において開示されている。

【0095】

核酸(例えば、RNA(mRNAまたはmiRNAなど)またはDNA)を、試料から単離または抽出した後で、単離された核酸の濃度の測定、分解されるかまたは損傷したRNAの修復(または回収)、RNAの逆転写、および/またはRNAもしくはDNAの増幅を含む、いくつかの任意選択の他のステップのうちのいずれかを実施して、このような核酸を検出のために調製することができる。

【0096】

他の例では、試料(例えば、FFPEメラニン細胞含有組織試料)を、緩衝液(例えば、溶解緩衝液)中に懸濁させ、懸濁させた試料中に存在する核酸(RNAまたはDNAなど)を、このような懸濁させた試料から単離または抽出(例えば、全体的にまたは部分的に精製)せず、このような懸濁液中で、1または複数の相補的な核酸プローブ(例えば、ヌクレアーゼ保護プローブ)と接触させ、これにより、核酸(例えば、RNA)を試料から単離または抽出する必要を排する。この実施形態は、懸濁させた試料中に存在する核酸(RNAまたはDNAなど)が、細胞構造へと架橋または固定されており、たやすく単離可能または抽出可能ではない場合に特に有利である。それらの検出のために、このようなプローブの伸長が要請されない、比較的短い(例えば、100塩基対未満、75~25塩基対、または50~25塩基対など)プローブは、いくつかの、抽出を伴わない方法の実施形態において有用である。当業者は、このような伸長のための核酸鋳型(例えば、RNA)が分解されるか、または他の形でアクセス可能でなくなる場合、プローブの伸長(例えば、PCRまたはプライマーの伸長)を必要とする方法は、信頼できないことを十分に理解する。試料中の核酸(例えば、RNA)を、このような核酸の既往の抽出を伴わずに

10

20

30

40

50

検出するための具体的な方法（例えば、q N P A）については、本明細書の別の箇所で詳細に記載する。

【 0 0 9 7 】

核酸ハイブリダイゼーション

いくつかの例では、開示されるバイオマーカー（表 4、11、および/または13のバイオマーカーなど）または正規化バイオマーカー（例えば、表3）の発現レベルを、本明細書で提示される方法において決定するステップは、試料を、複数の核酸プローブ（ヌクレアーゼ保護プローブ（N P P）または隣接ライゲーション可能プローブなど）または対合増幅プライマーと接触させることを包含することが可能であり、ここで、複数の中の各プローブ（またはライゲーション可能プローブのセット）または対合プライマーは、複数の核酸プローブまたは対合プライマーが、表4、11、および/または13のその/それらの相補的なバイオマーカーとハイブリダイズすることを可能とする条件下で、少なくとも2つの表4、11、および/もしくは13のバイオマーカーのうちの一つまたは表3の正規化バイオマーカーに対して特異的、かつ、相補的である。一例では、方法はまた、試料を複数の核酸プローブ（N P Pなど）と接触させた後で、試料を、一本鎖核酸分子を消化するヌクレアーゼと接触させることも包含しうる。他の例では、表4、11、および/もしくは13の、少なくとも2つのバイオマーカーの各々、または表3の正規化バイオマーカーを、各々のこのようなバイオマーカーに特異的な複数の（例えば、2つ、3つ、4つ、5つ、または6つの）プローブからなる「プローブセット」であって、そのデザインが、例えば、このような遺伝子産物から得られるシグナルを増大させるか、または同じ遺伝子産物の複数の変異体を検出するのに有用でありうる「プローブセット」と接触させる。

10

20

【 0 0 9 8 】

いくつかの例では、可変核酸（例えば、表4、11、および/または13）または正規化核酸（例えば、表3）を、核酸ハイブリダイゼーションにより検出する。核酸ハイブリダイゼーションは、プローブおよび標的核酸（例えば、表4、11、および/または13のもの）を、プローブとその相補的な標的とが、相補的な塩基対合を介して安定的なハイブリッド二重鎖を形成しうる条件下で変性させることを伴う。いくつかの例では、次いで、ハイブリッド二重鎖を形成しない核酸を除去し（例えば、洗い流すか、ヌクレアーゼにより消化させるか、または物理的に除去する）、典型的には、（直接的または間接的に）結合させた検出可能な標識の検出を介して、ハイブリダイズした核酸が検出されるようにする。具体例では、そのそれぞれの標的とハイブリダイズしない任意の過剰プローブ、およびプローブと相補的でない標的配列の領域など、ハイブリッド二重鎖を形成しない核酸は、ヌクレアーゼを添加することにより消化させ、相補的なプローブの標的配列のハイブリッド二重鎖だけを後に残すことができる。

30

【 0 0 9 9 】

核酸は、温度を上昇させることおよび/または核酸を含有する緩衝液の塩濃度を低下させることにより変性させることが一般に認識されている。低度な厳密性条件（例えば、低温および/または高塩濃度）下では、アニールした配列が完全に相補的でない場合であってもなお、ハイブリッド二重鎖（例えば、DNA : DNA、RNA : RNA、またはRNA : DNA）が形成される。したがって、低度な厳密性では、ハイブリダイゼーションの特異性が低減される。逆に、高度な厳密性（例えば、高温または低塩濃度）では、ハイブリダイゼーションの成功には、より少ないミスマッチが必要とされる。当業者は、ハイブリダイゼーション条件は、異なる程度の厳密性をもたらすようにデザインしうることを十分に理解する。ロックド核酸またはペプチド核酸を含ませることなど、プローブ内に非天然塩基を含ませることにより、ハイブリダイゼーションの厳密性を低下させずに、ハイブリダイゼーションの強度を増大させることができ、したがって、高度な厳密性の緩衝液中でハイブリダイゼーションの特異性を維持することができる。

40

【 0 1 0 0 】

一般に、ハイブリダイゼーションの特異性（厳密性）とシグナル強度との間ではトレー

50

ドオフが存在する。したがって、一実施形態では、洗浄は、一貫した結果を生み出し、かつ、バックグラウンド強度を約10%を超えるシグナル強度をもたらす最高の厳密性において実施する。したがって、ハイブリダイゼーション複合体（例えば、アレイの表面上で捕捉されるような）は、逐次的に高度な厳密性溶液で洗浄し、各洗浄間で読み取ることができる。このように生み出されるデータセットを解析することにより、洗浄の厳密性であって、それを上回れば、ハイブリダイゼーションパターンが感知可能な形で変化することはなく、対象の特定のオリゴヌクレオチドプローブの十分なシグナルをもたらす厳密性が明らかになる。

#### 【0101】

これらの方法により検出される核酸の発現および/または核酸の存在の変化は、例えば、このような核酸のレベル（量）の上昇もしくは低下または機能的活性の増大もしくは減少、それらの発現またはタンパク質への翻訳の増大または減少、あるいはそれらの局在化または安定性の増大または減少を含みうる。例えば、正規化バイオマーカー（例えば、表3を参照されたい）と比べた増大または減少は、例えば、表4、11、および/または13のうちのいずれかに示されるバイオマーカーに対応する核酸など、特定の核酸の発現および/または存在の、約1倍、2倍、3倍、4倍、5倍の変化（増大または減少）など、少なくとも1倍、少なくとも2倍、または少なくとも5倍でありうる。マルチプレックス化された方法の実施形態では、特に、各々のこのような遺伝子も同様に正規化されている（例えば、1または複数の共検出される正規子遺伝子の発現に対して；例えば、表3を参照されたい）場合はまた、非正規子遺伝子（例えば、可変遺伝子；例えば、表4、11、および/または13）の相対的な発現も比較することができる。よって、1つの可変遺伝子の正規化された発現は、別の可変遺伝子の正規化された発現よりも約1倍、1.5倍、2倍、3倍、4倍、5倍など、少なくとも1倍、少なくとも2倍、もしくは少なくとも5倍高いか、またはそれよりも約1倍、1.5倍、2倍、3倍、4倍、5倍など、少なくとも1倍、少なくとも2倍、もしくは少なくとも5倍低くてよい。

#### 【0102】

いくつかの実施形態では、遺伝子発現は、マルチプレックス化された方法および/またはハイスループット法を使用して測定する。マルチプレックス化された方法では、単一の試料中で複数の測定（例えば、遺伝子発現の測定）を行う。単一の試料中の多数の遺伝子のモニタリングを可能とする多様な技術が発展している（例えば、従来のマイクロアレイ、マルチプレックス化されたPCR、連続遺伝子発現分析（SAGE；例えば、米国特許第5,866,330号）、マルチプレックスライゲーション依存型プローブ増幅（MLPA）、ハイスループットシーケンシング、標識されたビーズベースの技術（例えば、米国特許第5,736,330号および同第6,449,562号）、デジタル式分子バーコード技術（例えば、米国特許第7,473,767号））。ハイスループット法では、複数の試料中の遺伝子発現を同時に測定する。また、ハイスループット法をマルチプレックス化する（すなわち、複数の試料の各々における複数の遺伝子を同時に検出する）こともできる。

#### 【0103】

いくつかの実施形態では、1または複数のバイオマーカー（表4、11、および/もしくは13のバイオマーカーのうちの2つ以上（例えば、表6、8、または14の任意の遺伝子の組合せ）ならびに/または表3の少なくとも1つのバイオマーカーなど）の発現レベルを、単一のメラニン細胞含有試料中または複数のメラニン細胞含有試料中（異なる被験体に由来する試料中など）で同時に決定する。一例では、表4、11、および/もしくは13で列挙されるバイオマーカーのうちの少なくとも2つ、少なくとも3つ、少なくとも4つ、少なくとも5つ、少なくとも6つ、少なくとも7つ、少なくとも8つ、少なくとも9つ、少なくとも10、少なくとも11、少なくとも12、少なくとも13、少なくとも14、少なくとも15、少なくとも16、少なくとも17、少なくとも18、少なくとも19、または、当てはまる場合、少なくとも20、少なくとも21、少なくとも22、少なくとも23、少なくとも24、少なくとも25、少なくとも26、少なくとも27、

10

20

30

40

50

少なくとも28、少なくとも29、少なくとも30、少なくとも31、もしくは全て(表4、11、および/または13のバイオマーカーのうちの2つ、3つ、4つ、5つ、6つ、7つ、8つ、9つ、10、11、12、13、14、15、16、17、18、19、20、または、当てはまる場合、21、22、23、24、25、26、27、28、29、30、31、または全てなど)、または、例えば、表6、8、もしくは14の遺伝子の組合せのうちのいずれかは、同じ試料中または複数の試料中で同時に検出することができ、いくつかの例では、表3で列挙される正規化バイオマーカー(または本明細書で提示される方法により同定される他の正規化バイオマーカー(複数可))のうちの少なくとも2つ、少なくとも3つ、少なくとも4つ、少なくとも5つ、または6つ全ても同時に、例えば、少なくとも2つの、表4、11、および/または13のバイオマーカーと同時に検出する。複数の試料は、少なくとも2つの異なる試料(例えば、少なくとも2つの異なる被験体に由来する試料および/または同じ被験体の腫瘍または身体の異なる領域に由来する試料)など、複数の異なる被験体に由来する場合もあり、かつ/または同じ被験体に由来する複数の試料に由来する場合もある。いくつかの例では、少なくとも2つ、少なくとも5つ、少なくとも10、少なくとも20、少なくとも50、少なくとも100、少なくとも500、少なくとも1000、少なくとも2000、少なくとも5000、なおまたは少なくとも10,000のメラニン細胞含有試料を同時に分析する(10~100、100~1000、1000~10000、1000~50000、または10000~10,0000のメラニン細胞含有試料を同時に分析するなど)。

10

#### 【0104】

20

本開示はまた、ハイスループットスクリーニングのための統合システムを活用する方法も含む。システムは、典型的には、流体を供給源から供給先へと輸送するロボット電機子、ロボット電機子を制御するコントローラー、検出器、検出を記録するデータ保存ユニット、および、例えば、1もしくは複数のプログラミングリンカーを含むマイクロ滴定プレート、または標的(表4、11、および/または13のバイオマーカーのうちの2つ以上、ならびに表3の正規化マーカーのうちの1または複数など)と直接ハイブリダイズしうる、1もしくは複数のオリゴヌクレオチドを含むマイクロ滴定プレートなどのアッセイ成分を含む。

#### 【0105】

アレイとは、遺伝子発現をマルチプレックス検出するための、1つの有用な(非限定的な)ツールのセットである。アレイとは、要素(例えば、分析物捕捉試薬(標的特異的オリゴヌクレオチドプローブ、アプタマー、または抗体など))の系統的配置であって、値(例えば、遺伝子発現値)のセットを検索表と関連させうる系統的配置である。配置される要素は、個別に同定可能な表面(例えば、フローチャンネルまたはビーズ)を使用して、またはこれらの組合せにより、単一の表面上で系統的に同定する(例えば、空間的マッピングまたは示差的タグ付け(differential tagging)により)ことができる。

30

#### 【0106】

開示されるバイオマーカーを検出するのに使用されうる方法およびアッセイシステムの他の例は、それらがハイスループットアッセイ技法について記載する限りにおいて、参照により本明細書に組み込まれる、国際特許公開第WO2003/002750号および同第WO2008/1221927号、同第WO1999/032663号、同第WO2000/079008号、同第WO/2000/037684号、および同第WO2000/037683号、ならびに米国特許第6,232,066号、同第6,458,533号、同第6,238,869号、および同第7,659,063号において開示されている、ハイスループットアッセイ技法である。

40

#### 【0107】

いくつかのアレイの実施形態では、表3、4、11、および/または13に示される遺伝子の1または複数の産物を(直接的または間接的に)捕捉するようにデザインされた核酸プローブ(オリゴヌクレオチドなど)を、マイクロチップ基板上にプレーティングまた

50

はアレイする。例えば、アレイは、表3、4、11、および/または13に示される遺伝子のうちの少なくとも2つ(このような遺伝子のうちの少なくとも3つ、少なくとも5つ、少なくとも10、少なくとも20、もしくは全て、または表6、8、もしくは14の遺伝子の組合せのうちのいずれか、または本明細書において他の形で開示される遺伝子など)ならびに、必要に応じて、表3に示される遺伝子のうちの少なくとも1つと相補的なオリゴヌクレオチドを含みうる。他の例では、アレイは、表3、4、11、および/または13に示される遺伝子のうちの少なくとも2つ(このような遺伝子のうちの少なくとも3つ、少なくとも5つ、少なくとも10、少なくとも20、または全て、または表6、8、もしくは14の遺伝子の組合せのうちのいずれか、または本明細書において他の形で開示される遺伝子など)ならびに、必要に応じて、表3に示される遺伝子のうちの少なくとも1つの産物と相補的なヌクレアーゼ保護プローブの部分と相補的なオリゴヌクレオチドを含みうる。

10

**【0108】**

次いで、アレイした配列を、被験試料(例えば、被験体から得られたメラニン細胞含有試料であって、良性の母斑または悪性黒色腫(例えば、原発性黒色腫)としての特徴付けが所望されるメラニン細胞含有試料)から単離された核酸(cDNA、miRNAまたはmRNAなど)とハイブリダイズさせる。一例では、被験試料から単離された核酸は、アレイ上の特異的な相補的なオリゴヌクレオチドとのそれらのハイブリダイゼーションを決定しうるように標識する。代替的に、被験試料の核酸を標識せず、アレイ上のオリゴヌクレオチドと標的核酸とのハイブリダイゼーションを、例えば、標的と相補的なさらなるオリゴヌクレオチドであって、標識されたオリゴヌクレオチドを使用するサンドイッチアッセイを使用して検出する。

20

**【0109】**

一実施形態では、試料核酸へと結合させるか、または標的核酸と直接的もしくは間接的にハイブリダイズする核酸プローブへと結合させた、1または複数の標識を検出することにより、ハイブリダイズした核酸を検出する。標識は、いくつかの方法のうちのいずれかにより組み込むことができる。一例では、標識は、試料核酸を調製するときの増幅ステップにおいて同時に組み込む。したがって、例えば、標識されたプライマーまたは標識されたヌクレオチドによるポリメラーゼ連鎖反応(PCR)は、標識された増幅産物をもたらす。一実施形態では、標識されたヌクレオチド(フルオレセインで標識されたUTPおよび/またはCTPなど)を使用する転写増幅により、標識を、転写される核酸へと組み込む。

30

**【0110】**

本開示を通じた実施形態における使用に適する、検出可能な標識は、分光的手段、光化学的手段、生化学的手段、免疫化学的手段、電気的手段、光学的手段、または化学的手段により検出可能な任意の組成物を含む。有用な標識は、標識されたストレプトアビジンコンジュゲートによる染色のためのビオチン、磁気ビーズ(例えば、DYNABEADS(商標))、蛍光色素(例えば、フルオレセイン、Texas red、ローダミン、緑色蛍光タンパク質など)、化学発光マーカー、放射性標識(例えば、 $^3\text{H}$ 、 $^{125}\text{I}$ 、 $^{35}\text{S}$ 、 $^{14}\text{C}$ 、または $^{32}\text{P}$ )、酵素(例えば、西洋ワサビペルオキシダーゼ、アルカリホスファターゼ、およびELISAにおいて一般に使用される他の酵素)、および金コロイドまたは有色ガラス製ビーズもしくは有色プラスチック(例えば、ポリスチレン、ポリプロピレン、ラテックスなど)製ビーズなどの比色標識を含む。このような標識の使用について教示する特許としては、米国特許第3,817,837号;米国特許第3,850,752号;米国特許第3,939,350号;米国特許第3,996,345号;米国特許第4,277,437号;米国特許第4,275,149号;および米国特許第4,366,241号が挙げられる。いくつかの実施形態では、標識は、多様な長さのスペーサーアームにより結合させて、潜在的な立体障害を低減する。

40

**【0111】**

このような標識を検出する手段もまた周知である。したがって、例えば、放射性標識は

50

、写真用フィルムまたはシンチレーションカウンターを使用して検出することができ、蛍光マーカは、発光した光を検出する光検出器を使用して検出することができる。酵素標識は典型的に、酵素に基質をあてがい、基質に対する酵素の作用によって生成する反応生成物を検出することにより検出し、比色標識は、有色標識を単に可視化することにより検出する。

#### 【0112】

標識は、ハイブリダイゼーションの前に、標的(試料)核酸(複数可)へと添加することもでき、ハイブリダイゼーションの後で、標的(試料)核酸(複数可)へと添加することもできる。いわゆる「直接的標識」とは、ハイブリダイゼーションの前に、標的(試料)核酸へと直接結合させるかまたは組み込む検出可能な標識である。これに対し、いわゆる「間接的標識」は、ハイブリダイゼーションの後で、ハイブリッド二重鎖へと接続する。間接的標識は、ハイブリダイゼーションの前に標的核酸へと結合させた、結合性部分へと結合させることが多い。したがって、例えば、標的核酸は、ハイブリダイゼーションの前にビオチン化することができる。ハイブリダイゼーションの後では、アビジンをコンジュゲートしたフルオロフォアが、ビオチンを保有するハイブリッド二重鎖に結合し、容易に検出される標識をもたらす(Laboratory Techniques in Biochemistry and Molecular Biology、24巻: Hybridization With Nucleic Acid Probes、P. Tijssen編、Elsevier、N.Y.、1993年を参照されたい)。

#### 【0113】

発色性インサイチュハイブリダイゼーション(CISH)または銀インサイチュハイブリダイゼーション(SISH)などのインサイチュハイブリダイゼーション(ISH)は、対象の遺伝子(表3、4、11、および/または13の対象の遺伝子など)の発現を検出および比較するための例示的な方法である。ISHとは、相補的な核酸を使用して、1もしくは複数の特異的な核酸配列を、組織の部分もしくは切片において(インサイチュ)、または、組織が十分に小さい場合は、組織の全体において(ホールマウントISH)位置特定するハイブリダイゼーションの種類である。RNA ISHを使用して、表4、11、および/または13のバイオマーカーの発現など、組織内の発現パターンをアッセイすることができる。試料細胞または試料組織は、それらの透過性を増大させて、表4、11、および/または13のバイオマーカーのうちの1または複数に特異的なプローブなどのプローブが、細胞に入ることを可能とするように処理することができる。プローブを、処理された細胞へと添加し、適正な温度でハイブリダイズさせ、過剰プローブを洗い流す。相補的なプローブは、例えば、オートラジオグラフィ、蛍光顕微鏡法、または免疫アッセイを使用して、組織内のプローブの位置および量を決定しうるように、放射性タグ、蛍光タグ、または抗原タグなど、検出可能な標識で標識する。

#### 【0114】

インサイチュ PCRとは、標的核酸配列のPCRベースの増幅に続く、標的およびアンプリコンのインサイチュにおける検出である。インサイチュ PCRの前に、細胞または組織試料は一般に、固定および透過処理して、形態を保存し、増幅される細胞内配列への、PCR試薬のアクセスを可能とし、必要に応じて、細胞内逆転写ステップを導入して、RNA鋳型からcDNAを生成し、これにより、低コピーのRNA配列の検出を可能とする。次に、標的配列のPCR増幅を実施し、次いで、細胞内PCR産物を、ISHまたは免疫組織化学検査により可視化する。

#### 【0115】

定量的ヌクレアーゼ保護アッセイ(qNPA)

本開示される方法の特定の実施形態では、定量的ヌクレアーゼ保護アッセイおよびアレイ(下記に記載されるアレイなど)を活用して、核酸を試料中で検出する。定量的ヌクレアーゼ保護アッセイは、それらの各々が参照によりそれらの全体において本明細書に組み込まれる、国際特許公開第WO99/032663号; 同第WO00/037683号; 同第WO00/037684号; 同第WO00/079008号; 同第WO03/002

10

20

30

40

50

750号;および同第WO08/121927号;ならびに米国特許第6,238,869号;同第6,458,533号;および同第7,659,063号において記載されている。また、Martelら、Assay and Drug Development Technologies、2002年、1巻(1-1号):61~71頁;Martelら、Progress in Biomedical Optics and Imaging、2002年、3巻:35~43頁;Martelら、Gene Cloning and Expression Technologies、Q. LuおよびM. Weiner編、Eaton Publishing, Natick(2002年);Seligmann, B., Pharmacogenomics、2003年、3巻:36~43頁;Martelら、「Array Formats」、「Microarray Technologies and Applications」、U. R. MullerおよびD. Nicolau編、Springer-Verlag, Heidelberg;Sawadaら、Toxicology in Vitro、20巻:1506~1513頁;Bakirら、Biorg. & Med. Chem Lett、17巻:3473~3479頁;Krisら、Plant Physiol.、144巻:1256~1266頁;Robertsonら、Laboratory Investigation、87巻:979~997頁;Rimszaら、Blood、2008年10月15日、112巻(8号):3425~3433頁;Pechholdら、Nature Biotechnology、27巻、1038~1042頁も参照されたい。これらの全ては、参照により本明細書に完全に組み込まれる。

10

20

## 【0116】

qNPA法を使用して、ヌクレアーゼ保護プローブ(NPP)を、標的配列とハイブリダイズさせ、これに続いて、一本鎖核酸分子を消化するヌクレアーゼとともに試料のインキュベーションを行う。したがって、プローブが検出される(例えば、プローブがヌクレアーゼにより消化されない)場合は、プローブの標的、例えば、表3、4、11、および/または13に示される標的核酸が、試料中に存在し、この存在は、検出する(例えば、定量化する)ことができる。NPPは、個々の標的についてデザインすることができ、アレイ上の同定のためのカクテルとしてアッセイへと添加することができ、したがって、複数の遺伝子標的を、同じアッセイ内および/またはアレイ内で測定することができる。

## 【0117】

いくつかの例では、メラニン細胞含有試料中の細胞を直接使用するか、またはまず水溶液中で(例えば、溶解緩衝液を使用して)溶解させるかまたは透過処理する。水溶液または溶解緩衝液は、洗浄剤(ドデシル硫酸ナトリウムなど)、および/または1もしくは複数のカオトロピック剤(ホルムアミド、グアニジンHCl、グアニジウムイソチオシアネート、または尿素など)を含む。溶液はまた、緩衝液(例えば、SSC)も含有しうる。いくつかの例では、溶解緩衝液は、約15%~25%のホルムアミド(v/v)、約0.01%~0.1%のSDS、および約0.5~6倍濃度のSSCを含む。緩衝液は必要に応じて、tRNA(例えば、約0.001~約2.0mg/ml)またはリボヌクレアーゼも含みうる。溶解緩衝液はまた、Phenol RedなどのpH指示薬も含みうる。特定の例では、溶解緩衝液は、20%のホルムアミド、3倍濃度のSSC(79.5%)、0.05%のDSD、1μg/mlのtRNA、および1mg/mlのPhenol Redを含む。細胞を、水溶液中で、細胞を溶解させるかまたは透過処理するのに十分な期間(約1分間~約60分間、例えば、約5分間~約20分間、または約10分間など)にわたり、十分な温度(約22~約115、例えば、約37~約105、または約90~約110など)でインキュベートする。いくつかの例では、検出される核酸がRNAであれば、溶解は、約95で実施する。他の例では、検出される核酸がDNAであれば、溶解は、約105で実施する。

30

40

## 【0118】

いくつかの例では、標的と相補的な核酸保護プローブ(NPP)(配列番号1~36および123~164に示されるNPPなど)を、試料へと、例えば、6倍濃度のSSPE

50

- T ( 0 . 9 M の N a C l 、 6 0 m M の N a H <sub>2</sub> P O <sub>4</sub> 、 6 m M の E D T A 、 および 0 . 0 5 % の T r i t o n X - 1 0 0 ) または溶解緩衝液 ( 上記で記載される ) などの緩衝液中に約 1 0 p M ~ 約 1 0 n M ( 約 3 0 p M ~ 5 n M 、 約 1 0 0 p M ~ 約 1 n M など ) の範囲の濃度で添加することができる。一例では、プローブを、約 3 0 p M の最終濃度で試料へと添加する。別の例では、プローブを、約 1 6 7 p M の最終濃度で試料へと添加する。さらなる例では、プローブを、約 1 n M の最終濃度で試料へと添加する。このような例では、N P P は、標的配列などの相補的な配列とハイブリダイズ ( と二重鎖を形成 ) すれば、S 1 などのヌクレアーゼにより消化されない。

#### 【 0 1 1 9 】

当業者は、N P P が、被験試料内に存在するその標的と特異的にハイブリダイズするのに十分な条件を同定することができる。例えば、当業者は、他の物質または分子との非特異的ハイブリダイゼーションを最小化しながら、核酸 ( 例えば、N P P ) が、別の核酸 ( 例えば、表 3 、 4 、 1 1 、 および / または 1 3 の標的核酸 ) と、選択された厳密性の条件下でハイブリダイズすることを可能とする特徴 ( 長さ、塩基組成、および相補性の程度など ) を、実験により決定することができる。典型的に、N P P の核酸配列は、N P P が、選択される厳密なハイブリダイゼーション条件下でハイブリダイズすること、例えば、約 3 7 以上 ( 約 3 7 、 4 2 、 5 0 、 5 5 、 6 0 、 6 5 、 7 0 、 7 5 以上など ) におけるハイブリダイゼーションを可能とするのに十分な、対応する標的配列に対する相補性を有する。様々でありうるハイブリダイゼーション反応のパラメーターの中には、塩濃度、緩衝液、p H、温度、インキュベーション時間、ホルムアミドなど、変性剤の量および種類がある。

#### 【 0 1 2 0 】

試料中の核酸を、変性させ ( 例えば、約 9 5 ~ 約 1 0 5 で約 5 ~ 1 5 分間にわたる ) 、N P P と、約 4 ~ 約 7 0 ( 例えば、約 3 7 ~ 約 6 5 、約 4 5 ~ 約 6 0 、または約 5 0 ~ 約 6 0 ) の範囲の温度で、約 1 0 分間 ~ 約 2 4 時間の間 ( 例えば、少なくとも約 1 時間 ~ 2 0 時間、または約 6 時間 ~ 1 6 時間 ) にわたりハイブリダイズさせる。いくつかの例では、プローブは、試料と共に、少なくとも約 4 0 、少なくとも約 4 5 、少なくとも約 5 0 、少なくとも約 5 5 、少なくとも約 6 0 、少なくとも約 6 5 、または少なくとも約 7 0 の温度でインキュベートする。一例では、プローブは、試料と共に、約 6 0 でインキュベートする。別の例では、N P P は、試料と共に、約 5 0 でインキュベートする。これらのハイブリダイゼーション温度は例示的なものであり、当業者は、N P P の長さおよびヌクレオチド組成などの因子に応じて、適切なハイブリダイゼーション温度を選択することができる。

#### 【 0 1 2 1 】

いくつかの実施形態では、方法は、核酸の精製を含まない ( 例えば、核酸の精製は、試料をプローブと接触させる前に実施せず、かつ / または核酸の精製は、試料をプローブと接触させた後でも実施しない ) 。いくつかの例では、細胞の溶解を除き、試料の前処理が要請されない。いくつかの例では、細胞の溶解および試料の N P P との接触は、逐次的に行い、いくつかの非限定的な例では、いかなるステップも介在させずに行う。他の例では、細胞の溶解および試料の N P P との接触を、併せて行う。

#### 【 0 1 2 2 】

1 または複数の N P P と試料中の核酸とのハイブリダイゼーションに続き、試料を、ヌクレアーゼ保護手順にかける。全長核酸とハイブリダイズした N P P は、ヌクレアーゼにより加水分解されず、その後検出することができる。

#### 【 0 1 2 3 】

1 または複数のヌクレアーゼによる処理は、試料中に存在する核酸分子へとハイブリダイズしたプローブ以外の核酸分子を破壊する。例えば、試料が、細胞抽出物または細胞溶解物を含む場合、対象の遺伝子以外のゲノム DNA、c DNA、t RNA、r RNA、および m RNA などの望ましくない核酸は、このステップで実質的に破壊することができる。当業者は、例えば、DNA を検出するのか、RNA を検出するのかに基づき、適切なヌ

10

20

30

40

50

クレーゼを選択することができる。試料中に存在するハイブリダイズした複合体および所望されない核酸の性質に応じて、豚臓RNアーゼ、緑豆(mung bean)ヌクレアーゼ、S1ヌクレアーゼ、RNアーゼA、リボヌクレアーゼT1、エキソヌクレアーゼIII、エキソヌクレアーゼVII、RNアーゼCLB、RNアーゼPhyM、RNアーゼU2などを含む、様々なヌクレアーゼのうちいずれかを使用することができる。特定の例では、ヌクレアーゼは、一本鎖核酸に特異的なヌクレアーゼ、例えば、S1ヌクレアーゼである。一本鎖核酸に特異的なヌクレアーゼを、本明細書で開示されるいくつかの方法の実施形態において使用することの利点は、このような一本鎖(「粘着性」)分子を、それらが不必要なバックグラウンドまたは交差反応性をもたらす後続の反応ステップから除去することである。S1ヌクレアーゼは、例えば、Promega、Madison、WI(型番:M5761); Life Technologies/Invitrogen、Carlsbad、CA(型番:18001-016); Fermentas、Glen Burnie、MD(型番:EN0321)などから市販されている。当技術分野では、これらの酵素のための反応条件が周知であり、経験的に最適化することができる。

10

#### 【0124】

いくつかの例では、適切な緩衝液(酢酸ナトリウム、塩化ナトリウム、硫酸亜鉛、および洗浄剤、例えば、0.25Mの酢酸ナトリウム、pH4.5、1.4MのNaCl、0.0225MのZnSO<sub>4</sub>、0.05%のKATHONを含む緩衝液など)中で希釈されたS1ヌクレアーゼを、ハイブリダイズしたプローブ混合物へと添加し、約50で約300~120分間(例えば、約60~90分間)にわたりインキュベートして、ハイブリダイズしなかった核酸および結合しなかったNPPを消化する。

20

#### 【0125】

試料は必要に応じて、ハイブリダイズしなかった材料を他の形で除去し、かつ/または残存する酵素を不活化もしくは除去するように処理する(例えば、フェノール抽出、沈殿、カラム濾過などにより)。いくつかの例では、試料は必要に応じて、標的核酸をプローブから解離させるように処理する(例えば、塩基による加水分解および加熱を使用して)。ハイブリダイゼーションの後、ハイブリダイズした標的は、例えば、ヌクレアーゼまたは化学的処理により分解し、NPPが標的へとハイブリダイズした量に直接比例してNPPを後に残すことができる。代替的に、試料は、標的の(一本鎖の)ハイブリダイズした部分、またはハイブリダイズした標的とプローブとにより形成される二重鎖を後に残すように処理してさらに解析することもできる。

30

#### 【0126】

次いで、NPPの存在(または残りの標的または標的:NPP複合体)を検出する。任意の適切な方法を使用して、プローブ(または残りの標的または標的:NPP複合体)を検出することができる。いくつかの例では、NPPは、検出可能な標識を含み、NPP(複数可)の存在を検出することは、検出可能な標識を検出することを含む。いくつかの例では、NPPを、同じ検出可能な標識で標識する。他の例では、NPPを、異なる検出可能な標識(各標的について異なる標識など)で標識する。他の例では、NPPを、例えば、標識された核酸とのハイブリダイゼーションにより間接的に検出する。いくつかの例では、NPPを、マイクロアレイ、例えば、検出可能に標識された(例えば、ビオチンまたは西洋ワサビペルオキシダーゼで標識された)核酸であって、NPPと相補的な核酸を含むマイクロアレイを使用して検出する。他の例では、NPPを、捕捉プローブおよびプロングラミングリンカーを含むマイクロアレイであって、プロングラミングリンカーの部分がNPPの部分と相補的なマイクロアレイを使用し、その後、検出リンカーであって、その部分が、NPPの個別の部分と相補的な検出リンカーと共にインキュベートして検出する。検出リンカーを、検出可能に標識しうるか、または検出リンカーの個別の部分が、検出可能な標識(ビオチンまたは西洋ワサビペルオキシダーゼなど)を含むさらなる核酸と相補的である。いくつかの例では、例えば、参照によりそれらの全体において本明細書に組み込まれる、国際特許公開第WO99/032663号;同第WO00/037683号;

40

50

同第WO00/037684号；同第WO00/079008号；同第WO03/002750号；および同第WO08/121927号；ならびに米国特許第6,238,869号；同第6,458,533号；および同第7,659,063号において記載されている通りに、NPPをマイクロアレイ上で検出する。

【0127】

略述すると、非限定的な一例では、ハイブリダイゼーションおよびヌクレアーゼ処理の後で、溶液を中和し、プログラム化されたARRAYPLATE (HTG Molecular Diagnostics, Tucson, AZ; ARRAYPLATEの各要素は、例えば、プレートへと結合させたアンカーおよびアンカーと会合させたプログラミングリンカーを活用して特異的プローブを捕捉するようにプログラム化されている)へと輸送し、インキュベーション(例えば、約50 で一晚)中にNPPを捕捉する。これとは別に、プローブは、QBEADアッセイと称するアッセイである、X-MAPビーズ(Lumines, Austin, TX)上で捕捉することもでき、所望に応じて、PCR増幅反応またはライゲーション反応を含め、さらに加工し、次いで、例えば、シーケンシングにより測定することもできる。インキュベーション(例えば、約50 で1時間)中にそれらのそれぞれの(捕捉された)プローブとハイブリダイズする、ARRAYPLATEおよびQBEADアッセイの場合には、媒質を除去し、プローブ特異的検出リンカーのカクテルを添加する。ARRAYPLATEおよびQBEADアッセイの場合に特有であるが、アレイまたはビーズを洗浄し、次いで、トリプルビオチンリンカー(triple biotin linker)(あらゆる検出リンカー上の共通の配列とハイブリダイズするオリゴヌクレオチドであって、その中に3つのビオチンが組み込まれたオリゴヌクレオチド)を添加し、インキュベートする(例えば、約50 で1時間)。ARRAYPLATE(mRNAアッセイ)では、HRPで標識されたアビジン(アビジン-HRP)またはストレプトアビジンポリHRPを添加し、インキュベートし(例えば、約37 で1時間にわたり)、次いで、洗浄して、結合しなかったアビジン-HRPまたはストレプトアビジンポリHRPを除去する。基質を添加し、プレートを画像化して、プレート内のあらゆる要素の強度を測定する。QBEADの場合は、アビジン-PEを添加し、ビーズを洗浄し、次いで、Lumines 200、FLEXMAP 3Dまたは他の適切な機器を使用するフローサイトメトリーにより測定する。当業者は、本明細書で開示される方法で活用される、NPPに基づく定量的ヌクレアーゼ保護アッセイにおける使用に適する捕捉プローブ、プログラミングリンカー、検出リンカー、および他の試薬をデザインすることができる。

【0128】

いくつかの例では、検出リンカーを使用する代わりに、NPPを直接ビオチン化する。

【0129】

核酸の増幅

いくつかの方法の例では、それらの検出の前に、またはそれらの検出のための手段として、核酸分子(核酸遺伝子産物(例えば、mRNA、miRNAまたはlncRNA)またはヌクレアーゼ保護プローブなど)を増幅する。いくつかの例では、増幅中に、例えば、リアルタイムRT-PCRを使用することにより、核酸の発現レベルを決定する。

【0130】

一例では、核酸試料は、ハイブリダイゼーション、例えば、アレイ上に存在する相補的なオリゴヌクレオチドとのハイブリダイゼーションの前に増幅することができる。定量的結果が所望される場合、増幅した核酸の相対頻度を維持または制御する方法を活用する。「定量的」増幅法は周知である。例えば、定量的PCRは、同じプライマーを使用して、既知量の対照配列を同時に共増幅することを伴う。これにより、PCR反応を校正するのに使用されうる内部基準がもたらされる。この場合、アレイには、増幅される核酸を定量化するための内部基準に特異的なプローブを含ませることができる。

【0131】

いくつかの例では、増幅のために使用されるプライマーを、対象の遺伝子産物(表3、

10

20

30

40

50

4、11、および/または13のうちのいずれかに示される遺伝子のRNAなど)の固有のセグメントを増幅するように選択する。他の実施形態では、増幅のために使用されるプライマーを、対象の遺伝子産物(表3、4、11、および/または13のうちのいずれかに示される遺伝子のRNAなど)に特異的なNPPを増幅するように選択する。可変遺伝子産物(例えば、表4、11、および/または13のうちのいずれかに示される)のほか、正規化遺伝子産物(例えば、表3を参照されたい)を増幅するのに使用されるプライマーは、市販されているか、または周知の方法に従いデザインすおよび合成することができる。

#### 【0132】

一例では、RT-PCRを使用して、メラニン細胞含有組織試料(例えば、皮膚生検)中のRNA(例えば、mRNA、miRNAまたはlncRNA)レベルを検出することができる。一般に、RT-PCRによる遺伝子発現プロファイリングの第1のステップは、RNA鋳型のcDNAへの逆転写に続く、PCR反応におけるその指数関数的増幅である。2つの一般に使用される逆転写酵素は、トリ骨髄芽球症ウイルス逆転写酵素(AMV-RT)およびモロニー Maus 白血病ウイルス逆転写酵素(MMLV-RT)である。逆転写ステップは典型的に、発現プロファイリングの状況および目的に応じて、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する。

10

#### 【0133】

PCRでは、様々な熱安定性DNA依存型DNAポリメラーゼを使用しうるが、典型的には、Taq DNAポリメラーゼを用いる。TaqMan(登録商標)PCRでは典型的に、その標的の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、TaqポリメラーゼまたはTthポリメラーゼの5'-ヌクレアーゼ活性を活用するが、同等な5'-ヌクレアーゼ活性を有する任意の酵素も使用することができる。2つのオリゴヌクレオチドプライマーを使用して、PCR反応に典型的なアンプリコンを生成する。第3のオリゴヌクレオチドまたはプローブは、2つのPCRプライマーの間に位置するヌクレオチド配列を検出するようにデザインする。プローブは、Taq DNAポリメラーゼ酵素により伸長不可能であり、レポーター蛍光色素およびクエンチャー蛍光色素で標識される。2つの色素が、プローブ上において、互いに近接して位置する場合、レポーター色素からの任意のレーザー誘導発光は、消光色素により消光される。増幅反応中に、Taq DNAポリメラーゼ酵素は、プローブを、鋳型依存的な様式で切断する。結果として得られるプローブ断片は、溶液中に解離し、放出されたレポーター色素からのシグナルは、第2のフルオロフォアの消光効果を免れる。合成される各々の新たな分子のために、レポーター色素の1つの分子が自由になり(liberalized)、消光されないレポーター色素を検出することにより、データを定量的に解釈するための基礎がもたらされる。

20

30

#### 【0134】

RT-PCRの変化形は、二重標識された蛍光性プローブ(例えば、Taqman(登録商標)プローブ)を介してPCR産物の蓄積を測定するリアルタイム定量的RT-PCRである。リアルタイムPCRは、各標的配列に対する内部コンペティターを正規化のために使用する定量的競合PCR、および試料中に含有される正規化遺伝子またはRT-PCRのための正規化遺伝子を使用する定量的比較PCRのいずれにも適合する(Heidら、Genome Research、6巻:986~994頁、1996年を参照されたい)。定量的PCRはまた、米国特許第5,538,848号においても記載されている。関連するプローブおよび定量的増幅手順は、米国特許第5,716,784号および米国特許第5,723,591号において記載されている。マイクロ滴定プレート内で定量的PCRを実行するための機器は、例えば、PE Applied Biosystems(Foster City, CA)から入手可能である。

40

#### 【0135】

代替的な定量的核酸増幅手順は、米国特許第5,219,727号において記載されている。この方法では、標的配列および内部基準である核酸セグメントを同時に増幅するこ

50

とにより、試料中の該標的配列（例えば、表4、11、および/または13のうちのいずれかに列挙される遺伝子の発現産物）の量を決定する。各セグメントから増幅される核酸の量を決定し、検量線と比較して、増幅する前に試料中に存在した標的核酸セグメントの量を決定する。

#### 【0136】

##### RNAシーケンシング

RNAシーケンシングは、マルチプレックス化された遺伝子発現情報を得るための別の方法を提供し、いくつかの実施形態では、ハイスループットの遺伝子発現情報を得るための別の方法を提供する。当技術分野では、RNAシーケンシングの多数の特定の方法が公知であり、かつ/または開発されている（1つの総論として、ChuおよびCorey、*Nuc. Acid Therapeutics*、22巻：271頁（2012年）を参照されたい）。全トランスクリプトームシーケンシング技法およびターゲティングRNAシーケンシング技法の各々が利用可能であり、開示される方法においても有用である。シーケンシングベースの遺伝子発現解析のための代表的な方法としては、遺伝子発現の連続解析（SAGE）、大規模並列処理特徴配列決定法（massively parallel signature sequencing）（MPSS）による遺伝子発現解析、全トランスクリプトームショットガンシーケンシング（別名、WTSSまたはRNA-Seq）、またはヌクレアーゼ保護シーケンシング（別名、qNPS、またはNPSeq；PCT公開第WO2012/151111号を参照されたい）が挙げられる。

#### 【0137】

##### 遺伝子発現を検出するためのタンパク質

本開示される方法のいくつかの実施形態では、メラニン細胞含有試料（例えば、皮膚生検）中の遺伝子発現のレベルを決定するステップは、試料中の1または複数のタンパク質を検出する（例えば、このようなタンパク質の相対量または実際量を決定することにより）ことを含む。当技術分野では、タンパク質を検出する日常的な方法が公知であり、本開示は、タンパク質検出の特定の方法に限定されない。

#### 【0138】

タンパク質の遺伝子産物（例えば、表4および/または11のうちのいずれかにおけるタンパク質の遺伝子産物）または正規化タンパク質（例えば、表3の正規化タンパク質）を検出することができ、試料中のタンパク質発現のレベルは、標的タンパク質（表3、4、および/または11のうちのいずれかの標的タンパク質など）に特異的なタンパク質特異的結合剤（抗体またはアプタマーなど）であって、ELISAアッセイなどのイムノアッセイ、イムノプロットアッセイ、フローサイトメトリーアッセイ、免疫組織化学アッセイ、酵素免疫アッセイ、ラジオイムノアッセイ、ウェスタンブロットアッセイ、免疫蛍光アッセイ、化学発光アッセイ、および他のペプチド検出戦略において使用されるタンパク質特異的結合剤により認識される新規のエピトープを介して決定することができる（Wongら、*Cancer Res.*、46巻：6029～6033頁、1986年；Luworら、*Cancer Res.*、61巻：5355～5361頁、2001年；Mishimaら、*Cancer Res.*、61巻：5349～5354頁、2001年；Ijazら、*J. Med. Virol.*、63巻：210～216頁、2001年）。一般に、これらの方法では、モノクローナル抗体またはポリクローナル抗体を使用する。

#### 【0139】

したがって、いくつかの実施形態では、検出可能に標識しうるその断片の抗体またはアプタマーなどの標的タンパク質特異的結合剤を使用して、生物学的試料に存在する標的タンパク質（表3、4、および/または11のうちのいずれかの標的タンパク質など）の発現レベルを検出し、したがって、発現したタンパク質の量を検出する。いくつかの実施形態では、特異的な結合剤は、ポリクローナル抗体またはモノクローナル抗体などの抗体であって、標的タンパク質（表3、4、および/または11のうちのいずれかの標的タンパク質など）に特異的に結合する抗体である。したがって、ある種の実施形態では、生物学的試料中のタンパク質のレベルまたは量を決定するステップは、被験体に由来する試料を

10

20

30

40

50

、タンパク質特異的結合剤（表3、4、および/または11のうちのいずれかに示されるタンパク質に特異的に結合する抗体など）と接触させること、結合剤に試料が結合するかどうかを検出すること、および、これにより、試料中に存在するタンパク質の量を測定することを包含する。一実施形態では、特異的な結合剤は、標的タンパク質（表3、4、および/または11のうちのいずれかの標的タンパク質など）に特異的に結合するモノクローナル抗体またはポリクローナル抗体である。当業者は、表3、4、および/または11のうちのいずれかの標的タンパク質などの標的タンパク質に対する抗体の商業的供給元が存在することを十分に理解する。

#### 【0140】

標的タンパク質（表3、4、および/または11のうちのいずれかの標的タンパク質など）の存在は、1つ、2つ、3つ以上の特異的結合剤など、複数の特異的結合剤により検出することができる。したがって、本方法では、複数の抗体を使用することができる。いくつかの実施形態では、抗体のうちの1つを、例えば、マルチウェルプレート（マイクロ滴定プレートなど）、ビーズ、膜などの固体支持体へと結合させる。実際、マイクロ滴定プレートは、固相として都合よく活用することができる。しかし、抗体反応はまた、液中中で行うこともできる。

#### 【0141】

いくつかの例では、方法は、試料を、標的タンパク質（表3、4、および/または11のうちのいずれかの標的タンパク質など）に特異的に結合する第一抗体に特異的に結合する第二抗体と接触させることを含む。いくつかの例では、第二抗体を、例えば、フルオロフォア（例えば、FITC、PE、蛍光タンパク質など）、酵素（HRPなど）、放射性標識、またはナノ粒子（金粒子または量子ドット（QDOT（登録商標））などの半導体ナノ結晶など）で検出可能に標識する。この方法では、抗体に結合させた酵素が、発色性基質など、適切な基質と、例えば、分光光度的手段、蛍光測定的手段により、または目視的手段により検出する化学的部分を生成させるような様式で反応する。抗体を検出可能に標識するのに使用する酵素としては、リンゴ酸デヒドロゲナーゼ、ブドウ球菌ヌクレアーゼ、デルタ-5-ステロイドイソメラーゼ、酵母アルコールデヒドロゲナーゼ、アルファ-グリセロリン酸デヒドロゲナーゼ、トリオースリン酸イソメラーゼ、西洋ワサビペルオキシダーゼ、アルカリホスファターゼ、アスパラギナーゼ、グルコースオキシダーゼ、ベータ-ガラクトシダーゼ、リボヌクレアーゼ、ウレアーゼ、カタラーゼ、グルコース-6-リン酸デヒドロゲナーゼ、グルコアミラーゼ、およびアセチルコリンエステラーゼが挙げられるがこれらに限定されない。検出は、酵素のための発色性基質を用いる比色法により達成することができる。

#### 【0142】

検出はまた、基質についての酵素反応の程度を、同様に調製された標準物質と比較して、目視比較することにより達成することもできる。抗体を、蛍光化合物で標識することもまた可能である。例示的な蛍光標識化合物としては、イソチオシアン酸フルオレセイン、ローダミン、フィコエリトリン、フィコシアニン、アロフィコシアニン、o-フタルアルデヒド、Cy3、Cy5、Cy7、イソチオシアン酸テトラメチルローダミン、フィコエリトリン、アロフィコシアニン、Texas Red、およびフルオレスカミンが挙げられる。抗体はまた、 $^{152}\text{Eu}$ またはランタニド系列の他の金属などの蛍光発光金属を使用して、検出可能に標識することもできる。抗体にコンジュゲートさせうる他の金属化合物は、フェリチン、超常磁性磁気ビーズコロイドなどの金コロイドが挙げられるがこれらに限定されない。これらの金属としては、ジエチレントリアミン五酢酸（DTPA）またはエチレンジアミン四酢酸（EDTA）などの金属キレート基を使用して、抗体へと結合させることができる。抗体はまた、化学発光化合物へとカップリングすることにより、検出可能に標識することもできる。化学発光標識化合物の例は、ルミノール、イソルミノール、テロマティックアクリジニウムエステル（theromatic acridinium ester）、イミダゾール、アクリジニウム塩、およびシュウ酸エステルである。同様に、生物発光化合物を使用して、抗体を標識することもできる。一例では、抗体は

10

20

30

40

50

、ルシフェリン、ルシフェラーゼ、またはエクオリンなどの生物発光化合物で標識する。抗体にコンジュゲートさせうるハプテンとしては、ビオチン、ジゴキシゲニン、オキサザロン、およびニトロフェノールが挙げられるがこれらに限定されない。抗体にコンジュゲートさせうるかまたは抗体に組み込みうる放射性化合物としては、テクネシウム $^{99m}$  ( $^{99m}\text{Tc}$ )、 $^{125}\text{I}$ 、ならびに $^{14}\text{C}$ 、 $^3\text{H}$ 、および $^{35}\text{S}$ が挙げられるがこれらに限定されない、任意の放射性核種 (radio nucleotide) を含むアミノ酸が挙げられるがこれらに限定されない。

#### 【0143】

一般に、タンパク質 (表3、4、および/または11のうちのいずれかのタンパク質など) についてのイムノアッセイは、典型的には、生物学的試料を抗体の存在下でインキュベートすること、および、結合した抗体を、当技術分野において周知のいくつかの技法のうちいずれかにより検出することを含む。一例では、生物学的試料 (メラニン細胞を含有する生物学的試料など) を、ニトロセルロースもしくはマルチウェルプレート、または細胞、細胞粒子、もしくは可溶性タンパク質を固定化することが可能な他の固体支持体などの固相支持体またはキャリアと接触させること、および、これに固定化することができる。次いで、支持体を、適切な緩衝液で洗浄した後、標的タンパク質 (表3、4、および/または11のうちのいずれかの標的タンパク質など) に特異的に結合する抗体で処理することができる。次いで、固相支持体を、緩衝液で再度洗浄して、結合しなかった抗体を除去することができる。抗体を直接標識する場合は、固体支持体上に結合した標識の量を、従来の手段により検出することができる。抗体を標識しない場合は、標識された第二抗体であって、標的タンパク質 (表3、4、および/または11のうちのいずれかの標的タンパク質など) に特異的に結合する抗体を検出する第二抗体を使用することができる。

#### 【0144】

代替的に、抗体を固体支持体へと固定化し、次いで、皮膚または眼に由来する組織生検など、生物学的試料から単離されたタンパク質と、抗体とタンパク質とが互いに特異的に結合することを可能とする条件下で接触させる。次いで、結果として得られる抗体：タンパク質複合体を、例えば、タンパク質に特異的な別の抗体を添加することにより (したがって、抗体：タンパク質：抗体サンドイッチを形成することにより) 検出することができる。添加される第二抗体を標識する場合、複合体を検出することもでき、または、代替的に、添加される第二抗体に特異的な標識二次抗体 (labeled secondary antibody) を使用することもできる。

#### 【0145】

固相支持体またはキャリアは、試料である抗原または抗体に結合することが可能な材料を含む。例示的な支持体としては、ガラス、ポリスチレン、ポリプロピレン、ポリエチレン、デキストラン、ナイロン、アミラーゼ、天然セルロースおよび変性セルロース、ポリアクリルアミド、斑糲岩、ならびに磁鉄鉱が挙げられる。キャリアの性質は、ある程度可溶性の場合もあり、不溶性の場合もある。支持体の材料は、カップリングさせた分子がその標的 (抗体またはタンパク質など) に結合することが可能な限りにおいて、実質的に任意の可能な構造形態を取りうる。したがって、支持体の形状は、ビーズにおけるような球形の場合もあり、試験管の内表面または棒の外表面におけるような円筒形の場合もある。代替的に、表面は、シートまたは検査ストリップなど、平らであってもよい。

#### 【0146】

一実施形態では、酵素結合免疫吸着測定法 (ELISA) を活用して、標的タンパク質を検出する (例えば、Voller, 「The Enzyme Linked Immunosorbent Assay (ELISA)」, Diagnostic Horizons, 2巻: 1~7頁、1978年を参照されたい)。例えば、標的タンパク質 (表3、4、および/または11のうちのいずれかの標的タンパク質など) に特異的に結合する抗体を使用することにより試料中のタンパク質の存在を検出するのに、ELISAを使用することができる。いくつかの例では、抗体は、例えば、直接コンジュゲートして、または二次抗体を介して酵素に連結することができ、酵素により検出可能なシグナルへと転

10

20

30

40

50

換しうる物質を添加する。

【0147】

検出はまた、様々な他のイムノアッセイのうちのいずれかを使用して、例えば、抗体または抗体断片に放射性標識することにより達成することもできる。別の例では、高感度で特異的なタンデム免疫放射測定アッセイを使用することもできる (ShenおよびTai、J. Biol. Chem.、261巻:25号、11585~11591頁、1986年を参照されたい)。放射性同位元素は、ガンマカウンターまたはシンチレーションカウンターの使用などの手段により検出することもでき、オートラジオグラフィーにより検出することもできる。

【0148】

一例では、分光法を活用して、標的タンパク質 (表3、4、および/または11のうちのいずれかの標的タンパク質など) の発現レベルを検出または定量化する。例示的な分光法としては、質量分析、核磁気共鳴分光法、およびこれらの組合せが挙げられる。一例では、質量分析を使用して、皮膚生検など、メラニン細胞含有試料中の標的タンパク質 (表3、4、および/または11のうちのいずれかの標的タンパク質など) の存在を検出する (例えば、Stemmannら、Cell、107巻(6号):715~26頁、2001年を参照されたい)。

【0149】

標的タンパク質 (表3、4、および/または11のうちのいずれかの標的タンパク質など) はまた、二次元ポリアクリルアミドゲル電気泳動 (2D-PAGE) により分離されるタンパク質の配列タグの、マトリックス支援レーザー脱離/イオン化飛行時間 (MALDI-TOF) 質量マッピングおよび液体クロマトグラフィー/四重極飛行時間エレクトロスプレーイオン化タンデム質量分析 (LC/Q-TOF-ESI-MS/MS) の使用による、イムノアフィニティー (immunoaffinity) アッセイにカップリングさせた質量分析アッセイにより検出することもできる (Kiernanら、Anal. Biochem.、301巻:49~56頁、2002年)。

【0150】

SELDIなどの定量的質量分析法を使用して、皮膚生検など、メラニン細胞含有試料中のタンパク質発現を解析することができる。一例では、例えば、ProteinChip (Ciphergen Biosystems, Palo Alto, CA) を使用することによりタンパク質発現を検出するのに、表面増強レーザー脱離イオン化飛行時間 (SELDI-TOF) 質量分析を使用する。当技術分野では、このような方法が周知である (例えば、米国特許第5,719,060号;同第6,897,072号;および同第6,881,586号を参照されたい)。略述すると、SELDIの1つの変化形では、表3、4、および/または11のうちのいずれかの分析物など、対象の分析物を選択的に捕捉する化学反応と共に、クロマトグラフィー表面を使用する。

【0151】

任意選択のアッセイ対照尺度

必要に応じて、遺伝子発現産物 (例えば、核酸 (mRNA、miRNA、lncRNA など) またはタンパク質) を検出するのに使用されるアッセイは、アッセイの性能を評価するのに使用される陽性および陰性の工程対照要素 (process control element) の両方を有する。

【0152】

陽性対照は、任意の既知の要素、好ましくは、性質が標的と類似する要素 (例えば、RNA標的であれば、RNA (またはcDNA) 陽性対照) であって、アッセイ (または試料) 中に含ませ、標的 (複数可) と並行して検出することができ、このような標的 (複数可) の検出に干渉 (例えば、交差反応) しない要素でありうる。一例では、陽性対照は、別個の試料として並行して実施されるかまたは各試料へと既知量で「スパイク」される、in vitro転写物 (IVT) である。IVT特異的結合剤 (例えば、ヌクレアーゼ保護プローブなどのオリゴヌクレオチドプローブ) および、該当する場合、IVT特異的

10

20

30

40

50

検出剤もまた、各アッセイ中に含めることで、このような *in vitro* 転写物についての陽性結果を確認する。別の例では、I V T 転写物は、*Methanobacterium* 属種 AL - 21 染色体 (NC\_015216) の非交差反応領域からデザインすることができる。

#### 【0153】

陰性工程対照要素は、遺伝子産物を検出するようにデザインまたは選択された分析物特異的結合剤（例えば、オリゴヌクレオチドまたは抗体）であって、該当する被験試料中では発現しないことが予測される分析物特異的結合剤を含みうる。例えば、ヒトトランスクリプトームまたはヒトプロテオームにおける任意の遺伝子発現産物を認識する分析物特異的結合剤（ヒト遺伝子発現産物が所望の標的である場合は、植物または昆虫または線虫それぞれの RNA またはタンパク質に特異的なオリゴヌクレオチドプローブまたは抗体など）を、マルチプレックス化されたアッセイに含ませることができる。この陰性対照要素は、該当するアッセイにおいてシグナルを発生させないものとする。このような陰性工程対照要素についての、任意のバックグラウンドを上回るシグナルは、アッセイの不成功の指標である。一例では、陰性対照は、ANT である。

#### 【0154】

遺伝子発現は、生物学的特質に起因して、かつ/または検体の安定性、検体の完全性、もしくは入力レベルのほか、アッセイ工程およびアッセイシステムとも関連する変動に起因して、試料型または被験体にわたり変化しうる。特定の実施形態では、変動の非生物学的関連ソース（とりわけ、マルチプレックス化されたアッセイにおいて）を最小化するために、対象の試料間で、著明に変化しないかまたはバイオインフォマティクス法により著明に変化しないことが見出されている遺伝子発現産物（例えば、「ハウスキーピング遺伝子 (housekeeper)」または正規子) を測定する。いくつかのこのような実施形態では、候補正規化遺伝子産物の発現レベルは、十分であり（例えば、バックグラウンドを上回り）、かつ/または不飽和の強度値を裏付ける。正規子遺伝子発現産物についてのさらなる考察は、本開示の別の箇所で見出される。

#### 【0155】

いくつかの状況では、例えば、正規子を解析することにより、異常なシグナルが、工程以外では制御されない予測外の工程関連問題から生じる可能性があり、したがって、いくつかの実施形態では、検体の安定性、検体の完全性、または入力レベルに関わらず、任意の検体についてのアッセイの成功または不成功を指し示すために、試料に依存しない工程対照要素（複数可）を含めることが有用である。核酸遺伝子発現産物を検出する方法の実施形態は、あらゆるアッセイにおいて既知濃度の RNA 試料（例えば、*in vitro* 転写物 RNA または I V T）を含みうる。このような対照要素（例えば、I V T）は、各アッセイにおいて測定され、アッセイ工程の品質対照として働く。

#### 【0156】

MAQC (マイクロアレイ品質管理 (Microarray Array Quality Control)) プロジェクトは、「Universal Human Reference RNA」が、マイクロアレイ遺伝子発現アッセイのための有用な外部対照基準でありうることを提起した。したがって、RNA 遺伝子発現産物を伴ういくつかの開示される方法の実施形態は、Universal Human Reference RNA を含有する並行加工試料を含みうるが、含む必要がない場合もある。このようなユニバーサル RNA 試料は、該当するアッセイによる検出の標的とされる RNA のうちの全部または一部を含む場合は、含められたこのような RNA についての陽性シグナルを予測することができ、これも、1つの（または別の）アッセイ工程品質管理として働きうる。

#### 【0157】

遺伝子発現データ

遺伝子発現データは、「疾患の防止および治療、生物学的進化の機構、ならびに薬物の発見に関連する基本的な問題に取り組むための鍵を含有する」ことが十分に受容されている

10

20

30

40

50

(LuおよびHan、Information Systems、28巻：243～268頁(2003年))。いくつかの例では、このようなデータから情報を抜き出すことは、1または複数の検出される遺伝子産物の存在、非存在、または定性量(例えば、高度、中程度、低度)から、定性的な決定を下すことと同じく単純である。他の例では、生の遺伝子発現データを、前加工する(例えば、バックグラウンドを減算する、対数変換する、かつ/または補正する)場合もあり、正規化する場合もあり、かつ/または分類アルゴリズムにかける場合もある。これらの側面については、下記でより詳細に記載する。

#### 【0158】

##### データの前加工

##### バックグラウンド減算

いくつかの方法の実施形態では、生の遺伝子発現データは、バックグラウンド減算する。この補正は、例えば、データを、マイクロアレイなど、マルチプレックス化された方法を使用して収集した場合に使用することができる。このような変換の1つの目的は、例えば、マイクロアレイ表面の1つの部分が、いかなる生物学的理由もなく、表面の別の部分より「明るく」見える場合に、局所的効果について補正することである。当技術分野では、バックグラウンド減算の方法が周知であり、例えば、(i)局所的バックグラウンド減算(例えば、スポットマスクの外側であるが、スポット中心を中心とするバウンディングボックス内にある全てのピクセルについて考慮する)、(ii)モルフォロジーオープニング(morphological opening)によるバックグラウンドの推定(元の画像から減算するためのバックグラウンド画像を作成するには、オープニングフィルタ、エロージョンフィルタ、ダイレーションフィルタ、およびランクフィルタなど、非線形モルフォロジーフィルタ(Soille、Morphological Image Analysis: Principles and Applications、Berlin、Springer-Verlag(1999年)を参照されたい)に依拠する)、(iii)定常バックグラウンド(全てのスポットについて、定常バックグラウンドを減算する)、Normexpによるバックグラウンド補正(共変量としてのバックグラウンド強度を使用して、正規分布および指数分布のコンボリューションを、フォググラウンド強度に適合させると、観察されるフォググラウンドを所与としたときに予測されるシグナルが補正強度となる)を含む。

#### 【0159】

##### データの変換

多くの生物学的変数(例えば、遺伝子発現データ)は、統計学的パラメトリック検定の仮定を満たさず、例えば、このような変数は、正規分布しないか、等分散でないか、またはこれらの両方である(Durbinら、Bioinformatics、18巻：S105頁(2002年))。いくつかの場合には、データを変換することにより、統計学的仮定によりよく適合する。いくつかの方法の実施形態では、有用なデータの変換は、(i)各観察値の対数、例えば、底を10とする対数、底を2とする対数、底をeとする対数(自然対数としてもまた公知である)を取ることからなる対数変換であって、このような対数は定数因子だけ異なるため対数の選択は重要でない対数変換;または例えば、Durbin(上記)により記載されている分散安定化変換を含みうる。具体例では、このような方法において検出される各バイオマーカー(例えば、少なくとも2つの表4、11および/もしくは13のバイオマーカー、ならびに/または少なくとも1つの正規化バイオマーカー)についての生の発現値を $\log$ (例えば、 $\log_2$ または $\log_{10}$ )変換する。他の実施形態では、正規化するステップは、少なくとも2つの表4、11および/または13のバイオマーカーの各々の $\log$ (例えば、 $\log_2$ または $\log_{10}$ )変換された生の発現値を、少なくとも1つの正規化バイオマーカーの $\log$ (例えば、 $\log_2$ または $\log_{10}$ )変換された生の発現値(複数可)で除算することを含みうる。

#### 【0160】

##### データフィルタ

いくつかの方法の実施形態では、遺伝子発現データにフィルタをかけて、信頼できない

10

20

30

40

50

と考えられうるデータを除去することができる。当技術分野では、遺伝子発現データの信頼性を評価するための多くの方法が公知であり、以下の非限定的な例は、代表的な例であるに過ぎないことが理解される。

【0161】

遺伝子が発現しないか、または検出不可能な（バックグラウンドを上回らない）レベルで発現するいくつかの場合には、遺伝子発現データを、開示される方法から除外してよい。逆に、陰性対照（例えば、ANT）遺伝子の発現が、標準のカットオフを超える（例えば、100、200、250、もしくは300相対光単位を超えるか、またはバックグラウンドを1%、2%、3%、4%、5%、6%、7%、8%、9%、もしくは10%超上回る）いくつかの場合には、遺伝子発現データを、解析から除外してよい。

10

【0162】

プローブセットまたは遺伝子が関係する実施形態では、いくつかの特定のデータフィルタであって、

(i) 信頼できないプローブセットから生じるデータは、プローブセットの信頼性を、一連の参照データセットに対してランク付けすることにより、解析から除外するように選択しうる。例えば、RefSeqおよびEnsembl (EMBL) は、極めて高品質の参照データセットであると考えられる。いくつかの場合には、RefSeq配列もしくはEnsembl配列にマッチするプローブセットに由来するデータはとりわけ、それらの予測される高い信頼性のために、マイクロアレイ解析実験に含めることができる。同様に、信頼できる参照データセットにそれほどマッチしないプローブセットに由来するデータは、さらなる解析から除外してもよく、ケースバイケースで包含について考慮する場合もあること；または

20

(ii) 分散を呈示しないかまたは呈示する分散が小さなプローブセットは、さらなる解析から除外してよいこと。低分散プローブセットは、カイ二乗検定により解析から除外してよい。プローブセットは、その変換された分散が、自由度  $N - 1$  のカイ二乗分布の99パーセントの信頼区間の左側にある場合、低分散であると考えられること；または

(iii) 所与の遺伝子クラスターもしくは転写物クラスターについてのプローブセットは、例えば、他のデータ前加工ステップの後で、それらが最小数未満のプローブを含有する場合、さらなる解析から除外してよいこと。例えば、いくつかの実施形態では、所与の遺伝子クラスターもしくは転写物クラスターについてのプローブセットは、それらが1つ、2つ、3つ、4つ、または5つ未満のプローブを含有する場合、さらなる解析から除外してよいこと

30

を含むデータフィルタが有用でありうる。

【0163】

必要に応じて、いくつかのレプリケートのうちの1つのデータが、平均値から「 $x$ 」標準偏差 (SD) (例えば、少なくとも2SDまたは少なくとも3SD) 隔たっているか、またはレプリケート間のCV%が指定された量 (例えば、対数変換空間内で少なくとも8%) を超えることにより判定されるなど、他のデータと比較して統計学的に外れ値であるのかどうかを決定する、統計学的外れ値プログラムを使用することができる。アレイベースのアッセイでは、外れ値は、アレイスポットのうちの1つに問題があることから生じる場合もあり、画像化のアーチファクトに起因する場合もある。外れ値の除去は、遺伝子ごとに基づいて実施することが典型的であり、1つのレプリケートにおける遺伝子の大半が外れ値である場合は、あらかじめ確立された規則であって、レプリケートの全体を除外する規則を適用することができる。例えば、極めて重要な試薬の不適正な添加を結果としてもたらずピペティング誤差であれば、レプリケートの全体を外れ値としうる。

40

【0164】

遺伝子発現を、試料レプリケート (例えば、三連) で測定するいくつかの例では、ペアワイズ相関のより、および試料のペアワイズ線形回帰により、再現性を測定することができる。相関  $r = 0.95$  を、レプリケート (例えば、三連) における再現性の採択域として使用することができる。より具体的な例では、ペアワイズ相関  $r = 0.90$  であるレプリ

50

ケートは、単純回帰モデルによりさらに精査することができるが、ここで、線形回帰の切片が、統計学的に有意にゼロと異なる場合は、レプリケートをさらなる検討から除外する。不成功のレプリケートが25%（例えば、4例中1例）以上、33%（例えば、3例中1例）以上、50%（例えば、4例中2例）以上、または67%（例えば、3例中2例）以上を上回る任意の試料は、「不成功試料」と考えることができ、さらなる解析から除外することができる。

#### 【0165】

##### 正規化

正規化の目的は、生物学的効果に起因する変動を、観察および定量化しうるように、実験誤差に起因する（例えば、ピペティング、プレートにおける位置、画像のアーチファクト、全RNA量の差違などに起因する）変動を除去することである。この工程は、異なる試料型の間で観察される差違が、試料の生物学的特質の差違に真に起因するものであり、何らかの技術アーチファクトに起因するものではないことを確認する一助となる。実験中には、誤差が導入されうるいくつかの点が存在するが、これは、正規化により除外することができる。当技術分野では、遺伝子発現データを正規化するための方法が十分に確立されている（例えば、Methods in Microarray Normalization、Phillip Stafford編、Baton Rouge、FL: CRC Press、Taylor & Francis Groupにより刊行中、2008年）。

#### 【0166】

正規化は典型的に、1または複数の表4、11および/または13のバイオマーカーの発現値などの実験値を、1または複数の正規化値または正規化係数と比較する（例えば、除算すること（または典型的に、対数変換の後で減算すること）により）ことを伴う。正規化値は、単一の正規化バイオマーカーの生の（または対数変換された）発現値の場合もあり、例えば、複数の正規子の発現値から計算するか、または当技術分野で公知の方法および計算を使用して計算する場合もある。いくつかの例では、正規化では、複数の正規化バイオマーカーの発現の平均値を使用して、検定される各々の表4、11および/または13のバイオマーカーについての正規化された発現値を生成する。いくつかの例では、正規化では、検定される表4、11および/または13のバイオマーカーの各々についての生の発現値、および少なくとも1つの表3の正規化マーカーについての生の発現値を使用して、検定される各々の表4、11および/または13のバイオマーカーの正規化された発現値を生成する。

#### 【0167】

いくつかの実施形態では、表3の1または複数の正規化バイオマーカーなど、1または複数の「正規化バイオマーカー」の発現を、決定または測定することができる。例えば、被験試料中のBMP-1、MFI2、NCOR2、RAP2b、RPS6KB2、SDHA、RPL19、RPLP0、およびALDOAのうちの1つ、2つ、3つ、4つ、5つ、6つ、7つ、8つ、または全ての発現を検出することができる。

#### 【0168】

代替的に、開示される方法において有用な1または複数の正規化バイオマーカーは、本明細書で提示される方法を使用して同定することができる。例えば、正規化バイオマーカーとは、任意の構成的に発現した遺伝子（またはタンパク質）であって、その発現に対して、別の発現した遺伝子（またはタンパク質）を比較（例えば、一方の遺伝子（またはタンパク質）の発現を、他方の遺伝子（またはタンパク質）の発現で除算すること（または典型的に、対数変換の後で減算すること）により）しうる遺伝子（またはタンパク質）である。他の例では、正規化バイオマーカーは、任意の遺伝子発現産物（例えば、mRNA、miRNA、またはタンパク質）であって、その発現が、母斑試料および黒色腫試料など、代表的な複数の試料にわたり有意に異ならない遺伝子発現産物でありうる。したがって、いくつかの方法では、正規化バイオマーカーは、表4、11、および/または13で列挙されていない任意の遺伝子発現産物であって、その発現が、メラニン細胞含有試料（

例えば、母斑試料および黒色腫試料の代表的な集団)の間で有意に異なる遺伝子発現産物でありうる。他の例では、少なくとも1つの正規化バイオマーカーは、複数の正規化バイオマーカーであって、それらの発現のいずれもが母斑試料と原発性黒色腫試料との間で統計学的に有意に異なる正規化バイオマーカーを含みうる。

【0169】

開示される方法において有用な正規化バイオマーカーを同定する別の方途は、生のデータを比較する場合に、推定正規子の発現が、互いに追跡する(すなわち、一方の正規化バイオマーカーが増大すれば、他方の正規化バイオマーカーも同様に増大するのか)のかどうかを決定することである。有用な正規子は、複数の対象の試料にわたり互いに追跡する。また、推定正規化バイオマーカーの間の比も決定することができ、それらの間の比が、

10

【0170】

いくつかの方法の実施形態は、例えば、本開示で記載されるとおり、正規化バイオマーカーを同定したら、表4、11、および/または13の少なくとも2つのバイオマーカーの各々についての生の(または対数変換された)発現値を、少なくとも1つの正規化バイオマーカーについての生の(または対数変換された)発現値に対して正規化するステップを含む。

【0171】

代替的に、正規化値を決定し、このような値を使用して、実験値(例えば、表4、11、および/または13からの少なくとも2つの異なるバイオマーカーの遺伝子発現値)を正規化することもできる。いくつかの開示される方法では、例えば、それらの発現値の範囲および分布が、対象の試料(複数可)(例えば、母斑試料および/または黒色腫試料などのメラニン細胞含有試料)のトランスクリプトームにおける遺伝子集団の発現の範囲および分布を表す、複数のバイオマーカーの集団CT(例えば、平均(算術平均または幾何平均など)、中央値、最頻値、または平均値)は、正規化値として働きうる。他の例では、複数のバイオマーカーにおける外れ値(例えば、集団CTから $\pm 1$ または2標準偏差)の発現値が、複数のバイオマーカーの集団CTについての元の計算から除外され、外れ値を含まない集団CTが、複数のバイオマーカーについて決定され、実験変数(例えば、表4、11、および/または13における少なくとも2つの遺伝子についての遺伝子発現値)

20

30

【0172】

他の具体例では、頑健なマルチアレイ平均(RMA)法を使用して、生のデータを正規化することができる。RMA法は、多くのマイクロアレイ上の各々のマッチした細胞についてのバックグラウンド補正強度を計算することにより始める。バックグラウンド補正值は、Irisarryら(Biostatistics、4巻:249頁(2003年))により記載されている通り、正の値に制限される。次いで、バックグラウンド補正の後、各々のバックグラウンド補正されたマッチした細胞の強度の、底を2とする対数を得る。次いで、各マイクロアレイ上の、バックグラウンド補正され、対数変換された、マッチした強度を、分位数正規化法であって、各入力アレイおよび各プローブの発現値について、アレイのパーセンタイルプローブ値を、全アレイのパーセンタイル平均値で置きかえる正規化法(この方法は、Bolstadら(Bioinformatics、19巻(2号):185頁(2003年))により、より完全に記載されている)を使用して正規化する。次いで、分位数正規化の後、正規化されたデータを、線形モデルへと適合させて、各マイクロアレイ上の各プローブについての発現尺度を得ることができる。

40

【0173】

いくつかの例では、第1の正規化は、処置内のレプリケートにわたる正規化の場合もあり、技法についてのレプリケート内の正規化の場合もある。第1の正規化とは、検定される全てのバイオマーカー(表4、11、および/または13のバイオマーカーのうち2つ以上など)に対する、そのレプリケートのセットについての全シグナルの定常レベルに

50

対して重み付けされた正規化である。このステップでは、レプリケートのセット内の各アッセイ（ウェルまたはビーズまたはレーンなど）についての全シグナル強度を、全てが等しくなるように調整する。全てのレプリケートについて全平均値シグナルを計算し、次いで、各試料についての正規化係数であって、そのレプリケートに由来する（form）全シグナルを、全てのレプリケートについての全平均値シグナルに対して調整する正規化係数を計算する。次いで、この正規化係数を使用して、そのレプリケート内の各バイオマーカーについてのシグナルを正規化する。

#### 【0174】

##### 特徴選択（FS）

分類アルゴリズムの性能は典型的に、数千の特徴（遺伝子/タンパク質）について最適未満である。したがって、特徴選択法（feature selection method）を使用して、表現型を最もよく予測する特徴を同定する。選択された遺伝子/タンパク質は、分類子または予測モデルへと提示する。特徴空間の次元の縮減からは、以下の利益：（i）分類精度さを改善する利益、（ii）データを生成した基底的概念のよりよい理解をもたらす利益、および（iii）特徴の数が多く、トレーニングパターンの方が比較的少ない場合に生じるデータの過剰適合の危険性を克服する利益、が得られる。特徴選択を使用して、開示される遺伝子セットを決定した（したがって、対応する分類子には、上記の利点が組み込まれている）。

#### 【0175】

特徴選択技法は、フィルタ技法（データの内在的特性を調査することにより、特徴の関与性を評価する）、ラッパー法（モデルの仮説を、特徴サブセット検索内に組み込む）、および埋込み技法（特徴の最適のセットについての検索を、分類子アルゴリズムへと組み込む）を含む。開示される方法において有用なフィルタFS技法は、（i）二標本t検定またはモデレートt検定（moderated t-test）（例えば、LIMMA）、ANOVA解析、ベイズフレームワーク、およびガンマ分布モデルの使用などのパラメトリック法、（ii）ウィルコクソン順位和検定、クラス間-クラス内二乗和検定、順位積法（rank products method）、ランダム順列法（random permutation method）、または2つのデータセットの間の発現の倍数変化の差違についての閾値点を設定し、次いで、誤分類（misclassification）の数を最小化する各遺伝子の閾値点を検出することを伴う誤分類総数（TNOM）の使用などのモデルフリー法、ならびに（iii）二変量法などの多変量法、相関ベースの特徴選択法（CFS）、最小冗長性最大関連性（minimum redundancy maximum relevance method）（MRMR）、マルコフブランケットフィルタ方法、ツリーベースの方法、および非相関収縮重心（uncorrelated shrunken centroid）法を含む。開示される方法において有用なラッパー法は、逐次検索法、遺伝子アルゴリズム、および分布アルゴリズムの推定を含む。本開示の方法において有用な埋込み法は、ランダムフォレスト（RF）アルゴリズム、サポートベクターマシンアルゴリズムの重みベクトル、およびロジスティック回帰アルゴリズムの重みを含む。Saey sらは、遺伝子発現解析における特徴選択について上記で提示したフィルタ技法の相対的な長所について記載している。いくつかの実施形態では、LIMMAソフトウェアパッケージ（Smyth、LIMMA: Linear Models for Microarray Data、Bioinformatics and Computational Biology Solutions、Gentlemanら編、New York、Springer、397~420頁（2005年））を使用することにより、特徴選択を提供する。

#### 【0176】

##### 分類子アルゴリズム

いくつかの方法では、遺伝子発現情報（例えば、表3、4、11、および/または13において記載されているバイオマーカーについての）を、アルゴリズムへと適用して、発現プロファイルを分類（例えば、メラニン細胞含有試料（皮膚生検など）が、良性の母斑

10

20

30

40

50

であるのか、原発性黒色腫であるのか、これらのいずれでもない（判定保留（ i n d e t e r m i n a n t ）など）のか）する。本明細書で開示される方法は、遺伝子発現ベースの分類子であって、メラニン細胞含有試料を、母斑または黒色腫として特徴付けるための分類子を含みうる。特定の分類子実施形態を、提供される遺伝子セットおよび分類法に基づき記載するが、他の分類子実施形態も今や可能である。

【 0 1 7 7 】

分類子とは、被験試料（例えば、メラニン細胞含有試料）を、このような試料中の遺伝子（表 4、1 1、および / または 1 3 の遺伝子など）の発現に基づき、クラス（または群）（例えば、母斑または黒色腫）へと分類するのに使用されうる予測モデル（例えば、アルゴリズムまたは規則のセット）である。クラスターの数があらかじめ知られていないクラスター解析とは異なり、分類子は、試料の 1 または複数のセットであって、所望のクラス値（複数可）（例えば、母斑または黒色腫）が既知であるセットについてトレーニングされている。トレーニングされたら、分類子を使用して、クラス値（複数可）をその後の観察値へと割り当てる。

10

【 0 1 7 8 】

開示される方法において有用な、例示的なアルゴリズムとしては、主成分分析アルゴリズム、部分最小二乗法、および独立成分分析アルゴリズムなど、変数の数を縮減する方法が挙げられるがこれらに限定されない。例示的なアルゴリズムとしては、統計学的方法および機械学習技法に基づく方法など、多数の変数を直接扱う方法もさらに挙げられるがこれらに限定されない。統計学的方法としては、罰金付きロジスティック回帰、マイクロアレイ予測解析（ P A M ）、収縮重心法に基づく方法、サポートベクターマシン解析、および正則化線形判別分析が挙げられる。機械学習技法としては、バギング手順、ブースティング手順、ランダムフォレストアルゴリズム、およびこれらの組合せが挙げられる。 B o u l e s t e i x ら（ C a n c e r I n f o r m . 、 6 巻： 7 7 頁（ 2 0 0 8 年 ） ） は、マルチプレックス化された遺伝子発現データの解析について上記で提示した分類法についての概要を提示している。

20

【 0 1 7 9 】

機械学習とは、コンピュータが、適応技術を使用してパターンを認識し、かつ行動を予測し、これにより、パターンを、莫大な量のデータを介して分取ならびに解析および同定する学習である。機械学習アルゴリズム（例えば、ロジスティック回帰（ L R ）アルゴリズム、ランダムフォレスト（ R F ）アルゴリズム、サポートベクターマシン（ S V M ）アルゴリズム、 K 近傍（ K N N ）アルゴリズム）は、アルゴリズムを手動でデザインするにはあまりに複雑な適用において、ソフトウェアを開発するのに有用でありうる。

30

【 0 1 8 0 】

いくつかの実施形態では、被験試料を、トレーニングされたアルゴリズムを使用して分類する。本開示のトレーニングされたアルゴリズムは、既知の母斑試料および黒色腫試料の参照セットを使用して開発されているアルゴリズムを含む。試料の類別に適するアルゴリズムとしては、 k 近傍アルゴリズム、概念ベクトルアルゴリズム、単純ベイズアルゴリズム、ニューラルネットワークアルゴリズム、隠れマルコフモデルアルゴリズム、遺伝子アルゴリズム、および相互情報量特徴選択アルゴリズム、または任意のこれらの組合せが挙げられるがこれらに限定されない。いくつかの場合には、本開示のトレーニングされたアルゴリズムには、本開示の、細胞学者または病理学者による評定または診断、開示される前分類子アルゴリズムもしくは遺伝子セットにより提供される情報、または被験試料が採取される被験体の病歴についての情報などであるがこれらに限定されない、遺伝子発現データ以外のデータも組み込むことができる。

40

【 0 1 8 1 】

いくつかの特定の実施形態では、サポートベクターマシン（ S V M ）アルゴリズム、ランダムフォレストアルゴリズム、またはこれらの組合せにより、試料（例えば、メラニン細胞含有試料）の、母斑または黒色腫（例えば、原発性黒色腫）および、必要に応じて、判定保留クラスへの分類が提供される。いくつかの実施形態では、同定されるマーカーで

50

あって、試料（例えば、母斑対黒色腫）を識別するマーカーを、統計学的有意性に基づき選択する。いくつかの場合には、統計学的有意性の選択は、偽発見率（FDR）についてのBenjamini Hochberg補正を適用した後で実施する（J. Royal Statistical Society, Series B (Methodological)、57巻：289頁（1995年）を参照されたい）。

#### 【0182】

いくつかの場合には、開示される分類子アルゴリズムは、Fishelら（Bioinformatics、23巻：1599頁（2007年））により記載されているメタ解析法などのメタ解析法で補足することができる。いくつかの場合には、分類子アルゴリズムを、再現性解析などのメタ解析法で補足することができる。いくつかの場合には、再現性解析により、少なくとも1つの予測発現産物のマーカーセットに現れるマーカーを選択する。

10

#### 【0183】

##### 例示的な決定木モデル

決定木アルゴリズムとは、各々の内部のノードは、属性についての検定を示し、ブランチは、検定のアウトカムを表す、フローチャート様のツリー構造である。リーフノードは、クラスラベルまたはクラス分布を表す。決定木を作成するには、全てのトレーニング例をルートにおいて使用し、ツリーのルートにおける論理検定を適用し、次いで、トレーニングデータを、論理検定の値に基づき、サブグループへと分割する。この工程を帰納的に適用し（すなわち、属性を選択し、分岐させ）、1本のブランチ内の全てのデータ要素が同じクラスとなったら終了する。未知の標本を分類するには、その属性値を決定木に対して検定する。

20

#### 【0184】

機械学習の一例として述べると、ランダムフォレストは、トレーニング時間において膨大な決定木を構築し、個々のツリーによるクラス出力の最頻値であるクラスを出力することにより作動する、分類（および回帰）のための集団による学習方法である。1つの特定のランダムフォレストアルゴリズム（Breiman, Machine Learning、45巻：5～32頁（2001年））では、各ツリーを、以下の通りに構築する。

1. トレーニング例数を、「N」とし、分類子内の変数の数を「M」とする。
2. 「m」とは、ツリーのノードにおいて決定を下すのに使用される入力変数の数であり、mは、M未満とする。
3. 置きかえを伴ってn回にわたり選択することにより、このツリーのトレーニングセットを、N例の利用可能なトレーニング例全てから選択する（すなわち、ブートストラップ標本を選ぶ）。残りのトレーニング例を使用して、それらのクラスを予測することにより、ツリーの誤差を推定する。
4. ツリーの各ノードについて、そのノードにおける決定に基づくm個の変数をランダムに選択する。トレーニングセット内のこれらのm個の変数に基づき、最良の分岐を計算する。
5. 各ツリーを十分に成長させ、剪定（正規ツリー分類子を構築するときになされる場合がある）はしない。

30

40

#### 【0185】

予測するには、新たな標本を、ツリーに沿って下方へと押し進める。標本には、それが最後には至る終端ノードにおいて、トレーニング標本のラベルを割り当てる。この手順を、集団内の全てのツリーにわたり繰り返し、全てのツリーの最頻度投票を、ランダムフォレストによるクラス予測として報告する。

#### 【0186】

##### 例示的なロジスティック回帰モデル

表4、11、および/または13の遺伝子を使用して、統計学的予測モデルを開発するための1つの代表的な方法は、二極分布（binary distribution）およびロジットリンク関数によるロジスティック回帰である。このようなモデルのための推

50

定は、フィッシャースコアリングを使用して実施することができる。しかし、正確ロジスティック回帰、経験的サンドイッチ推定量、または他の、バイアスを補正するか、分散を安定化させるか、もしくは他の形の補正推定技法により推定されるモデルもまた、多くの状況下で、類似のモデルであって、やや異なるパラメーターの推定値をもたらす一方で、定性的に一貫したパターンの結果をもたらすモデルを提供する。同様に、累積ロジット関数、相補的  $\log - \log$  関数、プロビット関数、または累積プロビット関数を含むがこれらに限定されない他のリンク関数も、同じ定性的パターンの結果を与える予測モデルをもたらすことが期待される。

予測モデル（アルゴリズム）の1つの代表的な形態は、

$$\text{Logit}(Y_i) = 0 + 1X_1 + 2X_2 + 3X_3 \cdots nX_n$$

[ 式中、 $0$  は、切片項であり、 $n$  は、係数の推定値であり、 $X_n$  は、所与の遺伝子についての発現値の対数（例えば、底を2とする対数、または底を10とする対数などの、任意の対数）である。典型的に、全ての  $0$  の値は、 $-1$ 、 $000$  を超え、かつ、 $1$ 、 $000$  未満となろう。切片項である  $0$  は、 $-200$  を超え、かつ、 $200$  未満となることが多く、 $-100$  を超え、かつ、 $100$  未満である例を伴う。さらなる  $n$  は、 $n > 0$  のとき、 $-100$  を超え、かつ、 $100$  未満でありうる ]

である。

#### 【0187】

特定の方法の実施形態では、 $\text{Logit}(Y_i)$  の出力を、少なくとも2つの表4、11、および/または13のバイオマーカーについての連結発現値（CEV: consolidated expression value）と称する。CEVは、(a) 少なくとも2つの表4、11、および/または13のバイオマーカーの発現レベルを、少なくとも2つの表4、11、および/または13のバイオマーカーの各々について所定の定数で重み付けすること、および、(b) 少なくとも2つの表4、11、および/または13のバイオマーカーの重み付けした発現レベルを組み合わせて、CEVを生成させることにより決定する。このような方法はまた、CEVを、既知の黒色腫（例えば、原発性黒色腫）試料を、既知の良性の母斑試料から識別する参照値と比較するステップも含みうる。一例では、方法は、CEVが、既知の黒色腫試料と、参照値の同じ側であれば、試料を悪性（例えば、原発性黒色腫）として特徴付けることをさらに含む。別の例では、方法は、CEVが、既知の良性の母斑試料と、参照値の同じ側であれば、試料を良性（例えば、母斑）として特徴付けることをさらに含む。

#### 【0188】

本明細書で想定される任意の予測モデルの性能は、当技術分野で公知のいくつかの検定であって、ワルドカイ二乗検定（全てのモデル適合にわたる）、およびホスマー-レメシヨウ不適合度検定（lack fit test）（モデルについて統計学的に検出可能な不適合度が見られない）を含むがこれらに限定されない検定により検証することができる。モデル内の各遺伝子についての予測子は、統計学的に有意（例えば、 $p < 0.05$ ）であるとする。

#### 【0189】

いくつかの交差検証法が、結果の再現性を確認するのに利用可能である。例示的な方法は、SAS Proc Logistic 分類表手順の一部として実装されるワンステップの最尤推定近似である。いくつかの例では、オープンソースパッケージである Weka における10倍の交差検証および66~33%の分岐検証を、結果の確認に使用することができる。他の例では、一個抜き（LOO）交差検証を含む  $n$  倍の交差検証および分岐標本トレーニング/検定により、結果の有用な確認がもたらされる。

#### 【0190】

いくつかの方法の実施形態では、アルゴリズム（別名、適合モデル）により、予測される事象の確率であって、例えば、メラニン細胞含有試料（例えば、皮膚生検）試料が、黒色腫（例えば、原発性黒色腫）である確率、悪性である確率、母斑である確率、または良性である確率がもたらされる。いくつかの場合には、当業者に公知のSAS計算法を使用

10

20

30

40

50

して、予測される確率のバイアス低減推定値を計算することができる (support.sas.com/document/cdl/en/statug/63347/HTML/default/viewer.htm#statug\_logistic sect 044.htm (2013年3月15日現在)を参照されたい)。他の例では、一連の閾値  $z$  [ここで、 $z$  は、 $0 \sim 1$ の間である] を、典型的には、モデルの所望の臨床的有用性または適用要件に基づき当業者が決定する通りに設定する。予測される確率であって、特定の試料について計算される確率が、あらかじめ設定された閾値である  $z$  を超えるかまたはこれに等しければ、試料を母斑群へと割り当て、そうでなければ、黒色腫群へと割り当てるか、またはこの逆を行った。他の例では、2つの閾値は、2つの閾値の間にある試料値が、「判定保留」または「割り当て不能 (not otherwise assigned)」などのラベルを割り当てられるセットでありうる。

10

## 【0191】

アルゴリズム出力に基づき、例えば、出力を、参照標準 (例えば、既知の悪性メラニン細胞含有試料および良性メラニン細胞含有試料から決定されるカットオフ) と比較することにより、被験試料 (例えば、皮膚試料) が悪性であるのか、良性であるのかについての決定を行う。いくつかの例では、出力をアルゴリズムから計算するステップ、および/または出力を参照標準と比較することにより、試料が悪性であるのか、悪性でないのかをアルゴリズム出力から決定するステップを、適切にプログラムされたコンピュータにより実施する。いくつかの例では、方法はまた、使用者に、アルゴリズム出力、または試料が悪性であるのか、悪性でないのか、もしくは「黒色腫と一致する」のか、「母斑と一致する」のか、「判定保留」であるのかなどの決定を含む報告を提供するステップも含みうる。いくつかの例では、報告は、解析される表4、11、および/または13からの少なくとも2つのバイオマーカーについてのCEVを含む。

20

## 【0192】

結果として得られる出力値は、カットオフ値と比較する。カットオフ値は、良性であるのか悪性であるのかがあらかじめ既知である複数のメラニン細胞含有試料中の、表4、11、および/または13からの、少なくとも2つのバイオマーカーについての正規化された発現値に対する機械学習解析またはロジスティック回帰分析により決定することができる。カットオフ値は、個々の使用者がケースバイケースで、例えば、使用される母斑-黒色腫分類子についての、特定の感度値および特異度値ならびに/またはAUC値を選択することにより決定することができる。カットオフ値を決定するための他の方法は、WO 2/103320号;ならびに米国特許第7,171,311号;同第7,514,209号;同第7,863,001号;および同第8,019,552号 (診断検査において有用なカットオフ値を決定するための方法について記載する限りにおいて、全てが参照により本明細書に組み込まれる) において提示されている。

30

## 【0193】

いくつかの例では、被験試料 (例えば、皮膚生検) を、アルゴリズム出力値が、複数の既知の良性試料と、カットオフ値の同じ側であれば、良性として特徴付けるか、または出力値が、複数の既知の悪性試料と、カットオフ値の同じ側であれば、悪性として特徴付ける。一例では、試料を、出力値が、カットオフ値を下回る場合は、良性として特徴付けるか、または出力値が、カットオフ値を上回る場合は、悪性として特徴付ける。別の例では、試料を、出力値が、カットオフ値を上回る場合は、良性として特徴付けるか、または出力値が、カットオフ値を下回る場合は、悪性として特徴付ける。

40

## 【0194】

分子プロファイリングおよび分類子出力

二値分類子を含む、開示される方法により、メラニン細胞含有試料など、生物学的試料を分類する場合、4つの可能なアウトカムが得られることが典型的である。予測からのアウトカムが  $p$  であり、実際の値もまた  $p$  である場合、アウトカムは、真陽性 (TP) と呼ばれるが、実際の値が  $n$  である場合、アウトカムは、偽陽性 (FP) と呼ばれる。逆に、予測アウトカムおよび実際の値の両方が  $n$  である場合は、真陰性が生じ、予測アウトカム

50

は  $n$  であるが、実際の値が  $p$  である場合は偽陰性が生じる。試料が黒色腫（例えば、原発性黒色腫）であるのかどうかを決定しようとする実施形態について考察する。この場合における偽陽性は、試料が、検査では陽性とされているが、実際には黒色腫（例えば、原発性黒色腫）でない場合に生じる。他方、偽陰性は、試料が、実際には黒色腫（例えば、原発性黒色腫）であるが、検査では陰性（すなわち、黒色腫ではない）とされている場合に生じる。いくつかの実施形態では、関連する比率で利用可能な試料に対してなされる誤差をリサンプリングすることにより、現実世界における亜型の有病率を仮定する ROC 曲線を作成することができる。

#### 【 0 1 9 5 】

黒色腫（例えば、原発性黒色腫）の陽性予測値（PPV）、または精度、または検査後確率とは、正確に黒色腫（例えば、原発性黒色腫）である陽性の検査結果を有する試料の比率である。PPVは、陽性検査が、検定される基本仮説（例えば、試料は、黒色腫（原発性黒色腫など）であるという仮説）を反映する確率を反映する。一例では、

偽陽性率（ ） =  $FP / (FP + TN)$  - 特異度

偽陰性率（ ） =  $FN / (TP + FN)$  - 感度

検出力 = 感度 =  $1 -$

陽性尤度比 = 感度 /  $(1 -$  特異度)

陰性尤度比 =  $(1 -$  感度) / 特異度

[ 式中、TNは、真陰性であり、FNは、偽陰性であり、TPおよびFPは、上記で定義した通りである ] である。

#### 【 0 1 9 6 】

陰性予測値（NPV）とは、陰性の検査結果（例えば、母斑または判定保留）を有する被験体または試料であって、正確に診断または亜型分類された被験体または試料の比率である。所与の検査についてのNPVが大きいことは、検査が陰性結果をもたらす場合、それがその評価において正確である可能性が極めて高いことを意味する。

#### 【 0 1 9 7 】

いくつかの実施形態では、本開示される方法の遺伝子発現解析の結果は、所与の診断（例えば、母斑または黒色腫または判定保留）が正確であるという統計学的信頼水準をもたらす。いくつかの実施形態では、このような統計学的信頼水準は、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、または99.5%を上回る。

#### 【 0 1 9 8 】

本開示の一態様では、別の方法（例えば、組織病理学的検査および/または免疫細胞化学的検査）により処理され、診断された試料を、次いで、第2の診断スクリーンとしての、開示される分子プロファイリングにかける。この第2の診断スクリーンは、少なくとも、1) 偽陽性および偽陰性の有意な低減、2) 結果として生じる病理の一因となる、根底にある遺伝子経路、代謝経路、またはシグナル伝達経路の決定、3) 統計学的確率を診断の精確さへと割り当てる能力、4) 両義的な結果を解明する能力、および5) かつて両義的だった試料を適正に特徴付ける能力を可能とする。

#### 【 0 1 9 9 】

いくつかの実施形態では、生物学的試料を、母斑または黒色腫（例えば、原発性黒色腫）として、75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、または99.5%を超える精確さで分類する。上記の文で使用した精確さという用語は、特異度、感度、陽性予測値、陰性予測値、および/または偽発見率を含む。

#### 【 0 2 0 0 】

他の場合には、受診者動作特性（receiver operator characteristic）（ROC）解析を使用して、最適のアッセイパラメーターを決定して、精確さ、特異度、陽性予測値、陰性予測値、および/または偽発見率の具体的レベルを達成することができる。ROC曲線とは、その弁別閾値を変化させるときの、二値分類子

10

20

30

40

50

システムの性能を例示するグラフプロットである。ROC 曲線は、陽性のうちの真陽性の割合 (TPR = 真陽性率) を、陰性のうちの偽陽性の割合 (FPR = 偽陽性率) と対比して、多様な閾値設定においてプロットすることにより作成する。

#### 【0201】

##### 方法の実装

本明細書で記載される、分類子を伴う方法などの方法は、多数の様式で実装することができる。いくつかの代表的な非限定的な実施形態を下記に記載する。

#### 【0202】

いくつかの方法の実施形態では、遺伝子発現データを、本明細書で記載される多様なアルゴリズムを適用するためのコンピュータまたは他のデバイス、マシン、または装置へと入力 (例えば、手動で、または自動で) するが、これは、膨大な遺伝子発現データ点を回収および加工する場合に特に有利である。他の実施形態は、通信基盤、例えば、インターネットの使用を伴う。多様な形態のハードウェア、ソフトウェア、ファームウェア、プロセッサ、またはこれらの組合せは、具体的な分類子および方法の実施形態を実装するのに有用である。ソフトウェアは、プログラム保存デバイス上に明確に統合されたアプリケーションプログラムとして実装することもでき、ユーザーのコンピュータ環境およびレビューア (この場合、レビューアは、関連のリモートサイトに (例えば、サービスプロバイダーの施設に) 配属される) のコンピュータ環境に実装されるソフトウェアの異なる部分として (例えば、アプレットとして) 実装することもできる。

#### 【0203】

例えば、ユーザーによるデータ入力の際に、またはユーザーによるデータ入力の後で、データ加工の一部を、ユーザー側のコンピュータ環境で実施することができる。例えば、ユーザー側のコンピュータ環境は、尤度「スコア」を提示するように定義された検定コードを提供するようにプログラムすることができ、ここで、スコアは、加工されるかまたは部分的に加工された応答として、レビューアのコンピュータ環境へと、レビューアのコンピュータ環境において結果を提示し、かつ/または報告を作成する 1 または複数のアルゴリズムをその後に実行するための検定コードの形態で伝送される。スコアは、数値スコア (数値を表す) の場合もあり、数値または数値の範囲を表す非数値スコア (例えば、アウトカムの 90 ~ 95 % の尤度を表す「A」) の場合もある。

#### 【0204】

本明細書で記載されるアルゴリズムを実行するためのアプリケーションプログラムは、任意の適切なアーキテクチャーを含むマシンへとアップロードし、マシンにより実行することができる。一般にマシンは、1 または複数の中央処理装置 (CPU)、ランダムアクセスメモリ (RAM)、および入力/出力 (I/O) インターフェース (複数可) などのハードウェアを有するコンピュータプラットフォームを伴う。コンピュータプラットフォームはまた、オペレーティングシステムおよびマイクロ命令コードも含む。本明細書で記載される多様なプロセスおよび関数は、マイクロ命令コードの一部の場合もあり、オペレーティングシステムを介して実行されるアプリケーションプログラムの一部 (またはこれらの組合せ) の場合もある。加えて、多様な他の周辺デバイスは、さらなるデータ保存デバイスおよび印字デバイスなどのコンピュータプラットフォームへと接続することもできる。

#### 【0205】

コンピュータシステムとして、システムは一般に、プロセッサユニットを含む。プロセッサユニットは、検査データ (例えば、応答遺伝子のレベル、参照遺伝子産物 (複数可) のレベル; 応答遺伝子の正規化されたレベル) を含むことが可能であり、また、患者データなどの他のデータも含みうる情報を受信するように作動する。この受信された情報は、データベース内に少なくとも一時的に保存することができ、データを解析して、上記で記載した報告を作成することができる。

#### 【0206】

入力データおよび出力データの一部または全部はまた、電子的に送信することもできる

10

20

30

40

50

(ある種の出力データ(例えば、報告)は、電子的に送信することもでき、電話で(例えば、fax backなどのデバイスを使用して、ファクシミリにより)送信することもできる)。例示的な出力受信デバイスは、表示エレメント、プリンター、ファクシミリデバイスなどを含みうる。伝送および/または表示の電子的形態は、電子メール(email)、インタラクティブテレビなどを含みうる。一実施形態では、入力データの全部もしくは一部および/または出力データの全部もしくは一部(例えば、通常は少なくとも最終報告)は、典型的なブラウザによるアクセス、好ましくは、機密アクセスのために、ウェブサーバー上に維持される。データは、所望される保健従事者にアクセス可能であるかまたは保健従事者に送信することができる。最終報告の全部または一部を含む入力データおよび出力データを使用して、医療施設における機密データベース内に存在しうる患者の医療記録に追加することができる。いくつかの例では、方法は、報告を作成するステップを含む。いくつかの例では、報告は、黒色腫には「+」もしくは「M」または母斑には「-」もしくは「N」など、試料の分類を指し示すアイコンを含む。

10

#### 【0207】

本明細書に記載される方法における使用のためのシステムは一般に、少なくとも1つのコンピュータプロセッサ(例えば、方法を単一のサイトでその全体において実行する場合)または少なくとも2つのネットワーク接続されたコンピュータプロセッサ(例えば、ユーザー(本明細書ではまた、「クライアント」とも称する)がデータを入力し、リモートサイトの、解析のための第2のコンピュータプロセッサへと伝送する場合であって、第1のコンピュータプロセッサと第2のコンピュータプロセッサとをネットワークにより、例えば、イントラネットまたはインターネットを介して接続する場合)を含む。システムはまた、入力のためのユーザーコンポーネント(複数可)ならびにデータ、作成された報告、および手動の介入を精査するためのレビューコンポーネント(複数可)も含みうる。システムのさらなるコンポーネントは、サーバーコンポーネント(複数可); およびデータを保存するためのデータベース(複数可)(例えば、報告エレメント、例えば、解釈レポートエレメントのデータベース、またはユーザーによるデータ入力およびデータ出力を含みうるリレーショナルデータベース(RDB)における)を含みうる。コンピュータプロセッサは、パーソナルデスクトップコンピュータ(例えば、IBM、Dell、Macintosh)、携帯用コンピュータ、メインフレーム、ミニコンピュータ、または他のコンピュータデバイスにおいて典型的に見出されるプロセッサでありうる。

20

30

#### 【0208】

ネットワーク接続されたクライアント/サーバーアーキテクチャーは、所望に応じて選択することができ、例えば、古典的な2階層または3階層のクライアントサーバーモデルでありうる。アプリケーションサーバーコンポーネントの一部としての、または個別のコンポーネント(RDBマシン)としてのリレーショナルデータベース管理システム(RDBMS)は、データベースとのインターフェースを提供する。

#### 【0209】

一例では、アーキテクチャーを、データベース中心のクライアント/サーバーアーキテクチャーであって、一般に、クライアントアプリケーションが、アプリケーションサーバーであり、要請に応じた多様な報告エレメント、特に、解釈レポートエレメント、とりわけ、解釈的文章表現および警告を伴う報告を追加するように、データベース(またはデータベースサーバー)への要求を行うアプリケーションサーバーからサービスを要求するアーキテクチャーとして提供する。サーバー(複数可)(例えば、アプリケーションサーバーマシンまたは個別のRDB/リレーショナルデータベースマシンの一部としての)は、クライアントの要求に応答する。

40

#### 【0210】

入力クライアントコンポーネントは、アプリケーションを動作させるための、多岐にわたる動力および特徴を備える、完全な独立型パーソナルコンピュータでありうる。クライアントコンポーネントは通常、任意の所望されるオペレーティングシステム下で作動し、

50

通信エレメント（例えば、ネットワークに接続するためのモデムまたは他のハードウェア）、1または複数の入力デバイス（例えば、情報またはコマンドを転送するのに使用されるキーボード、マウス、キーパッド、または他のデバイス）、保存エレメント（例えば、ハードドライブ、または他のコンピュータで読取り可能な保存媒体、コンピュータで書き込み可能な保存媒体）、および表示エレメント（例えば、モニター、テレビ、LCD、LED、または情報をユーザーへと伝える他の表示デバイス）を含む。ユーザーは、入力コマンドを、入力デバイスを介して、コンピュータプロセッサへと入力する。一般に、ユーザーインターフェースは、ウェブブラウザアプリケーションのために書き込まれたグラフィカルユーザーインターフェース（GUI）である。

#### 【0211】

サーバーコンポーネント（複数可）は、パーソナルコンピュータ、ミニコンピュータ、またはメインフレームであることが可能であり、データ管理（management）、クライアント間の情報の共有、ネットワーク管理（administration）、およびセキュリティを提供する。使用されるアプリケーションおよび任意のデータベースは、同じサーバー上にある場合もあり、異なるサーバー上にある場合もある。

#### 【0212】

クライアントおよびサーバー（複数可）のための他のコンピュータの配置であって、メインフレームなど、単一のマシン上、マシンの集団上、または他の適するコンフィギュレーション上での加工を含む配置が想定される。一般にクライアントマシンとサーバーマシンとは、本開示の加工を達成するように、一体となって働く。

#### 【0213】

使用される場合、データベース（複数可）は通常、データベースサーバーコンポーネントへと接続され、データを保持する任意のデバイスでありうる。例えば、データベースは、コンピュータ用の任意の磁気保存デバイスまたは光学保存デバイス（例えば、CDROM、内蔵型のハードドライブ、テープドライブ）でありうる。データベースは、サーバーコンポーネントに対してリモート（ネットワーク、モデムなどを介するアクセスを伴う）に配置することもでき、サーバーコンポーネントに対してローカルに配置することもできる。

#### 【0214】

システムおよび方法において使用する場合、データベースは、データアイテム間の関係に従い構成され、アクセスされる、リレーショナルデータベースでありうる。リレーショナルデータベースは一般に、複数のテーブル（エンティティ）からなる。テーブルの行は、レコード（個別のアイテムについての情報の集合）を表し、列は、フィールド（レコードの特定の属性）を表す。その最も単純な構想では、リレーショナルデータベースとは、少なくとも1つの共通なフィールドを介して互いと「関連する」データエントリーの集合である。

#### 【0215】

いくつかの実施形態では、所望の場合、データを入力するサービス提供時点において、コンピュータおよびプリンターを装備するさらなるワークステーションを使用し、適切な報告を作成することもできる。コンピュータは（複数可）、所望に応じて、データの入力、伝送、解析、報告の受領などの開始を容易とするように、アプリケーションを立ち上げるためのショートカット（例えば、デスクトップ上の）を有しうる。

#### 【0216】

コンピュータで読取り可能な保存媒体

本開示はまた、コンピュータで読取り可能な保存媒体（例えば、CD-ROM、メモリキー、フラッシュメモ리카ード、ディスクなど）であって、コンピュータ環境内で実行されると、本明細書で記載される応答尤度評価の結果の全部または一部を実行するアルゴリズムの実装を提供するプログラムをその上に保存した保存媒体も想定する。コンピュータで読取り可能な媒体が、本明細書で記載される方法を実行するための完全なプログラムを含有する場合、プログラムは、収集、解析、および出力の生成のためのプログラムの

10

20

30

40

50

命令を含み、一般に、コンピュータで読取り可能なコードデバイスであって、本明細書で記載されるとおり、ユーザーと対話する、そのデータを解析情報と共に加工する、および、そのユーザーのために固有の印字媒体または電子媒体を作成するためのコードデバイスを含む。

#### 【0217】

保存媒体が、本明細書で記載される方法の一部（例えば、方法のユーザー側の態様（例えば、データ入力能力、報告受領能力など））の実装を提供するプログラムをもたらす場合、プログラムは、ユーザーにより入力されたデータの、リモートサイトにおけるコンピュータ環境への伝送（例えば、インターネットを介する、イントラネットを介するなど）を提供する。報告を作成するためのデータの加工または加工の完了は、リモートサイトにおいて実行することができる。完全な報告を提供するために、報告を精査し、必要とされる任意の手動の介入を完了した後に、次いで、完全な報告を、電子文書または印字された文書（例えば、ファクスまたは郵送される紙の報告）として、ユーザーへと返送することができる。本明細書で記載されるプログラムを含有する保存媒体は、適切な基材またはこのような指示を得ることができるウェブアドレス上に記録された指示（例えば、プログラムのインストール、使用などについての指示）と共にパッケージングすることができる。コンピュータで読取り可能な保存媒体はまた、応答尤度評価を実行するための1または複数の試薬（例えば、プライマー、プローブ、アレイ、または他のこのようなキットの構成要素）と組み合わせて提供することもできる。

#### 【0218】

出力

いくつかの実施形態では、特定の試料（患者）についてのスコアを決定したら、そのスコアを、医師または他の介護者へと表示し、かつ/または伝えることができる。例えば、検査の結果は、認知可能な出力であって、検査の結果についての情報を提供する出力により、ユーザー（医師または他の保健従事者、検査室関係者、または患者など）へと提供する。いくつかの例では、出力は、紙への出力（例えば、書面化された出力または印字された出力）、スクリーン上の表示、グラフ式の出力（例えば、グラフ、図表、または他の図）、または聴覚的出力である。したがって、出力は、作成された報告を含みうる。

#### 【0219】

例えば、出力は、文章表現によるスコア（必要に応じて、対応するスコアを伴う）でありうる。例えば、文章表現による出力は、「母斑と一致する」などの場合もあり、「黒色腫と一致する」など（「原発性黒色腫と一致する」など）の場合もあり、「判定保留」（例えば、母斑とも黒色腫とも一致しない）などの場合もある。このような文章表現による出力は、例えば、良性試料（例えば、母斑）または悪性試料（例えば、原発性黒色腫）の診断を下すのに使用することもでき、単に母斑を黒色腫（例えば、原発性黒色腫）から識別するときに医師を支援するのに使用することもできる。

#### 【0220】

他の例では、出力は、試料中の遺伝子発現もしくはタンパク質発現（表3、4、11、および/または13のうちのいずれかの遺伝子発現またはタンパク質発現など）の量、または試料中の遺伝子発現もしくはタンパク質発現（4、11、および/または13のうちのいずれかの遺伝子発現またはタンパク質発現など）の、対照と比較した相対量などの数値（例えば、定量的出力）である。さらなる例では、出力は、グラフ式の表示、例えば、被験体に由来する試料中の遺伝子発現またはタンパク質発現（表3、4、11、および/または13のうちのいずれかの遺伝子発現またはタンパク質発現など）の値（量または相対量など）を、検量線上に指し示すグラフである。特定の例では、出力（グラフ式の出力など）は、被験試料を母斑または黒色腫（例えば、原発性黒色腫）として特徴付けるカットオフ値またはレベルを示すまたは提供する。他の例では、出力は、試料が母斑として分類される場合の「N」もしくは「-」、試料が黒色腫として分類される場合の「M」もしくは「+」、または試料が判定保留（例えば、母斑とも黒色腫とも一致しない）として分類される場合の「I」もしくは「?」などのアイコンである。いくつかの例では、出力を

10

20

30

40

50

、例えば、物理的手段、聴覚的手段、または電子的手段（例えば、郵送、電話、ファクシミリ伝送、電子メール、または電子的医療記録との通信による）を介して提供することにより、ユーザーに連絡する。

#### 【0221】

さらなる例では、出力は、遺伝子発現もしくはタンパク質発現（表4、11、および/または13のうちのいずれかの遺伝子発現またはタンパク質発現など）の、対照と比べた増大、遺伝子発現もしくはタンパク質発現（表4、11、および/または13のうちのいずれかの遺伝子発現またはタンパク質発現など）の、対照と比べた減少、または遺伝子発現もしくはタンパク質発現（表4、11、および/または13のうちのいずれかの遺伝子発現またはタンパク質発現など）の、対照と比べた無変化の存在を同定することなど、試料中の遺伝子発現またはタンパク質発現（表3、4、11、および/または13のうちのいずれかの遺伝子発現またはタンパク質発現など）の相対量に関する定性的情報をもたらさう。

10

#### 【0222】

いくつかの例では、出力に、データを解釈するための指針、例えば、原発性黒色腫の存在または非存在を指し示す数値限界または他の限界を添付する。指針は、母斑または黒色腫（例えば、原発性黒色腫）が存在するのか存在しないのかを指定する必要はないが、このような診断を含む場合がある。出力内の指標（*indicia*）は、例えば、正常または異常の範囲もしくはカットオフを含む場合があり、次いで、出力の受領者は、これらを、結果を解釈する、例えば、診断計画または処置計画に達するのに使用することができる。他の例では、出力は、推奨される治療レジメンをもたらさう。いくつかの例では、検定は、他の臨床情報の決定（試料中の1または複数のさらなる黒色腫バイオマーカーの量を決定することなど）も含みうる。

20

#### 【0223】

##### 臨床使用ステップ

開示される方法の結果として、メラニン細胞含有試料（例えば、皮膚生検）を、良性（例えば、母斑）試料もしくは悪性（例えば、原発性黒色腫などの黒色腫）試料、または判定保留の試料もしくは疑わしい試料（例えば、がん、疾患、または状態を示唆する試料）、あるいは非診断（*non-diagnostic*）試料（例えば、がん、疾患、または状態の存在または非存在に関してもたらず情報が不十分である試料）として特徴付けることができる。これらの（および他の可能な）結果の各々は、訓練された臨床従事者に有用である。いくつかの方法の実施形態は、下記でより詳細に記載されるとおりの臨床的に関与するステップを含む。

30

#### 【0224】

##### 診断指標

診断は、被験体（例えば、患者）に、いかなる疾患または状態を有するまたは有しうるのかの情報を与える。本開示を通してより具体的に記載される通り、メラニン細胞含有試料を特徴付ける、任意の開示される方法の任意の結果は、例えば、被験体または保健従事者に、診断として提供することができる。したがって、いくつかの方法の実施形態では、診断（良性（例えば、母斑）もしくは悪性（例えば、原発性黒色腫などの黒色腫）、または判定保留の、もしくは疑わしい（例えば、がん、疾患、または状態を示唆する）、あるいは非診断（例えば、がん、疾患、または状態の存在または非存在に関してもたらず情報が不十分である）など）を、被験体または保健従事者に提供することが想定される。

40

#### 【0225】

##### 予後診断指標

予後診断とは、その試料が特定の検査結果（例えば、黒色腫と対比した母斑）を受けた被験体についての、起こり得る健康転帰である。予後診断の不良とは、被験体についての長期にわたる見通しが良好でないこと、例えば、1、2、3、または5年間の生存率が、50%以下（例えば、40%、30%、25%、20%、15%、10%、5%、2%、または1%以下）であることを意味する。他方、予後診断の良好とは、被験体についての

50

長期にわたる見通しが、かなり良好であること、例えば、1、2、3、または5年間の生存率が、30%、40%、50%、60%、70%、75%、80%、または90%を超えることを意味する。

【0226】

そのメラニン細胞含有試料が悪性（例えば、黒色腫）として特徴付けられた被験体の予後診断（その疾患または状態に関する）は、そのメラニン細胞含有試料が良性（例えば、母斑）として特徴付けられた被験体より不良である可能性がある。したがって、特定の方法の実施形態は、悪性（例えば、原発性黒色腫などの黒色腫など）として特徴付けられる被験試料が採取される被験体についての転帰（上記を参照されたい）が比較的に不良であることを予後診断するステップを含む。逆に、他の例示的な方法は、良性（例えば、母斑など）として特徴付けられる被験試料が採取される被験体についての転帰（上記を参照されたい）が比較的に良好であることを予後診断するステップを含む。

10

【0227】

治療（予測）指標

開示される方法は、試料が黒色腫（例えば、原発性黒色腫）として診断された場合に、黒色腫（例えば、原発性黒色腫）のための処置について被験体を選択するステップもさらに含む。代替的に、開示される方法は、試料が良性の母斑として診断された場合に、処置をしないことについて被験体を選択することもさらに含む。

【0228】

いくつかの実施形態では、開示される診断法は、患者の診断に応じて、以下：a) 被験体の決定された診断が、原発性黒色腫について陽性である場合に、被験体のための処置レジメン（1もしくは複数の化学療法剤、さらなる組織を除去するためのさらなる手術、またはこれらの組合せによる処置など）を処方するステップ；b) 被験体の決定された診断が原発性黒色腫について陰性であるか、もしくは良性の母斑について陽性である場合に、被験体のための処置レジメンを処方しないステップ；c) 被験体の決定された診断が、原発性黒色腫について陽性である場合に、処置（1もしくは複数の化学療法剤、さらなる組織を除去するためのさらなる手術、またはこれらの組合せによる処置など）を、被験体へと投与するステップ；およびd) 被験体の決定された診断が、原発性黒色腫であるか、もしくは良性の母斑について陽性である場合に、処置レジメンを被験体へと投与しないステップ、のうちの1または複数を含む。代替的な実施形態では、方法は、(a)～(d)のうちの1または複数推奨するステップを含む。したがって、開示される方法は、被験体に由来する試料が原発性黒色腫として特徴付けられる場合に、被験体を、原発性黒色腫について処置するステップもさらに含む。

20

30

【0229】

いくつかの例では、開示される方法を使用して黒色腫を有すると診断された被験体を処置するのに、化学療法を使用する。がんの処置では、化学療法とは、腫瘍細胞またはがん細胞など、急速に増殖する細胞の繁殖を停止させる（kill）かまたは緩徐にする、1または複数の薬剤（化学療法剤）の投与を指す。特定の例では、化学療法とは、1または複数の薬剤の投与であって、被験体における腫瘍細胞の数を、少なくとも約50%など、著明に低減する投与を指す。「化学療法剤」としては、がんの処置における治療有用性を有する任意の化学的薬剤が挙げられる。化学療法剤の例は、例えば、Fischerら（編）、The Cancer Chemotherapy Handbook、6版、Philadelphia、Mosby、2003年；ならびに/またはSkeelおよびKhleif（編）、Handbook of Cancer Chemotherapy、8版、Philadelphia、Lippincott, Williams & Wilkins（2011年）において見出すことができる。

40

【0230】

黒色腫を処置するのに典型的に使用される化学療法としては、インターロイキン2（IL2）、ダカルバジン、インターフェロン、イピリムマブ、タキソールを伴うカルボプラチン、顆粒球マクロファージコロニー刺激因子（GM-CSF）、および/またはベムラフ

50

エニブが挙げられる。被験体における化学療法剤の使用は、黒色腫など、がんの徴候または症状を低下させる場合もあり、転移の阻害など、がんの進行、転移、および/または成長を低減するか、停止させるか、または逆転する場合もある。

#### 【0231】

##### アレイ

本明細書では、発現（表4、11、および/または13の試料型特異的バイオマーカーのうち2つ以上の発現など）を検出するのに使用されうるアレイ、例えば、上記で論じたとおり良性の母斑または原発性黒色腫としてのメラニン細胞含有試料の特徴付けにおける使用のためのアレイが開示される。いくつかの実施形態では、開示されるアレイはまた、1または複数の正規化バイオマーカー（例えば、表3の正規化バイオマーカー）の発現を検出するのにも使用することができる。他の実施形態では、開示されるアレイはまた、表6、8、または14の遺伝子セットなど、本開示を通して記載される遺伝子セットの発現を検出するのにも使用することができる。特定の例では、アレイ表面は、プレート、ビーズ（または複数のビーズ）、またはフローセル（例えば、複数のチャンネルを有するフローセル）を含む。

#### 【0232】

いくつかの実施形態では、アレイは、明確に分離した領域またはアドレス可能な位置を含む固体表面であって、各領域が、表4、11、および/または13のバイオマーカーと直接ハイブリダイズすることが可能であり、いくつかの例では、表3に示される正規化遺伝子と直接ハイブリダイズすることが可能である、少なくとも1つの固定化されたオリゴヌクレオチドを有する固体表面を含みうる。いくつかの例では、アレイは、表4で列挙される32のバイオマーカー全て、および表3の全ての正規化バイオマーカー、または表11で列挙されるバイオマーカーの全て、および表3の全ての正規化バイオマーカーと、直接的または間接的に、特異的にハイブリダイズすることが可能な、固定化された捕捉プローブを含む。オリゴヌクレオチドプローブは、アレイ上の位置により同定可能である。別の例では、アレイは、各領域が、少なくとも1つまたは少なくとも2つの固定化された捕捉プローブを有する、明確に分離した領域を含みうる。固定化された捕捉プローブは、少なくとも2つの異なる、表4、11、および/または13のバイオマーカーと、直接的または間接的に、特異的にハイブリダイズすることが可能であり、いくつかの例では、表3に示される正規化遺伝子と、直接的または間接的に、特異的にハイブリダイズすることが可能である。捕捉プローブは、アレイ上の位置により同定可能である。アレイ上のプローブは、アドレス可能な様式で表面へと結合させることができる。例えば、各アドレス可能な位置は、個別に同定可能なビーズまたはフローセル内のチャンネルでありうる。

#### 【0233】

例えば、アレイは、少なくとも3つのアドレス可能な位置であって、各位置が、同じ特異性を有する固定化された捕捉プローブを有し、各位置が、各他の位置における捕捉プローブとは異なる特異性を伴う捕捉プローブを有する位置を含みうる。少なくとも3つの位置のうち2つにおける捕捉プローブは、表4、11、および/または13に列挙されるバイオマーカーに、直接的または間接的に、特異的にハイブリダイズすることが可能であり、少なくとも3つの位置のうち1つにおける捕捉プローブは、表3に列挙される正規化バイオマーカーに、直接的または間接的に、特異的にハイブリダイズすることが可能である。加えて、各捕捉プローブの特異性も、アレイのアドレス可能な位置により同定可能である。いくつかの例では、アレイは、少なくとも2つの分離した領域（マルチウェル表面上のこのようなウェル、またはフローセル内のチャンネル）であって、各領域が、少なくとも3つのアドレス可能な位置を有する分離した領域もさらに含む。いくつかの例では、このようなアレイは、表4、6、8、11、13、または14で列挙される全てのバイオマーカーおよび表3の少なくとも2つの正規子（例えば、RPS6KB2およびSDHA）と、直接的または間接的に、特異的にハイブリダイズすることが可能な、固定化された捕捉プローブを含む。いくつかの例では、捕捉プローブ（複数可）は、少なくとも2つの、表4、11、および/または13に列挙されるバイオマーカーおよび少なくとも1つの

10

20

30

40

50

、表3の正規化バイオマーカーと、核酸プログラミングリンカーを介して間接的にハイブリダイズし、ここで、プログラミングリンカーは、捕捉プローブ（複数可）と相補的な第1の部分およびヌクレアーゼ保護プローブ（NPP）と相補的な第2の部分を含むヘテロ二官能性リンカーであり、NPPは、表4、11、および/または13で列挙される少なくとも2つのバイオマーカーのうちの1つ、または少なくとも1つの、表3の正規化バイオマーカーと相補的である。したがって、いくつかの例では、アレイはまた、核酸プログラミングリンカーも含む。

#### 【0234】

いくつかの実施形態では、アレイは、表4、11、および/または13のバイオマーカーのうちの、少なくとも2つ、少なくとも3つ、少なくとも4つ、少なくとも5つ、少なくとも6つ、少なくとも7つ、少なくとも8つ、少なくとも9つ、少なくとも10、少なくとも11、少なくとも12、少なくとも13、少なくとも14、少なくとも15、少なくとも16、少なくとも17、少なくとも18、少なくとも19、少なくとも20、または、当てはまる場合、少なくとも21、少なくとも22、少なくとも23、少なくとも24、少なくとも25、少なくとも26、少なくとも27、少なくとも28、少なくとも29、少なくとも30、少なくとも31、または全て（表4、11、および/または13のバイオマーカーのうちの2つ、3つ、4つ、5つ、6つ、7つ、8つ、9つ、10、11、12、13、14、15、16、17、18、19、または、当てはまる場合、20、21、22、23、24、25、26、27、28、29、30、31、または全てなど）と相補的なオリゴヌクレオチドを含むかまたはこれらから本質的になるオリゴヌクレオチドを含む。いくつかの例では、アレイは、表3のバイオマーカーのうちの少なくとも1つ、少なくとも2つ、少なくとも3つ、少なくとも4つ、少なくとも5つ、少なくとも6つ、少なくとも7つ、少なくとも8つ、または全てなど（表3の正規化バイオマーカーのうちの1つ、2つ、3つ、4つ、5つ、6つ、7つ、8つ、もしくは9つ、またはRPS6KB2およびSDHAなど）、正規化バイオマーカーと相補的なオリゴヌクレオチドをさらに含む。いくつかの例では、アレイは、1または複数の対照オリゴヌクレオチド（1つ、2つ、3つ、4つ、5つ、6つ、7つ、8つ、9つ、10以上の対照オリゴヌクレオチドなど）、例えば、1または複数の陽性対照および/または陰性対照をさらに含む。いくつかの例では、対照オリゴヌクレオチドは、DEADボックスポリペプチド5（DDX5）、グリセルアルデヒド-3-リン酸デヒドロゲナーゼ（GAPDH）、フィブリリン1（FBN1）、または*Arabidopsis thaliana* AP2様エチレン応答性転写因子（ANT）のうちの1または複数と相補的である。

#### 【0235】

いくつかの実施形態では、アレイは、空間的に分離した領域（マルチウェル表面上のウェルまたはフローセル内のチャンネルなど）を有する表面であって、各領域が、表面へと安定的に（例えば、共有結合的に）結合させたアンカーおよび核酸プログラミングリンカーを含み、プログラミングリンカーが、捕捉プローブ（複数可）と相補的な第1の部分およびヌクレアーゼ保護プローブ（NPP）と相補的な第2の部分を含むヘテロ二官能性リンカーであり、NPPが、標的核酸（表4、11、および/または13のバイオマーカーなど）と相補的である表面を含みうる。いくつかの実施形態では、アレイは、第1の部分がアンカーと相補的であり、第2の部分がNPPと相補的な二官能性リンカーであって、NPPが、表4、11、および/または13のバイオマーカーのうちの少なくとも2つ、少なくとも3つ、少なくとも4つ、少なくとも5つ、少なくとも6つ、少なくとも7つ、少なくとも8つ、少なくとも9つ、少なくとも10、少なくとも11、少なくとも12、少なくとも13、少なくとも14、少なくとも15、少なくとも16、少なくとも17、少なくとも18、少なくとも19、または、当てはまる場合、少なくとも20、少なくとも21、少なくとも22、少なくとも23、少なくとも24、少なくとも25、少なくとも26、少なくとも27、少なくとも28、少なくとも29、少なくとも30、少なくとも31または32全て（表4、11、および/または13のバイオマーカーのうちの2つ、3つ、4つ、5つ、6つ、7つ、8つ、9つ、10、11、12、13、14、15、

10

20

30

40

50

16、17、18、19、または、当てはまる場合、20、21、22、23、24、25、26、27、28、29、30、31、または32など)のうちの1つと相補的である二官能性リンカーを含むかまたはこれらから本質的になる。いくつかの例では、アレイは、第1の部分アンカーと相補的であり、第2部分が、表3のバイオマーカーのうち少なくとも1つ、少なくとも2つ、少なくとも3つ、少なくとも4つ、少なくとも5つ、少なくとも6つ、少なくとも7つ、少なくとも8つ、または全てなど(表3のバイオマーカーのうち1つ、2つ、3つ、4つ、5つ、6つ、7つ、8つ、または9つなど)の正規化バイオマーカーと相補的なNPPと相補的な、二官能性リンカーをさらに含む。このようなアレイは、NPPと相補的でない二官能性リンカーの少なくともセグメントとハイブリダイズするアンカーを、アレイへと結合させている。別の例では、アレイは、二官能性リンカーの第2部分が、対照遺伝子(DDX5、GAPDH、FBN1、またはANTなど)と相補的なNPPと相補的な二官能性リンカーをさらに含む。このようなアレイは、(1)プログラミングリンカーの第1の部分とハイブリダイズさせたアンカープローブ、(2)プログラミングリンカーの第2の部分(いくつかの例では、ピオチン化される)とハイブリダイズさせたNPP、(3)第1部分をNPPとハイブリダイズさせ、第2部分を検出プローブとハイブリダイズさせた二官能性検出リンカー、(4)検出プローブ、(5)標識(アビジンHRPなど)、またはこれらの組合せもさらに含む。

#### 【0236】

いくつかの例では、最大で47の異なる捕捉(すなわち、アンカー)オリゴヌクレオチドの集合を、表面上の空間的に異なる位置に配置し、誘導体化した表面と安定的に会合させる(例えば、誘導体化した表面へと共有結合的に結合させる)ことができる。任意の特定のアッセイでは、捕捉プローブの所与のセットを使用して、各ウェルの表面を、47もの異なる対象の標的またはアッセイ種類に特異的であるようにプログラム化し、異なる被験試料を、各プレート内の96のウェルの各々へと適用することができる。同じ捕捉プローブのセットを使用して、ウェルの表面を、対象の他の標的およびアッセイのために、複数回にわたり再プログラム化することもできる。

#### 【0237】

##### アレイ基板

アレイの固体支持体は、有機ポリマーから形成することができる。固体支持体に適する材料としては、ポリプロピレン、ポリエチレン、ポリブチレン、ポリイソブチレン、ポリブタジエン、ポリイソプレン、ポリビニルピロリジン、ポリテトラフルオロエチレン(polytetrafluoroethylene)、ポリビニリデンジフルオリド(polyvinylidene difluoride)、ポリフルオロエチレン-プロピレン、ポリエチレンビニルアルコール、ポリメチルペンテン、ポリクロロトリフルオロエチレン(polychlorotrifluoroethylene)、ポリスルホン(polysulfone)、ヒドロキシル化二軸延伸ポリプロピレン、アミノ化二軸延伸ポリプロピレン、チオール化二軸延伸ポリプロピレン、エチレンアクリル酸、エチレンメタクリル酸(ethylene methacrylic acid)、およびこれらのコポリマーのブレンド(米国特許第5,985,567号を参照されたい)が挙げられるがこれらに限定されない。本明細書で開示されるアレイに適する基材の他の例としては、ガラス(官能化ガラスなど)、Si、Ge、GaAs、GaP、SiO<sub>2</sub>、SiN<sub>4</sub>、変性シリコン、ニトロセルロース、ポリスチレン、ポリカーボネート、ナイロン、繊維、またはこれらの組合せが挙げられる。アレイ基材は、硬く比較的不可撓性(例えば、ガラスまたは支持膜)の場合もあり、可撓性(ポリマー膜など)の場合もある。

#### 【0238】

一般に、固体支持体表面を形成するのに使用されうる材料の適切な特徴としては、活性化すると、支持体の表面が、オリゴヌクレオチドなどの生体分子を、表面へと安定的に(例えば、共有結合的に、静電的に、可逆的に、不可逆的に、または恒久的に)結合させることが可能となるような表面活性化に適すること;「インサイチュ」における生体分子の合成に適すること;オリゴヌクレオチドまたはタンパク質(抗体など)により占有されて

10

20

30

40

50

いない支持体上の領域が非特異的な結合に適さないように、または、非特異的な結合が生じる場合は、オリゴヌクレオチドまたはタンパク質（抗体など）を除去することなく、このような材料を表面からたやすく除去しうるように、化学的に不活性であることが挙げられる。

#### 【0239】

別の例では、表面活性化有機ポリマーを、固体支持体表面として使用する。表面活性化有機ポリマーの一例は、ラジオ周波数プラズマ放電を介してアミノ化されたポリプロピレン材料である。また、カルボキシル化基、ヒドロキシル化基、チオール化基、または活性エステル基など、他の反応性基も、使用することができる。

#### 【0240】

##### アレイフォーマット

アレイ内で、各々の配置された試料は、その位置を、アレイの次元（例えば、少なくとも二次元）内において、信頼できる形で、かつ、一貫した形で決定しうる点でアドレス可能である。アレイ上の特徴の適用位置は、異なる形状を呈する場合がある。例えば、アレイは、規則的（均一の行内および列内に配置されるか、または複数の個別に同定可能なビーズにより提示されるなど）な場合もあり、不規則的な場合もある。したがって、秩序化アレイ内では、試料をアレイへと適用するときに、各試料の位置を試料へと割り当て、各位置を適切な標的または特徴の位置と関連させるためのキーを与えることができる。秩序化されたアレイは、対称性の格子パターンで配置することが多いが、試料は、他のパターン（放射状に分布させた直線で、渦巻き線で、または秩序化クラスターで、など）で配置することもできる。アドレス可能なアレイは通常、コンピュータは、アレイ上の特定のアドレスを、その位置における試料についての情報（例えば、シグナル強度などを含む、ハイブリダイゼーションデータまたは結合データ）と関連させるようにプログラムしうるという点において、コンピュータで読取り可能である。コンピュータで読取り可能なフォーマットのいくつかの例では、アレイ内の個々の特徴を、例えば、コンピュータによりアドレス情報と関連させうるデカルト格子パターンで規則的に配置する。

#### 【0241】

一例は、当技術分野では一般にディップスティックと称する、オリゴヌクレオチドバンドの直線状アレイを含む。別の適切なフォーマットは、離れたセルの二次元パターン（ $64 \times 64$  アレイによる  $4096$  個の四角など）を含む。一例では、アレイは、反応チャンパー1つ当たり最大で  $47$ （例えば、 $5$ 、 $5 \sim 16$ の間、 $5 \sim 47$ の間、 $16$ 、 $16 \sim 47$ の間）のアドレス可能な位置を含み、したがって、 $96$  ウェルアレイ内には、各反応チャンパー内のアドレス可能な位置（例えば、ウェル）が同じであるのか異なるのかにより、 $96 \times 5$ 、 $96 \times 16$ 、 $96 \times 47$  のアドレス可能な位置を存在させることができる（例えば、プログラム化可能なアレイ技術を使用して）が、アレイ技術の技術分野では、普遍的にプログラム化可能なアレイは、対象のアレイ表面上に物理的にプリントされうるアドレス可能な位置の数までの任意の数の分析物を捕捉するように柔軟にプログラム化しうるということが理解される。他の実施形態は、物理的に離れた表面であって、組み合わせられるとアドレス可能なアレイを作製する表面のセット、例えば、各々が、特異的分析物を捕捉するようにプログラム化またはプリントされた、個別に同定可能な（例えば、アドレス可能な）ビーズのセットへと一体に組み合わせられた表面を含むアレイを含む。当業者により十分に理解される通り、スロット（方形）アレイおよび円形アレイを含むがこれらに限定されない他のアレイフォーマットも、使用に同等に適する（米国特許第  $5,981,185$  号を参照されたい）。いくつかの例では、アレイは、マルチウェルプレート（ $96$  ウェルプレートなど）である。一例では、アレイは、糸、膜、または薄膜であるポリマー媒体上に形成する。有機ポリマー媒体の例は、厚さが約  $1$  ミル（ $0.001$  インチ）～約  $20$  ミルのオーダーの厚さを有するポリプロピレンシートであるが、薄膜の厚さはそれほど重要ではなく、かなりの広範囲にわたり変化させることができる。アレイは、それらの耐久性に加えて、バックグラウンド蛍光が少ない二軸延伸ポリプロピレン（BOPP）薄膜を含みうる。

10

20

30

40

50

## 【0242】

本開示のレイフォーマットは、様々な異なる種類のフォーマットに含ませることができる。「フォーマット」としては、例えば、マイクロ滴定プレート（例えば、マルチウェルプレート）、試験管、無機シート、ディップスティック、ビーズなど、固体支持体を固定しうる任意のフォーマットが挙げられる。例えば、固体支持体がポリプロピレン系である場合は、1または複数のポリプロピレン系を、プラスチック製のディップスティック型デバイスへと固定することができ、ポリプロピレン膜は、ガラススライドへと固定することができる。特定のフォーマットそれ自体は、重要ではない。必要な全てのことは、固体支持体またはその上に吸着させた任意の生体ポリマーの機能的な挙動に影響を及ぼすことなく、それへと固体支持体を固定しうること、およびフォーマット（ディップスティックまたはスライドなど）が、デバイスを導入するいかなる材料（臨床試料およびハイブリダイゼーション溶液など）に対しても安定的なことである。

10

## 【0243】

本開示のレイは、様々な手法により調製することができる。一例では、オリゴヌクレオチド配列を個別に合成し、次いで、固体支持体へと結合させる（米国特許第6,013,789号を参照されたい）。別の例では、配列を支持体上に直接合成して、所望のレイを提供する（米国特許第5,554,501号を参照されたい）。当業者には、オリゴヌクレオチドを固体支持体へとカップリングする、およびオリゴヌクレオチドを支持体上に直接合成するのに適する方法が公知であり、適切な方法の概要は、Matsonら、Anal. Biochem., 217巻:306~10頁、1994年において見出すことができる。一例では、オリゴヌクレオチドを固体支持体上で調製するための従来の化学的技法を使用して、オリゴヌクレオチドを支持体上に合成する（PCT出願第WO85/01051号および同第WO89/10977号、または米国特許第5,554,501号など）。

20

## 【0244】

適切なレイは、レイのセル内でオリゴヌクレオチドを合成する自動式手段を使用して、4つの塩基の前駆体を所定のパターンで配置することにより作製することができる。略述すると、複数チャンネル型自動式化学的送達システムを援用して、オリゴヌクレオチドプローブ集団を、基材にわたり並行する行（送達システム内のチャンネルの数に対応する数の行）で作製する。第1の方向におけるオリゴヌクレオチドの合成の完了後、次いで、基材を90°回転させて、今度は第1のセットに垂直な第2の行のセット内で合成が進行することを可能とすることができる。このプロセスにより、その交点（intersection）が複数の離れたセルを生成する、複数チャンネルレイが作製される。

30

## 【0245】

オリゴヌクレオチドは、オリゴヌクレオチドの3'端またはオリゴヌクレオチドの5'端を介して、支持体に結合させることができる。一例では、オリゴヌクレオチドは、3'端を介して、固体支持体に結合させる。しかし、当業者は、オリゴヌクレオチドの3'端の使用が固体支持体への結合に適するののか、オリゴヌクレオチドの5'端の使用が固体支持体への結合に適するののかを決定することができる。一般に、3'端および5'端領域におけるオリゴヌクレオチドプローブの内部の相補性が、支持体への結合を決定する。

40

## 【0246】

キット

本明細書ではまた、発現（表4、11、および/または13のバイオマーカーのうちの2つ以上の発現など）を検出するのに使用されうる、例えば、上記で論じたとおりの良性の母斑または原発性黒色腫としての試料の特徴付けにおける使用のためのキットも開示される。いくつかの実施形態では、開示されるキットはまた、1または複数の正規化バイオマーカー（例えば、表3の正規化バイオマーカー）の発現を検出するのにも使用することができる。特定の例では、キットは、本明細書で提示されるレイのうちの1または複数を含む。

## 【0247】

50

いくつかの例では、キットは、表4、11、および/または13のバイオマーカーのうち2つ以上など、核酸またはタンパク質の発現を検出するためのプローブおよび/またはプライマーを含み、いくつかの例では、1または複数の表3の正規化バイオマーカーを含む。いくつかの例では、キットは、表4、11、および/または13に列挙されるバイオマーカーに特異的に結合する抗体を含む。例えば、キットは、本明細書で開示されるバイオマーカーを検出するためのアレイを構築するのに必要とされる1または複数の核酸プローブを含みうる。

いくつかの例では、キットは、核酸プログラミングリンカーを含む。プログラミングリンカーは、アレイ上の捕捉プローブ(複数可)と相補的な第1の部分およびヌクレアーゼ保護プローブ(NPP)と相補的な第2の部分を含むヘテロ二官能性リンカーであり、NPPは、表4、11、および/または13で列挙される少なくとも2つのバイオマーカーのうち1つ、または少なくとも1つの、表3の正規化バイオマーカーと相補的である。一例では、表面が、アドレス可能な固定化された捕捉プローブおよび核酸プログラミングリンカーを含むように、プログラミングリンカーを、捕捉プローブとあらかじめハイブリダイズさせ、それらを共有結合的に結合させないようにする。このような例では、キットは、プログラミングリンカーを伴う個別の容器を有さない。

#### 【0248】

いくつかの例では、キットは、NPPを含む。NPPは、プログラミングリンカーの第2の部分と相補的である。例示的なNPPは、配列番号1~36、および123~164に示される。

#### 【0249】

いくつかの例では、キットは、二官能性検出リンカーを含む。このようなリンカーは、検出プローブで標識することができ、NPPまたは標的(表4、11、および/または13のものなど)と特異的にハイブリダイズすることが可能である。

#### 【0250】

いくつかの例では、キットは、本明細書で開示されるアレイ、ならびに緩衝液(溶解緩衝液など)を含有する容器;一本鎖核酸に特異的なヌクレアーゼを含有する容器;核酸プログラミングリンカーを含有する容器;NPPを含有する容器;複数の二官能性検出リンカーを含有する容器;検出プローブ(三重ビオチン化された検出プローブなど)を含有する容器;および検出試薬(アビジンHRPなど)を含有する容器のうち1または複数を含む。

#### 【0251】

一例では、キットは、正常皮膚細胞(例えば、良性の母斑)においてもしくは原発性黒色腫において予測される、表4、11、および/もしくは13のバイオマーカーの予測された値もしくは値の範囲、または臨床的に有用なカットオフを示すグラフまたは表を含む。いくつかの例では、キットは、表4、11、および/または13のバイオマーカーについての、特定の量の核酸またはタンパク質などの対照試料をさらに含む。

#### 【0252】

キットは、さらなる構成要素、例えば、指示材料およびさらなる試薬、例えば、酵素ベースの検出系(例えば、西洋ワサビペルオキシダーゼまたはアルカリホスファターゼおよび適切な基質を含む検出試薬)、二次抗体(例えば、表3、4、11、および/または13の標的(例えば、タンパク質)に特異的に結合する一次抗体に特異的に結合する抗体)、または抗体に標識するための手段もさらに含みうる。キットはまた、キットがデザインされた特定の適用を容易とするさらなる構成要素(例えば、マイクロ滴定プレート)も含みうる。一例では、キットは、対照核酸をさらに含む。当業者には、このようなキットおよび適切な内容物が周知である。指示材料は、書面の場合もあり、電子形態(コンピュータディスクまたはコンパクトディスクなど)の場合もあり、視覚形態(ビデオファイルなど)の場合もある。

#### 【0253】

以下の実施例は、ある種の特徴および/または実施形態を例示する目的で提示す

10

20

30

40

50

る。これらの実施例は、本発明を、記載される特定の特徴または実施形態へと限定するとはみなさないものとする。

#### 【実施例】

#### 【0254】

##### (実施例1)

臨床的に特徴付けられた皮膚試料の発見セットを使用した遺伝子の選択

母斑および黒色腫細胞は、全ての細胞と同様に大量の遺伝子を発現するが、それらの多くは、これらグループ間の識別には無関係である。したがって、有用な遺伝子情報を抽出して次元数を小さくするために、この実施例は、2600種より多くのmRNA標的発現の最初のスクリーニングにより、ヒト被験体から生検によって得られたホルマリン固定パラフィン包埋(「FFPE」)皮膚試料中で有意に差次的に発現されたmRNAが同定されることを記載する。さらに実施例全体で使用された方法論的な詳細も記載する。

10

#### 【0255】

それぞれ厚さおよそ5 $\mu$ mであり顕微鏡スライド上に載せられた39個のFFPE組織切片の発見セットは、John Wayne Cancer Institute(JWCI)組織バンクから提供された。このセットには、14個の正常な皮膚試料、10個の母斑試料、5個の原発性黒色腫試料、および10個の黒色腫転移試料が包含されていた。

#### 【0256】

##### 試料の調製および溶解

簡単に言えば、各FFPE組織切片を測定して、そのおよその面積( $\text{cm}^2$ )を決定した。次いでかみそり刃を使用してスライド上の過量のパラフィンが入らないように組織切片を掻き取り、それを、ラベルを付けたエッペンドルフチューブに入れた。適用可能な組織切片0.3 $\text{cm}^2$ 当たり25 $\mu$ lの予熱した(50 $^\circ\text{C}$ )ホルムアミドおよびSDSを含むSSC緩衝液に試料を懸濁した。次いで組織懸濁液上に、界面活性剤(例えば、Brij-97)を含有する500 $\mu$ lの鉱油(「非水層」)を載せ、この溶解反応物を95 $^\circ\text{C}$ で10~15分インキュベートした。この反応混合物を短時間で冷却した後、プロテイナーゼKを1mg/mlの最終濃度まで添加し、インキュベートを50 $^\circ\text{C}$ で30~60分続けた。溶解反応物の一部を即座にヌクレアーゼ保護アッセイ(nuclease protection assay)に使用するか(以下を参照)、または溶解反応物(またはそれらの残りの部分)を-80 $^\circ\text{C}$ で凍結して保存した。その後の使用の前に、凍結した溶解反応物を50 $^\circ\text{C}$ で10~15分で融解させた。

20

30

#### 【0257】

##### ヌクレアーゼ保護アッセイ(「NPA」)

25 $\mu$ lの溶解させた各溶解反応混合物を96ウェルプレートのウェル中に入れ、70 $\mu$ lの非水層を上に乗せた。各ウェルに、5 $\mu$ lのヌクレアーゼ保護プローブ(NPP)ミックスを添加した。NPPミックス中に、検出しようとする複数のmRNA標的のそれぞれに相補的な1nMの(過量の)NPPが存在した。ArrayPlate検出用のNPPは、(i)長さが50塩基対であり、NPPのそれぞれの半分が40~75の範囲のTm(および全長は60~85の範囲のTm)を有し、さらに(ii)それぞれのmRNA標的への特異性に関して、インシリコで(NCBI BLASTを使用して)、インビトロで転写物を用いてテストされた(NPA反応において、他のNPP、他の標的、または他の分析物との交差反応性は実質的になかった)。ArraySlide検出用のNPPは、NPPの3'末端に向けて偏りのある内部のビオチン化塩基(T)を含有するという点のみにおいて異なっている。NPPはさらに、他の実施例で特異的に同定された遺伝子に関連して記載される。

40

#### 【0258】

96ウェルのNPAプレートを95 $^\circ\text{C}$ で10~15分加熱して、核酸を変性させ、次いで60 $^\circ\text{C}$ で6~16時間インキュベートして、NPPをそれぞれのmRNA標的にハイブリダイズさせた。

#### 【0259】

50

ハイブリダイゼーション工程後、各ウェルの水相に酢酸ナトリウム緩衝液中の $20\mu\text{l}$ の過量のS1ヌクレアーゼ( $2.5\text{U}/\mu\text{l}$ )を添加した。S1反応を $50^\circ\text{C}$ で $90\sim 120$ 分進行させて、未結合のmRNAおよび未結合のNPPを消化した。いくつかの反応セットで、分子生物学グレードの水中のBSAを最終濃度が $40\text{mg}/\text{ml}$ になるまで添加した。

#### 【0260】

S1消化工程中に、96ウェルのNPAプレート中の反応に対応する各ウェルに $0.1\text{M}$ のEDTAおよび $1.6\text{N}$ のNaOHを含有する $10\mu\text{l}$ の溶液を添加することによって、96ウェルの「ストップ」プレートを調製した。96ウェルのNPAプレート中の各反応物の全体積(約 $120\mu\text{l}$ )を、第二の96ウェルのストッププレート中の対応するウェルに移した。ストッププレートを $95^\circ\text{C}$ で $15\sim 20$ 分インキュベートし、次いで室温で $5\sim 10$ 分冷却し、その後、 $10\mu\text{l}$ の $1.6\text{N}$ のHClを添加して、予め各反応物に添加されていたNaOHを中和した。

10

#### 【0261】

この実施例におけるヌクレアーゼ保護アッセイ反応を、(i)各ウェル中で34種の推定のヒト黒色腫関連mRNA(または対照)の発現が検出されるように特注設計された、第1の96ウェルプレートベースのアレイ(Array Plate 番号1)、(ii)33種の推定のヒト黒色腫関連mRNA(または対照)が検出されるように特注設計された第2のArray Plate(すなわち、番号2)、(iii)各ウェル中で1829種の推定のヒトがん関連mRNA(または対照)の検出が可能な、第一のガラス製スライドベースの21ウェル(Array Slide 番号1)の特注「がんトランスクリプトーム」アレイ、および(iv)推定でヒトトランスクリプトームを代表する各2600種のmRNAにおいて検出が可能な、第二のArray Slide(すなわち、番号2)「全トランスクリプトーム」の特注アレイを使用して直接的に(例えば、標的mRNA分析物の精製または逆転写を行わないで)調査した。上記のアレイのそれぞれで調査しようとする標的を、例えば文献検索および公共の知識に基づき決定した。

20

#### 【0262】

Array Plateでの捕捉および検出

SDSを含有するSSC緩衝液(「SSC-S」)中 $5\text{nM}$ の50塩基対プログラミングリッカー(「PL」) $40\mu\text{l}$ を用いて、Array Plate 番号1および2をプログラム化した。PLは、一部分がアレイ表面に取り付けられたユニバーサルアンカー配列に相補的であり、他の部分が特定のアレイ位置に割り当てられた特定のNPPに相補的な、人工的な25塩基対の二官能性合成オリゴヌクレオチドコンストラクト(アダプター)であった。プログラミング工程後、ストッププレートからの各反応物の全水相( $60\sim 65\mu\text{l}$ )をプログラム化されたArray Plateの対応するウェルに添加し、 $50^\circ\text{C}$ で $16\sim 24$ 時間インキュベートし、未消化のNPP(これは、ヌクレアーゼ工程中に標的に結合するため、試料中に存在する標的の定量可能な代理物である)を捕捉した。その後、各反応物に1%乾燥脱脂乳を含むSSC-S中の $5\text{nM}$ の二官能性検出リンカー(「DL」)を添加し、続いて $60^\circ\text{C}$ で1時間インキュベートした。DLは、一部分がそれぞれのNPPに相補的であり、他の部分がビオチン標識検出プローブ(「DP」)の1つまたは複数の(例えば、2または3つの)コピー(ここで、これらのDPは全てのDLに設計されている検出領域に特異的に結合できる)に相補的な人工的な25塩基対の二官能性合成オリゴヌクレオチドコンストラクトであった。検出のための「サンドイッチ」を完成させるために、反応物に $40\mu\text{l}$ の $3\text{nM}$ のDPを添加し、続いて $50^\circ\text{C}$ で $45\sim 60$ 分インキュベートした。次に、1%乾燥脱脂乳を含むSSC-S中の $40\mu\text{l}$ のアビジンペルオキシダーゼ( $1:600$ )を添加し、続いて $37^\circ\text{C}$ で $30\sim 45$ 分インキュベートした。最後に、ペルオキシダーゼ酵素の存在下で光を発し、HTG OMI X(商標)イメージャーを使用してその光が捕捉される化学発光基質ミックスを添加した。遺伝子発現は、Array Plateのアドレス可能な各位置における放出された光の強度と直接関連する。

30

40

50

## 【0263】

Array Slideでの捕捉および検出

次いで、NPP捕捉のために、各ヌクレアーゼ保護アッセイ反応物(60~65 $\mu$ l)の全水相を、Array Slide番号1または番号2に50 で16~24時間ハイブリダイズさせた。ビオチン化NPP捕捉後、それぞれのArray Slideを、1% Tweenを含有する1xSSC(「洗浄緩衝液」)で徹底的に洗浄した。検出酵素緩衝液(1xSSC-S、0.05% Tweenおよび乾燥脱脂乳)中のアビジン-ペルオキシダーゼ(1:600)50 $\mu$ lを37 で45分かけて添加した。Array Slideを洗浄し、続いて検出のために増幅用希釈剤(Perkin Elmer)中のTSA-Plus Cy3試薬を添加した。室温で3分インキュベートした後、Array Slideを洗浄緩衝液中で洗浄することによってTSA-Plus Cy3反応を止めた。最後に、Array Slideを回転させて乾燥させ、GenePix 4200ALマイクロアレイスライドスキャナー(Molecular Devices、Sunnyvale、CA)を使用して5 $\mu$ mの解像度でスキャンした。以下に記載した通りに、分析のためにNimbleScan2.5ソフトウェア(Roche NimbleGen、Madison、WI)を使用してTIFF画像からプローブ強度を抽出した。

10

## 【0264】

データ分析

この実施例における各アレイからの生データを、BRBアレイツール(インターネットで、2012年6月4日現在、[linus.nci.nih.gov/~brb/download\\_full\\_v4\\_2\\_1\\_stable.html](http://linus.nci.nih.gov/~brb/download_full_v4_2_1_stable.html)で調査用途のために自由に利用可能)を使用して加工した。簡単に言えば、データを最小強度で閾値化し、分位数により正規化して、ある特定のデータフィルタを適用して非差次的データポイントを更なる分析から除去した。データをlog<sub>2</sub>変換し分析することにより、p値および対数の倍率変化値に基づいて、グループアレイ間で統計学的有意な差次的遺伝子を見出した。

20

## 【0265】

さらなる研究のために、(a)原発性黒色腫試料に対して有意な(p<0.05)母斑における差次的な発現、および、いくつかの場合では(b)このような発現が測定された各試料集団において生のシグナル強度が3000を超えるmRNA発現に基づいて、78種の遺伝子を選択した。また推定の対照としてもともと包含されていた追加の4種の遺伝子(SDHA、RPS6KB2、RPL37A、およびTFRC)もさらなる研究に繰り越して用いられた。

30

## 【0266】

(実施例2)

臨床的に特徴付けられた皮膚試料の第二のセットで有意に差次的に発現された遺伝子-4種への正規化  
この実施例は、JWCI組織バンクによって母斑または原発性黒色腫のいずれかとして特徴付けられたヒト皮膚生検間でそのmRNA発現が有意に異なる32種の遺伝子のセットを同定することを記載する。

## 【0267】

実施例1で同定されたかまたは実施例1から繰り越された82種のmRNA標的に加えて、経路分析で同定されたかまたは陰性対照として使用される6種の追加の標的の発現を測定するために、2つの特注Array Plate(Array Plate番号3および4と称する)を構築した。以下の表1にArray Plate番号3および4に関する遺伝子リストを示す。

40

## 【表 1 - 1】

表1.ArrayPlateの遺伝子リスト

ArrayPlate番号3	
記号	GenBank 受託番号
SDHA	NM_004168
RPS6KB2	NM_003952
RPL37A	NM_000998
TFRC	NM_003234
ANT	NM_119937
MAGEA2	NM_005361
PAX3	NM_181457
CDK2	NM_001798
PRAME	NM_206953
MFI2	NM_005929
MCM6	NM_005915
S100B	NM_006272
PDIA4	NM_004911
SOX4	NM_003107
BRAF	NM_004333
PPIA	NM_021130
MAGED2	NM_014599
GALNTL1	NM_001168368
PTEN	NM_000314
HRAS	NM_005343
TP53	NM_000546
CTNNB1	NM_001904
TYR	NM_000372
TEX13A	NM_031274
BMP1	NM_001199
TGFB1	NM_000660
NR4A1	NM_002135
PIP4K2A	NM_005028
PDLIM7	NM_213636
TADA3L	NM_006354
B4GALT1	NM_001497
RAP2B	NM_002886

ArrayPlate番号4	
記号	GenBank 受託番号
SDHA	NM_004168
RPS6KB2	NM_003952
RPL37A	NM_000998
TFRC	NM_003234
ANT	NM_119937
BIRC7	NM_139317
BIRC5	NM_001168
MET	NM_001127500
HIF1A	NM_001530
ALK	NM_004304
DAZAP2	NM_014764
EVI2B	NM_006495
LDHA	NM_005566
ERCC1	NM_001983
ESR1	NM_000125
ALDOA	NM_000034
CTNNB1	NM_001904
ARID1A	NM_139135
NPHP1	NM_001128179
AF090940	AF090940
DUX4	NM_033178
POLR2J3	NM_001097615
HADHA	NM_000182
AK027225	AK027225
IGFBP5	NM_000599
BC017937	BC017937
OAZ1	NM_004152
TACSTD2	NM_002353
ATXN2L	NM_148416
PLIN2	NM_001122
PFDN6	NM_014260
HMGA1	NM_002131

10

20

30

40

【表 1 - 2】

表1.ArrayPlateの遺伝子リスト

ArrayPlate番号3	
記号	GenBank 受託番号
B2M	NM_004048
NCOR2	NM_001077261
SP100	NM_003113
SAT1	NM_002970
STAT2	NM_005419
RUNX1	NM_001001890
GNAS	NM_016592
SOCS3	NM_003955
BAX	NM_004324
CREBBP	NM_001079846
HIST1H2BN	NM_003520
HP1BP3	NM_016287
LZTS1	NM_021020
SQSTM1	NM_003900
TPSAB1	NM_003294

ArrayPlate番号4	
記号	GenBank 受託番号
ZFYVE16	NM_014733
AF168811	AF168811
BAX	NM_004324
AU159040	AU159040
BRD7P3	NR_002730
RNF126	NM_194460
ETV2	NM_014209
TPSAB1	NM_003294
ZFPL1	NM_006782
COX16	NM_016468
AK023563	AK023563
BEST1	NM_004183
PICALM	NM_001008660
NOP56	NM_006392
PTMS	NM_002824

10

20

## 【0268】

39個の母斑試料（まだ黒色腫になっていない患者（melanoma-naive patients）由来）および61個の原発性黒色腫試料からなる100個のFFPE組織切片で、mRNA発現を測定した。

## 【0269】

試料の調製および溶解、ヌクレアーゼ保護アッセイ、ならびにアレイ捕捉および検出を、実質的に実施例1のArrayPlate番号1および2で記載されたようにして行った。

## 【0270】

表2に、(i)この実施例で、母斑試料と黒色腫試料とで有意に差次的に発現されたことが見出された標的、および(ii)発現が正規化に使用された標的に関するNPP配列を示す。開示された発明において有用な他のNPP配列は他所に記載されており、またはこの開示で示された指針および開示された標的の公共的に利用可能な配列（例えば、表11および13に示した配列番号）を使用して当業者によって決定してもよい。

30

40

【表 2 - 1】

表2.例示的なヌクレアーゼ保護プローブ配列

遺伝子の 名称	受託番号	NPP配列 (5'から3';改行で折りたたむ)	配列 番号
B2M	NM_004048	CTGCTGGATGACGTGAGTAAACCTGAA TCTTTGGAGTACGCTGGATAGCC	1
B4GALT1	NM_001497	GTCTTGGAACCTGAGCCCAGGCTGGAC CTGGCAAAGGCGCTCAGTGGTAG	2
BMP1	NM_001199	CCGCAAGGTCGATAGGTGAACACAATA TAGCTGTCCTCGTCAGTGCGCTC	3
BRAF	NM_004333	GTAAGTGGAACATTCTCCAACACTTCC ACATGCAATTCTTCTCCAGTAAG	4
CDK2	NM_001798	CAAGTTCAGAGGGCCCACCTGAGTCCA AATAGCCCAAGGCCAAGCCTGGT	5
CREBBP	NM_001079846	CCTGGGTTGATACTAGAGCCGCTGCCT CCTCGTAGAAGCTCCGACAGTTG	6
CTNNB1	NM_001904	CAGCATCTGTGATGGTTCAGCCAAACG CTGGACATTAGTGGGATGAGCAG	7

【表 2 - 2】

表2.例示的なヌクレアーゼ保護プローブ配列

遺伝子の 名称	受託番号	NPP配列 (5'から3';改行で折りたたむ)	配列 番号
GALNTL1	NM_001168368	GGGCTCAGCTTGTCACTCTCCAGCTGGT TGAAGGCGTGCTGTCTGTAGGG	8
GNAS	NM_016592	CTCGCTGAGTCTTAGATTCCGCAGCCTA AGACTCGAGAGAGGTGCCTCCG	9
MAGEA2	NM_005361	CTCAGGCTCTCCACCTGGATGCTTGGCA GATCCTAGAACCACTGCATCTG	10
MAGED2	NM_014599	CTTCACCTTTCGGGCTTTCCTTGGCTTTG ACCTTGGGCCCAGTATCCTGAT	11
MCM6	NM_005915	TCCTGGTGTGCTAAGCTTGGAGACGTC AGGCACAACAATCAGTGTCCCTG	12
MFI2	NM_005929	GCTGGCATTGAAGAACTCGCTCACTGC TGTGAGGACGTCACAGTCCTTGG	13
NCOR2	NM_001077261	CCCGGTACAGCAGCGGGTACACAGCAC TCCGGGAGTGCCCTGGCTCCGTC	14
NR4A1	NM_002135	CGCCACAGCTGCCACGTGCTCCTTCAG GCAGCTGGCGATGCGGTTCTGCA	15
PDIA4	NM_004911	CACATCAAACCTGCTGGCCAGCACAGA CGCTGAGGTTGCATCGATCTTGG	16
PDLIM7	NM_213636	CTTCGATGTGTGTGAGGCTACCCGCATT CTCGCCATCGATGCTCAGCACC	17
PIP4K2A	NM_005028	ATTCACTCACTCACTCACTCACTCATT ATTCGGCCATAGCTGGAATCAA	18
PPIA	NM_021130	TGGTATCACCCAGGGAATACGTAACCA GACAACACACAAGACTGAGATGC	19
PRAME	NM_206953	GTCTGGCTGTGTCTCCCGTCAAAGGCTG CCATGAAGAGTGGCGGGAAGAG	20
PTEN	NM_000314	CTTCACCTTTAGCTGGCAGACCACAAA CTGAGGATTGCAAGTTCGCCAC	21
RAP2B	NM_002886	CCTCTCCTCCTGCTCCTTCATATGGTTC TCCCGGACTTCCCTCCATGTAT	22
RPL37A	NM_000998	CTGATGGCGGACTTTACCGTGACAGCG GAAGTGGTATTGTACGTCCAGGC	23
RPS6KB2	NM_003952	GCTTCACATACGTGGCGCCGTCTGTCT GGACAGCATCAAGGAGGGCTTC	124
RUNX1	NM_001001890	GCAGAGTCACACACATGCAAACACGCA CTCTTCGGAAGGCAGCCACTGTC	24
SAT1	NM_002970	ATTTCAAACATGCAACAACGCCACTGG TAATAAAGCTTTGGAATGGGTGC	25
SDHA	NM_004168	GAAGAAGCCCTTTGAGGAGCACTGGAG GAAGCACACCCTGTCCTATGTGG	123
SOCS3	NM_003955	GTCTTCTCTACCAGGAGCCTGAGGTGA AAGATGTCCCGTCTCCTCCATCC	26
SOX4	NM_003107	CTCCGCCTCTCGAATGAAAGGGATCTT GTCGCTGTCTTTGAGCAGCTTCC	27

10

20

30

40

【表 2 - 3】

表2.例示的なヌクレアーゼ保護プローブ配列

遺伝子の名称	受託番号	NPP配列 (5'から3';改行で折りたたむ)	配列番号
SP100	NM_003113	CCATGGTTGTGTAGCTCTGCCTCTGGGC TTTCTTCATCACAGGGCAACGG	28
SQSTM1	NM_003900	CCCAGGAAACATCAGCACACACACACA CAGGGACCCCTCCCTTCATGTCAC	29
STAT2	NM_005419	CGGGATTCAATCTCATGTTGCTGGCTCT CCACAGGTGTTTCGAGAACTGG	30
TADA3	NM_006354	CTACCCATCCAGCAGCTTCAGGATGCT CTCACGCTCCTTCAGAGTCTTCC	31
TEX13A	NM_031274	AGTATGAGTATGAGGCAGGGAGCTGGA CAGGAAGAGGTTCTGATGAGGCT	32
TFRC	NM_003234	GACGTGCTGCAGGGAAGTCCTCTCCTG GCTCCTCCCTCACTGGAGACTCG	33
TGFB1	NM_000660	GGTAGTGAACCCGTTGATGTCCACTTG CAGTGTGTTATCCCTGCTGTCAC	34
TP53	NM_000546	CCCGGGACAAAGCAAATGGAAGTCCTG GGTGCTTCTGACGCACACCTATT	35
TPSAB1	NM_003294	CGCCAGCAGCAGCAGATTCAGCATCCT GGCCGCTCCCTGTTCTTCTACC	36

10

20

## 【0271】

## データ分析

特に他の定めがない限り、実施例2および3の全ての分析はSASバージョン9.3で行われた。

## 【0272】

## A. 変換および品質管理

HTG OMI X (商標) イメージングデバイスを使用してデータを加工し、16ビットの画像を抽出した。ゲノム調査での標準的な実施と同様に、データのスケールをより線形にするために、生の強度値を、2を底とする対数で変換した。各遺伝子に3回の独立した観察を行い、(算術平均も同様に等しく役立つが)幾何平均を用いて3つ全ての観察を平均して、遺伝子ごとに2を底とする対数で示した発現値の複合平均値を作製した。植物遺伝子ANT (AP2様エチレン応答性転写因子; GenBankにおけるmRNAのRefSeq NM\_119937; 配列番号122)を、各アレイでの陰性対照として使用した。バックグラウンドを上回るANTが検出可能であった試料を使用して、アッセイの失敗をスクリーニングして除去した。また記述的な統計分析も行い、データファイル中の誤差に関してスクリーニングを行った。

30

## 【0273】

## B. 正規化のための遺伝子の選択

あらゆる遺伝子は、全ての試料タイプまたは被験体にわたりその発現において一定を保つ(すなわち、ユニバーサル「ハウスキーパー」遺伝子)という科学界の定説は、支持を失いつつある(例えば、Avison, Measuring Gene Expression, Psychology Press, 2007年、128頁)。したがって、特にマイクロアレイデータの正規化に好適な遺伝子を選択する他の選択肢が開発されている。いくつかの好適な方法が本明細書に記載されており、他の方法は当業者公知である。

40

## 【0274】

「正規化」遺伝子の発現を使用して、例えば試料量中の細胞の含量などの制御不可能なプロセス変数に対してデータを正規化した。この実施例に関して候補の正規化遺伝子ス

50

クリーニングする第一の工程は、S a t t e r t h w a i t e の t 検定を行って、対象の集団中の試料間で、すなわち母斑試料と原発性黒色腫試料との間でこのような候補の正規化遺伝子の発現に統計学的有意差がないことを決定することであった。まず、この分析を3連の生の平均発現値を使用して行い、その後、正規化された発現値を用いて確認した。有意性がないことを決定するために、0.05を超えるp値を下界として設定した。

【0275】

次いで候補の正規化遺伝子の発現レベルを検査して、十分な不飽和の強度値を確認した。十分な不飽和の強度値を、バックグラウンドより高く飽和未満の1.5の発現単位と定義した。

【0276】

最小の標準偏差に基づいて候補の正規化遺伝子も選択した。上界2.0の発現単位を、カットオフとして設定した。このカットオフよりも大きい標準偏差を有する候補の正規化遺伝子を考察から除外した。残った候補の正規化遺伝子の中から、対象の試料集団間で(すなわち、母斑試料と原発性黒色腫試料との間で)最も低い標準誤差を有するものを選択することを目的とした。

【0277】

上記の分析および他の適用可能な分析において、標準偏差の代わりに変動係数(CV)も使用できることに留意されたい。CVは、測定の単位に関わりなくデータまたは変数の分散を記載するための統計的方法である。CVは、標準偏差(またはいくつかのSAS法では二乗平均平方根誤差)を平均値で除算することにより計算され、アレイ全体の遺伝子の発現測定値の単位は、極めて類似していることから、CVまたは標準偏差を使用しても、ハウスキーパーの妥当性に関して異なる結論が引き出されるかもしれない定性的に異なる結果のパターンが生じることはほとんどない。

【0278】

表3に、これらの実施例全体にわたって選択された、ヒト母斑および原発性黒色腫の皮膚生検に代表的な例示的な正規化遺伝子(また「正規子」とも称される)を示す。図2Aおよび2Bに、代表的な正規子のBMP-1、MF12、NCOR2、およびRAP2bに関するボックスプロット、平均値のプロット、およびSASディフォグラム(diffogram)を示す。SASディフォグラムにおいて、例えば破線の斜線(左下から右上;青色)は $p = 0.5$ を示し;x軸およびy軸は、正規化された平均 $\log_2$ 強度値をプロットし;および各軸上の線は、示されたように各グループについての平均の正規化された平均 $\log_2$ 強度値を示す。実線の斜線(左上から右下;赤色)は破線の $p = 0.5$ の参照ラインと交わっており、これは、 $p > 0.05$ であり母斑試料と一次試料との間に統計学的有意差がないことを示す。

10

20

30

【表3】

表3.母斑試料および原発性黒色腫試料について同定された代表的な正規化遺伝子

記号	名称	GenBank参照番号	配列番号
BMP-1	ヒト骨形成タンパク質1、変異体1	NM_001199 (変異体1) NR_033404 (変異体5/nc); NR_033403 (変異体4/nc); NM_006129 (変異体3)	40-43
MF12	モノクローナル抗体133.2および96.5(MF12)により同定されたヒト抗原p97(黒色腫関連)	NM_005929 (変異体1) NM_033316 (変異体2)	37, 38
NCOR2	ヒト核内受容体コリプレッサー2	NM_001077261 (変異体2) NM_001206654 (変異体3); NM_006312 (変異体1)	44-46
RAP2b	ヒトRAP2B、RASがん遺伝子ファミリーの一種	NM_002886	39
RPS6KB2	ヒトリボソームタンパク質S6キナーゼ、70kDa、ポリペプチド2	NM_003952	120
SDHA	ヒトコハク酸デヒドロゲナーゼ複合体、サブユニットA、フラビントタンパク質(Fp)(SDHA)、ミトコンドリアタンパク質をコードする核遺伝子	NM_004168	121
RPL19	リボソームタンパク質L19	NM_000981	
RPLP0	大リボソームリンタンパク質P0	NM_001002 (変異体1); NM_053275 (変異体2)	
ALDOA	フルクトース-ニリン酸アルドラーゼA(別名、フルクトース-1,6-ニリン酸アルドラーゼA;ALDA;アルドラーゼ1;フルクトアルドラーゼA	NM_000034 (変異体1); NM_184041 (変異体2); NM_184043 (変異体3); NM_001127617 (変異体4) NM_001243177 (変異体6)	

## 【0279】

## C. 単変量の遺伝子スクリーニング

上記の正規子を用いてデータを正規化するために、各遺伝子の全ての繰り返しに関する平均の  $\log_2$  発現値を、BMP-1、MF12、NCOR2、およびRAP2b正規子(これはまた、「数種への正規化」としても当業者に公知であり、本明細書では「4種への正規化」と称される場合もある)の幾何平均で除算した。これまでに述べたように、算術平均も、上記の目的には十分である。得られた値に10の定数を乗算した。

## 【0280】

10

20

30

40

50

正規化の後、互いの（非正規子）遺伝子をスクリーニングして、その遺伝子の発現において母斑試料と原発性黒色腫試料との間に統計学的有意差があるかどうかを決定した。統計学的有意差は、その遺伝子が2つのグループを鑑別する何らかの能力を有することを示す。アルファ拡大と複数検定から保護するために、ボンフェローニ補正を使用してアルファの公称レベル（有意性に関するp値カットオフ）を選択した。S a t t e r t h w a i t eのt検定を使用して、各遺伝子を単変量の様式でスクリーニングした。S a t t e r t h w a i t eの補正されたt検定を使用して、グループ間の不等分散の場合における正確な推定を確認した。

【 0 2 8 1 】

表4に、mRNA発現において母斑試料と原発性黒色腫試料との間に統計学的有意差があることが見出された遺伝子のリストを示す。

【表4 - 1】

表4.母斑試料と原発性黒色腫試料との間で差違がある遺伝子

記号	名称	GenBank参照番号	P-値	配列番号
B2M	ヒトベータ-2-ミクログロブリン	NM_004048	<0.01	119
B4GALT1	ヒトUDP Gal:ベータGlcNAcベータ1,4-ガラクトシルトランスフェラーゼ、ポリペプチド1	NM_001497	<0.01	50
BRAF	ヒトv-rafマウス肉腫ウイルスがん遺伝子ホモログB1	NM_004333	<0.01	63
CDK2	ヒトサイクリン依存性キナーゼ2	NM_001798 (変異体1); NM_052827 (変異体2)	<0.01	112, 113
CREBBP	ヒトCREB結合タンパク質	NM_004380 (変異体1); NM_001079846 (変異体2)	<0.01	109, 110
CTNNB1	ヒトカテニン(カドヘリン関連タンパク質)、ベータ1	NM_001904	<0.01	83
GALNTL1	ヒトUDP-N-アセチル-アルファ-D-ガラクトサミン:ポリペプチドN-アセチル-ガラクトサミニルトランスフェラーゼ様1	NM_001168368 (変異体1); NM_020692 (変異体2)	<0.01	103, 104
GNAS	ヒトGNAS複合体遺伝子座	NM_000516 (変異体1); NM_080425 (変異体2); NM_080426 (変異体3); NM_016592 (変異体4); NM_001077488 (変異体6); NM_001077489 (変異体7); NR_003259 (変異体8/nc)	<0.01	85-91
MAGEA2	ヒト黒色腫抗原ファミリーA、2	NM_005361 (変異体1); NM_175742 (変異体2); NM_175743 (変異体3)	<0.01	105-107
MAGED2	ヒト黒色腫抗原ファミリーD、2	NM_014599 (変異体1); NM_177433 (変異体2); NM_201222 (変異体3)	<0.01	95-97
MCM6	ヒト微小染色体維持複合体(minichromosome maintenance complex)の構成要素6	NM_005915	<0.01	82

10

20

30

40

【表4 - 2】

表4.母斑試料と原発性黒色腫試料との間で差違がある遺伝子

記号	名称	GenBank参照番号	P-値	配列番号
NR4A1	ヒト核内受容体サブファミリー4、グループA、メンバー1	NM_002135 (変異体1); NM_173157 (変異体2); NM_001202233 (変異体3)	<0.01	47-49
PDIA4	ヒトタンパク質ジスルフィドイソメラーゼファミリーA、メンバー4	NM_004911	<0.01	81
PDLIM7	ヒトPDZおよびLIMドメイン7(エニグマ(enigma))	NM_005451 (変異体1); NM_203352 (変異体2); NM_213636 (変異体4)	<0.01	70-72
PIP4K2A	ヒトホスファチジルイノシトール-5-リン酸4-キナーゼ、II型、アルファ	NM_005028	<0.01	79
PPIA	ヒトペプチジルプロリルイソメラーゼA(シクロフィリンA)	NM_021130	<0.01	93
PRAME	黒色腫においてヒトで優先的に発現される抗原	NM_006115 (変異体1); NM_206953 (変異体2); NM_206954 (変異体3); NM_206955 (変異体4); NM_206956 (変異体5)	<0.01	98-102
PTEN	ヒトホスファターゼおよびテンシンホモログ	NM_000314	<0.01	94
RPL37A	ヒトリボソームタンパク質L37a(RPL37A)	NM_000998	<0.01	84
RUNX1	ヒトrunt関連転写因子1	NM_001754 (変異体1); NM_001001890 (変異体2); NM_001122607 (変異体3)	<0.01	66-68
SAT1	ヒトスペルミジン/スペルミンN1-アセチルトランスフェラーゼ1	NM_002970 (変異体1) NR_027783 (変異体2/nc)	<0.01	51, 52
SOCS3	サイトカインシグナル伝達のヒトサプレッサー3	NM_003955	<0.01	69
SOX4	ヒトSRY(性決定領域Y)-ボックス4	NM_003107	<0.01	80

10

20

30

40

【表 4 - 3】

表4.母斑試料と原発性黒色腫試料との間で差違がある遺伝子

記号	名称	GenBank参照番号	P-値	配列番号
SP100	ヒトSP100核抗原	NM_001080391 (変異体1); NM_003113 (変異体2); NM_001206701 (変異体3); NM_001206702 (変異体4); NM_001206703 (変異体5); NM_001206704 (変異体6)	<0.01	73-78
SQSTM1	ヒトシーケストソーム (sequestosome)1	NM_003900 (変異体1); NM_001142298 (変異体2); NM_001142299 (変異体3)	<0.01	116-118
STAT2	ヒトシグナルトランスデューサ ーおよび転写活性化因子2	NM_005419 (変異体1); NM_198332 (変異体2)	<0.01	114, 115
TADA3	ヒト転写アダプター3	NM_006354 (変異体1); NM_133480 (変異体2)	<0.01	61, 62
TEX13A	発現されたヒト精巣13A	NM_031274	<0.01	108
TFRC	ヒトトランスフェリン受容体 (p90、CD71)	NM_003234 (変異体1); NM_001128148 (変異体2)	<0.01	64, 65
TGFB1	ヒトトランスフォーミング増殖 因子、ベータ1	NM_000660	<0.01	92
TP53	ヒト腫瘍タンパク質p53	NM_000546 (変異体1); NM_001126112 (変異体2); NM_001126114 (変異体3); NM_001126113 (変異体4); NM_001126115 (変異体5); NM_001126116 (変異体6); NM_001126117 (変異体7); NM_001126118 (変異体8)	<0.01	53-60
TPSAB1	ヒトリプターゼアルファ/ベ ータ1	NM_003294	<0.01	111

10

20

30

40

【0282】

母斑を示す「0」と原発性黒色腫を示す「1」とでバイナリコードのダミー変数である疾患変数で正規化されたデータに関する共分散マトリックスを作製した。表5に、提示された各遺伝子の発現がどのように疾患変数と共変動するかを示す。

【表5】

表5.疾患の共分散

記号	共分散対 疾患変数	記号	共分散対 疾患変数	記号	共分散対 疾患変数
B2M	0.1253	NR4A1	0.766	SOX4	0.1853
B4GALT1	0.2552	PDIA4	0.165	SP100	0.203
BRAF	0.3014	PDLIM7	0.1693	SQSTM1	-0.092
CDK2	0.146	PIP4K2A	0.4079	STAT2	0.0731
CREBBP	0.086	PPIA	0.1539	TADA3	0.184
CTNNB1	0.205	PRAME	0.3603	TEX13A	0.2098
GALNTL1	0.1324	PTEN	0.164	TFRC	0.2609
GNAS	0.205	RPL37A	-0.409	TGFB1	0.1114
MAGEA2	0.3195	RUNX1	0.2568	TP53	0.265
MAGED2	0.129	SAT1	0.4122	TPSAB1	0.188
MCM6	0.2381	SOCS3	0.4256		

10

20

## 【0283】

表5で示されるように、表4における各遺伝子についての平均発現値は、その平均値が原発性黒色腫と比較すると母斑のほうが高いRPL37AおよびSQSTM1（陰性値）を除いて、母斑よりも原発性黒色腫において高い（陽性値）。言い換えれば、表4の遺伝子は、指摘した場合を除いて、それらの母斑における発現と比較して原発性黒色腫でアップレギュレートされる傾向がある。

## 【0284】

これらの遺伝子を個々にまたは組み合わせて使用することにより、検定試料を良性母斑または原発性黒色腫として特徴付けること（例えば診断すること）を可能にする予測モデル（例えば、回帰モデル、またはより具体的な実施例では線形回帰モデル）が得られる。開示された方法、アレイまたはキットで使用するための例示的で非限定的な遺伝子の組合せは、MAGEA2、PRAME、PDIA4、NR4A1、PDLIM7、B4GALT1、SAT1、RUNX1、および/またはSOCS3のうち少なくとも2種、3種、4種、5種、6種、7種、8種、または9種全てである。

30

## 【0285】

全体的な有意性に加えて、32種の遺伝子のセットの中からモデルの組合せを選択する場合、予測モデルを形成するために、いくつもの尺度を使用して、どの遺伝子が対になっているのか、またはうまく一緒に組み合わされているのかの決定に役立てた。1つの具体的な方法は、あるモデルにおいて各Xn変数の遺伝子の分散拡大要因（VIF）によって測定した場合、そのモデルにおける予測子（すなわち、Xn変数；以下を参照）間の多重共線性を最小化することであった。全ての予測子Xn変数が10未満の分散拡大要因（VIF）を有する表4、11、および/または13における遺伝子（例えばmRNAまたはmiRNA）のあらゆる組合せは、良性母斑からの試料と原発性黒色腫からの試料とを鑑別するのに有用な予測値を有する可能性があることから、本開示で想定される。

40

## 【0286】

## D. ロジスティック回帰モデル

表4の遺伝子を使用して統計予測モデルを開発するのに使用される原理は、二極分布お

50

よびロジットリンク関数を用いたロジスティック回帰であった。モデルの推定を、フィッシャーのスコア付けを使用して行った。しかしながら、正確なロジスティック回帰、実験的なサンドイッチ推定量もしくは他の補正されたバイアス、安定化された分散、または別の方式による補正推定技法を用いて推定されたモデルも、多くの環境下では、わずかに異なるパラメーター推定を生じつつも定性的には一貫した結果のパターンを生じる類似のモデルをもたらす。同様に、これらに限定されないが、累積ロジット、相補的な両対数、プロビットまたは累積プロビットなどの他のリンク関数も、同じ定性的な結果パターンをもたらす予測モデルを生じると期待され得る。

この実施例におけるモデル（アルゴリズム）の一次形態は：

$$\text{Logit}(Y_i) = 0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 \cdots \beta_n X_n \quad 10$$

であり、式中  $0$  は、切片項であり、 $\beta_n$  は、係数推定値（coefficient estimate）であり、 $X_n$  は、所定の遺伝子に関する  $2$  を底とする対数の発現値である。典型的には、全ての  $\beta_n$  の値は、 $-1$ 、 $0$  より大きく  $1$ 、 $0$  より小さい。多くの場合、 $0$  切片項は、 $-2$  より大きく  $2$  より小さく、いくつかの場合においては  $-1$  より大きく  $1$  より小さい。追加の  $\beta_n$  は、 $n > 0$  の場合、 $-1$  より大きく  $1$  より小さい可能性がある。

【0287】

モデル性能を検証するためにいくつかの検定を行った。ワルドのカイ二乗検定が使用されたが、この検定は、全てのモデル当てはめに関して統計学的に有意な結果を示すのに必要であった。ホスマーおよびレメシヨウの不適合度検定は、モデルに関して統計学的に検出できない不適合度を示すのに必要であった。モデルにおける各遺伝子の予測子は、 $p < 0.05$  の有意性であることが必要であった。

【0288】

多数の交差検証法を使用して、結果の再現性を確認した。一次的な方法は、SAS Procロジスティックの分類表法の一部として実施された1工程での最尤推定値の近似であった。また追加の結果確認のためのオープンソースパッケージのWekaで10回の交差検証および66～33%のスプリット検証も行った。この実施例においてロジスティック回帰は数学的基盤であるが、他の統計学的手法、数学的手法、およびデータマイニング手法（プロビット回帰、サポートベクターマシンまたはクラスタリングアルゴリズムなど）は、同じ定性的な結果パターンをもたらすモデルを作製できる。

【0289】

本実施例においてデータにロジスティック回帰モデリングを適用して、以下の表6に、母斑試料と原発性黒色腫試料とを正確に鑑別する非限定的な遺伝子の組合せと、対応する予測アルゴリズムに関する値とを示す。

10

20

30

【表 6 - 1】

表6.アルゴリズム値を用いた(4種に対し正規化された)例示的な予測の組合せ

実施形態	遺伝子の組合せ (左から右に、各遺伝子は、 $X_1, X_2 \dots X_n$ を示し、当てはまる場合、このアルゴリズムにおいて:出力= $\beta_0 + \beta_1 X_1 + \beta_2 X_2 \dots \beta_n X_n$ である)	回帰係数および切片( $\beta_0$ )
B1	N4RA1, B4GALT1 一例として: 出力= $-59.0958 + 1.5998(NR4A1) + 4.2115(B4GALT1)$	$\beta_0 = -59.0958$ $\beta_1 = 1.5998$ $\beta_2 = 4.2115$
B2	NR4A1, SOX4	$\beta_0 = -39.1063$ $\beta_1 = 2.0554$ $\beta_2 = 1.8234$
B3	NR4A1, SOX4, B4GALT1	$\beta_0 = -75.3582$ $\beta_1 = 1.9674$ $\beta_2 = 1.5141$ $\beta_3 = 4.0622$
B4	NR4A1, SOX4, SQSTM1, B2M	$\beta_0 = -34.8327$ $\beta_1 = 2.2925$ $\beta_2 = 2.2998$ $\beta_3 = -3.2193$ $\beta_4 = 2.1559$
B5	MAGED2, SAT1, SOX4	$\beta_0 = -49.3358$ $\beta_1 = -0.2919$ $\beta_2 = 3.0513$ $\beta_3 = 2.3171$
B6	N4RA1, BRAF	$\beta_0 = -43.4593$ $\beta_1 = 2.1785$ $\beta_2 = 2.3159$
B7	NR4A1, RPL37A	$\beta_0 = -9.6524$ $\beta_1 = 3.2965$ $\beta_2 = -2.1656$

10

20

30

40

【表 6 - 2】

表6.アルゴリズム値を用いた(4種に対し正規化された)例示的な予測の組合せ

実施形態	遺伝子の組合せ (左から右に、各遺伝子は、X1、X2...Xnを示し、当てはまる 場合、このアルゴリズムにおいて:出力= $\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n$ である)	回帰係数および 切片( $\beta_0$ )
B8	NR4A1, SQSTM1, TPSAB1	$\beta_0 = -7.6589$ $\beta_1 = 2.7873$ $\beta_2 = -3.6387$ $\beta_3 = 1.6122$
B9	NR4A1, TFRC, SAT1	$\beta_0 = -43.3177$ $\beta_1 = 1.5862$ $\beta_2 = 0.111$ $\beta_3 = 2.354$
B10	TFRC, SAT1	$\beta_0 = -40.4475$ $\beta_1 = 1.3975$ $\beta_2 = 2.5618$
B11	SOCS3, TFRC, BRAF	$\beta_0 = -42.6409$ $\beta_1 = 1.7603$ $\beta_2 = -0.032$ $\beta_3 = 3.048$
B12	SOCS3, TFRC	$\beta_0 = -34.9546$ $\beta_1 = 1.1505$ $\beta_2 = 2.4557$
B13	SOCS3, SOX4, SAT1, BRAF	$\beta_0 = -93.1404$ $\beta_1 = 2.3007$ $\beta_2 = 0.9978$ $\beta_3 = 4.3683$ $\beta_4 = 1.8052$

10

20

30

40

## 【0290】

図3～7は、試料が母斑または原発性黒色腫として適切に特徴付けられる(例えば、診断される)かどうかを正確に決定するためにこの実施例に記載されたNR4A1およびB4GALT1の組合せを使用したモデルの特定の結果を示す。

## 【0291】

表6で開示されたアルゴリズムを使用して、試験用FFPE皮膚生検を原発性黒色腫または母斑として特徴付けた。アルゴリズム(別名、当てはめモデル)により予測事象の確率が提供されるが、この確率は、この実施例では試料が原発性黒色腫である確率(probably)である。当業者公知のSAS計算法を使用して、予測確率のバイアス縮小推定量を計算した(support.sas.com/documentation/cd

50

l/en/statug/63347/HTML/default/viewer.htm#statug\_logistic\_sect044.htmを参照(2012年6月22日現在)。

#### 【0292】

一連の閾値  $z$  (ここで  $z$  は 0 から 1 の間である) を設定した。特定の試料に関して計算された予測確率が予め設定された閾値  $z$  を超えるかまたはそれに等しい場合、試料を原発性黒色腫グループに割り当て、そうでなければ母斑グループに割り当てた。次いでそれぞれのグループへの割り当てを既知の臨床データに照合確認して、中でも特に真陽性、真陰性、偽陽性、および偽陰性を決定した。これらの結果を、例えば分類表(例えば図5および6に記載の分類表)に示す。

10

#### 【0293】

閾値は、当業者によりモデルの望ましい臨床的有用性に基づいて決定できる。図5および6において、代表的なNR4A1およびB4GALT1(4種に対し正規化された)モデルを使用してこの点を実証する。より高い閾値は、原発性黒色腫とコールされるように設定できる。例えば、図5で強調表示された領域を参照されたい(確率レベル0.700~0.780)。この場合、偽陽性率は比較的低い、または特に明記されていなければ、感度よりも高い特異度が存在した。逆に言えば、試料を原発性黒色腫とコールするための閾値は低くし得る; 例えば図6で強調表示された領域を参照(確率レベル0.240~0.620)。その閾値レベルにおいて、モデルは、より少ない偽陰性またはより高い感度を与えるが、特異度は幾分より低かった。

20

#### 【0294】

表6および8に提供されたモデル方程式からの「出力」は、ロジットであることに留意されたい。ロジットは、試料が原発性黒色腫であることに関するオッズ比の対数である。それに代わる方法として、当業者は、試料を原発性黒色腫または母斑とコールするための閾値としてロジットを使用してもよい。どちらの方法でも同等の結果が得られる。例えば、0.5の予測確率を選択することは、0のロジットを使用した場合の結果と同じになる。これは、1のオッズ比は試料が原発性黒色腫である確率または母斑である確率が等しいと解釈されるため、理解が可能である。1の対数はゼロに等しいため、ゼロのロジットは1のオッズ比に等しい。1のオッズ比またはゼロのロジットは、試料が原発性黒色腫である確率または母斑である確率が50/50であることを考えれば、0.5の予測確率は、同等の結果である。したがって当てはめモデルまたはロジットの出力からの予測確率のいずれかを使用することは、同じ結果をもたらす。

30

#### 【0295】

(実施例3)

2つの代替の正規子を用いた分析により、母斑か原発性黒色腫かを予測することに関するモデルの頑健性を実証する

さらに実施例2に記載された予測(例えば、診断)遺伝子の組合せの頑健性を、データを正規化するために正規化遺伝子の代替のセットを使用することにより示した。この実施例で示される中でも特に、このような分析は、[N4RA1、B4GALT1]予測モデルに対して有意な効果を有さず、この結果は、実施例2に記載された全ての予測モデルの代表例であると考えられる。

40

#### 【0296】

実施例1および2で推定の「ハウスキーパー」遺伝子として記載されたアレイには、RPL37A、RPS6KB2、SDHA、およびTFRCが含まれていた。上記で論じられたように、「ハウスキーパー」遺伝子(例えば、その発現が試料タイプ全てにおいて変わらない遺伝子)の概念は支持を失いつつあるため、実際に試験すべきである(試験した)。

#### 【0297】

これらの候補の「ハウスキーパー」遺伝子それぞれについての複合的な2を底とする平均対数発現値(実施例2を参照)を、他の候補の「ハウスキーパー」遺伝子のそれぞれに

50

ついでにそのような値に正規化した。各正規化「ハウスキーパー」についての変動係数（CV）および標準偏差（SD）を計算したが、それは表7に示す通りである。

【0298】

【表7】

表7:各正規化「ハウスキーパー」に関する変動係数(CV)および標準偏差。

候補の「ハウスキーパー」	CV	SD
RPL37A	0.101384	0.894
RPS6KB2	0.023556	0.365
SDHA	0.022369	0.339
TFRC	0.070521	0.719

10

【0299】

候補の「ハウスキーパー」のCVが他の候補の最も低いCVのCVよりも2倍超大きい場合、その候補の「ハウスキーパー」を合理的な正規子として廃棄した。したがって、この実施例ではRPS6KB2およびSDHAだけを正規子として選択した。

【0300】

この代表的なヒト母斑および原発性黒色腫の皮膚生検に関する2つの正規化遺伝子のセットに関する一般的な情報は、上記の表3に記載されている。

20

【0301】

表4の遺伝子それぞれに関する複合平均の2を底とする対数発現値を、RPS6KB2およびSDHAに関する複合平均の2を底とする対数発現値に対して正規化し、実施例2に記載したロジスティック回帰分析を、表6に示した予測の遺伝子の組合せのそれぞれについて繰り返した。

【0302】

以下の表8に、2種（すなわちRPS6KB2およびSDHA）に対して遺伝子（X<sub>n</sub>）を正規化した予測アルゴリズムそれぞれについての対応する切片および係数を示す。

30

【表 8 - 1】

表8.アルゴリズム値を用いた(2種に対し正規化された)例示的な予測の組合せ

実施形態	遺伝子の組合せ (左から右に、各遺伝子は、X1、X2・・・Xnを示し、当てはまる場合、このアルゴリズムにおいて:出力= $\beta_0 + \beta_1 X_1 + \beta_2 X_2 \dots + \beta_n X_n$ である)	回帰係数および切片( $\beta_0$ )
B1	NR4A1, B4GALT1	$\beta_0 = -39.9861$ $\beta_1 = 1.9964$ $\beta_2 = 2.1807$
B2	NR4A1, SOX4	$\beta_0 = -25.7153$ $\beta_1 = 2.1472$ $\beta_2 = 0.4994$
B3	NR4A1, SOX4, B4GALT1	$\beta_0 = -39.4785$ $\beta_1 = 1.9795$ $\beta_2 = -0.4937$ $\beta_3 = 2.5948$
B4	NR4A1, SOX4, SQSTM1, B2M	$\beta_0 = -12.6489$ $\beta_1 = 2.5444$ $\beta_2 = 0.6808$ $\beta_3 = -3.7649$ $\beta_4 = 1.9304$
B5	MAGED2, SAT1, SOX4	$\beta_0 = -26.6321$ $\beta_1 = -2.5752$ $\beta_2 = 4.2547$ $\beta_3 = 1.3862$
B6	NR4A1, BRAF	$\beta_0 = -38.9049$ $\beta_1 = 2.2051$ $\beta_2 = 2.1615$
B7	NR4A1, RPL37A	$\beta_0 = -2.9081$ $\beta_1 = 3.0926$ $\beta_2 = -2.4906$
B8	NR4A1, SQSTM1, TPSAB1	$\beta_0 = -13.1057$ $\beta_1 = 3.004$ $\beta_2 = -3.7264$ $\beta_3 = 2.0336$

10

20

30

40

【表 8 - 2】

表8.アルゴリズム値を用いた(2種に対し正規化された)例示的な予測の組合せ

実施形態	遺伝子の組合せ (左から右に、各遺伝子は、X1、X2...Xnを示し、当てはまる場合、このアルゴリズムにおいて:出力= $\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n$ である)	回帰係数および切片( $\beta_0$ )
B9	NR4A1, TFRC, SAT1	$\beta_0 = -62.1051$ $\beta_1 = 1.6769$ $\beta_2 = 2.2591$ $\beta_3 = 2.6097$
B10	TFRC, SAT1	$\beta_0 = -69.1937$ $\beta_1 = 3.8901$ $\beta_2 = 3.6063$
B11	SOCS3, TFRC, BRAF	$\beta_0 = -59.3456$ $\beta_1 = 1.4668$ $\beta_2 = 2.4699$ $\beta_3 = 3.1692$
B12	SOCS3, TFRC	$\beta_0 = -35.5872$ $\beta_1 = 0.9893$ $\beta_2 = 3.0369$
B13	SOCS3, SOX4, SAT1, BRAF	$\beta_0 = -108.5$ $\beta_1 = 4.1954$ $\beta_2 = -7.3183$ $\beta_3 = 6.3842$ $\beta_4 = 8.8727$

10

20

30

## 【0303】

表8に開示されたアルゴリズムを決定して、実施例2に記載されているようにして検定した。

## 【0304】

図8および9は、(i)全体的なB4GALT1およびNR4A1(2-正規子)のモデル当てはめは高い有意性を保つこと；(ii)AUCは、97.67であること(これは、例えば、B4GALT1およびNR4A1(2-正規子)予測モデルが、無作為に選択された陽性例を無作為に選択された陰性例よりも高くランク付けする確率が97.67%であることを示す)；(iii)SAS交差検証後に起こり得る閾値の約50%を超える90%以上の正しい分類；(iv)SAS交差検証後に起こり得る閾値の範囲全体にわたり低下した特異度；および(v)これと同じ分子署名に関する実施例2のモデルと比較して、SAS交差検証後の最大感度における適度な増加を示す。

40

## 【0305】

まとめると、この実施例は、試料を分子レベルで母斑または原発性黒色腫として特徴付

50

けるためのB4GALT1およびNR4A1モデルは、あらゆる正規化方法で繰り返すことができることを実証する。類似の計算方法を使用した実施例2および3に(および本明細書の他所に)記載した他の予測署名について類似の結果が期待される。それによれば、検定試料を母斑または原発性黒色腫として特徴付けることに関するこれらの代表的なモデルおよび他の開示されたモデルの頑健性および有用性が明確に実証される。

#### 【0306】

(実施例4)

機械学習法を使用した黒色腫および母斑の分類

この実施例では、Array Plate番号3および4(実施例2および3を参照)を含む4つのArray Plateのセットを使用して、ヒト被験体から生検によって得られたFFPE皮膚試料の第三のセットにおけるmRNAおよびmiRNA発現を決定した。このようなデータを使用して、黒色腫および母斑試料で有意に差次的に発現されたmRNAおよびmiRNAのセットをうまく同定し、機械学習(例えば、ランダムフォレスト(Breiman, Machine Learning, 45(1): 5~32頁(2001年))黒色腫-母斑分類器をトレーニングした。

10

#### 【0307】

それぞれ厚さおよそ5 $\mu$ mであり顕微鏡スライド上に載せられた115個のFFPE組織切片のセットは、John Wayne Cancer Institute(JWCI)組織バンクから提供された。このセットには、56個の母斑試料および59個の原発性黒色腫試料が含まれていた。

20

#### 【0308】

4つのArray Plateのセットにおける181種の(対照を含む)mRNAまたはmiRNAの発現に関して試料を分析した。Array Plate番号3~5(mRNAのみ)に関するアッセイおよび検出方法は、実質的に実施例1に記載された通りであった。Array Plate番号6(miRNAおよびmRNAの共検出)に関するアッセイおよび検出方法は、実質的にPCT公報WO2013/049231に記載された通りであった。Array Plate番号6の正規子および陰性対照(ANT)mRNAに関するNPPは、これらの実施例の他所に記載したそれぞれの対照NPPの最も3'側の25ヌクレオチドに対応する25merであった。Array Plateのセットは、Array Plate番号3および4(表1を参照)、mRNA標的に特異的な「Array Plate番号5」、ならびにmiRNA標的(加えてmRNA対照)に特異的な「Array Plate番号6」を包含していた。表9に、Array Plate番号5および6に関する標的の列挙を示す。Array Plate番号3~6上にいくつか共通のmRNA標的が存在したため、101種の異なるmRNA(対照を含む)および42種の異なるmiRNAに関してデータを集めた。

30

【表 9 - 1】

表9.ArrayPlate標的のリスト

ArrayPlate番号5	
識別子	GenBank受託番号
MAGEA2	NM_005361
PAX3	NM_181457
GALNTL1	NM_001168368
MAGEA1	NM_004988
PanMAGEA3-12	
BIRC7	NM_139317
BIRC5	NM_001168
XIAP	NM_001167 (v1); NM_001204401 (v2); NR_037916.1 (非コード v3)
PRAME	NM_206953
MET	NM_001127500
MFI2	NM_005929
MCAM	NM_006500 (GI:71274106)
BAD	NM_004322 (v1); NM_032989 (v2)
BCL2	NM_000633 (アルファ); NM_000657 (ベータ)
HIF1A	NM_001530
MIB1	NM_020774
TOP2A	NM_001067
WT1	例えば NM_000378(変 異体A)またはさらに他 の変異体(例えばB-F)
MCM2	NM_004526
MCM6	NM_005915
ALK	NM_004304
S100B	NM_006272
PDIA4	NM_004911

ArrayPlate番号6	
識別子	GenBankまたは miRBase*識別子 (当てはまる場合)
SDHA	NM_004168
RPS6KB2	NM_003952
RPL37A	NM_000998
TFRC	NM_003234
ANT	NM_119937
23b	hsa-miR-23b
211	hsa-miR-211
1224-3p	hsa-miR-1224-3p
193A-5P	hsa-miR-193a-5p
146A	hsa-miR-146a
513b	hsa-miR-513b
133A	hsa-miR-133a
182	hsa-miR-182
205	hsa-miR-205
665	hsa-miR-665
1254	hsa-miR-1254
200C	hsa-miR-200c
292	mmu-miR-292-3p
200A	hsa-miR-200a
21	hsa-miR-21
140-3p	has-miR-140-3p
140-5p	has-miR-140-5p
29C	hsa-miR-29c

10

20

30

40

【表 9 - 2】

表9.ArrayPlate標的のリスト

ArrayPlate番号5		ArrayPlate番号6		
識別子	GenBank受託番号	識別子	GenBankまたは miRBase*識別子 (当てはまる場合)	
SOX4	NM_003107	142-5P	hsa-miR-142-5p	10
XRCC5	NM_021141	595	hsa-miR-595	
DAZAP2	NM_014764	207	mmu-miR-207	
EVI2B	NM_006495	106a	hsa-miR-106a	
LDHA	NM_005566	122	hsa-miR-122	
BRAF	NM_004333	1304	hsa-miR-1304	
ERCC1	NM_001983	155	hsa-miR-155	
ESR1	NM_000125	191	hsa-miR-191	20
RPL19	NM_000981	375	hsa-miR-375	
SDHA	NM_004168	612	hsa-miR-612	
ALDOA	NM_000034	650	hsa-miR-650	
RPLP0	NM_001002; NM_053275	1180	hsa-miR-1180	
PPIA	NM_021130	183	hsa-miR-183	
ANT	他の実施例を参照	203	hsa-miR-203	
MAGEB1	NM_002363 (v1); NM_177404 (v2); NM_177415 (v3);	1293	hsa-miR-1293	30
MAGEC2	NM_016249	342-3p	hsa-miR-342-3p	
MAGED2	NM_014599	1294	hsa-miR-1294	
		19b	hsa-miR-19b	
		557	hsa-miR-557	
		1198-5p	mmu-miR-1198-5p	40
		let-7a	hsa-let-7a	
		1291	hsa-miR-1291	
		29b	hsa-miR-29b	
		150	hsa-miR-150	

\*KozomaraおよびGriffiths-Jones, Nuc.Acids Res., 39(データベース版):D152(2011)

【 0 3 0 9 】

正規化遺伝子は、ArrayPlate番号3、4、および6上ではSDHAおよびRPS6KB2であり、ArrayPlate番号5上ではSDHA、RPL19、RPL

P0、およびALDOAであった(表3も参照)。これらの正規子のうち、これらの実施例の他所に記載したように、全試料タイプで何らかの有意差を示したものはなかった。

利用できる試料に限界があり、各アレイで全ての試料を処理できたわけではないため、全ての生データは以下のような厳重な品質管理(すなわち予備的な加工)を受けた:生データからバックグラウンドを減算し、 $\log_2$ 変換した。陰性対照遺伝子ANTに関して200より大きいRLUが測定されたあらゆる試料を失敗とみなし、それらの特定のウェルからの全てのデータをさらなる考察から除外した。各遺伝子での繰り返しの発現値に関して変動係数(CV)を決定した。試料の繰り返しのCVが8%を超える場合、平均から最も遠い繰り返しを外れ値として除外した。繰り返しの再現性をペアワイズ相関およびペアワイズ線形単回帰によって測定した。相関が $r = 0.90$ を示し、線形回帰の切片がゼロとは統計学的に有意差がない場合、このような繰り返しは認められ、その他の場合は失敗とみなされた。2つより多くの失敗した繰り返しを示した全ての試料を、失敗試料と定義した。品質標準を満たさなかったデータを分析から除外した。表10に要約を示す。

【表10】

表10:分析された試料および遺伝子の要約

ArrayPlate 番号	標的番号	試料のクラス	試料の処理 回数	データQC後の 試料数
3	47	黒色腫	59	57
		母斑	56	54
4	47	黒色腫	59	53
		母斑	40	34
5	40	黒色腫	58	46
		母斑	35	32
6	47	黒色腫	59	59
		母斑	50	50

【0310】

単変量解析

加工データのいくつかの単変量解析(例えば、対数倍率変化、2試料のt検定(偽発見率(FDR)調整p値)、およびAUCロジスティック回帰分析)を行って、各データセットにおいて試料タイプ間で特定の遺伝子が有意に差次的に発現されたのかどうかを評価した。

【0311】

図10Aに3つのmRNAアレイ(すなわちArrayPlate3~5)に関する単変量解析の結果を示す。行われたそれぞれ3つの検定の値((i)受診者動作特性(receiver operating characteristic)(ROC)曲線下面積(AUC)、(ii)対数倍率変化(fch)、および(iii)2試料のt検定(FDR調整p値;FDR.pvalue))を示す。x軸に発現データが集められた遺伝子を示し、y軸にそれぞれの単変量統計の値を示す。AUC分析については、より高い値が望ましい。この場合、0.75(点線で)を、統計的有意性のカットオフとして割り当てた。そのラインを超えるAUCを有する遺伝子が、母斑と黒色腫とを識別する候補である。対数倍率変化分析については、陰性1および陽性1(これらはそれぞれ、母斑と黒色腫との間の発現において2倍の差があることに等しい)を、統計的有意性のカットオフとして割り当てた。陽性1より大きく陰性1より小さい対数倍率変化を有する遺伝子は、母斑と黒色腫とを識別する候補である。FDR調整p値については、より低い値が望ましく、0.05(点線で)を、統計的有意性のカットオフとして割り当てた。そのラインより

下の発現を示す遺伝子が、母斑と黒色腫とを識別する候補である。

【0312】

図10BにmiRNAアレイに関する単変量解析の結果を示す。上述したように、x軸に列挙したmiRNAのそれぞれについてROC曲線より下のAUC、対数倍率変化、および2試料のt検定(FDR調整p値)を決定した。それぞれの単変量統計の値をy軸に示す。「HK」と標識したデータは、ArrayPlate番号6において「ハウスキーパー」に対し正規化されたmiRNA発現データを示し、「NO.HK」は、正規化されていないmiRNA発現データを示す。正規化されたデータと正規化されていないデータは一般的に類似の結果を提供する。AUC、倍率変化、および2試料のt検定(FDR調整p値)についての統計のカットオフは、上記のmRNA分析の場合と同じであった。0.75より大きいAUC、陽性1より大きい倍率変化、およびカットオフラインより下のFDR調整p値を有するmiRNAが、母斑と黒色腫とを識別する候補である。

10

【0313】

単変量解析の陽性の結果(すなわち、黒色腫と母斑との間で有意に差次的に発現されたmRNAおよびmiRNAの同定)は、より多くの資源を消費する(more resource-consuming)多変量解析を用いて進められた決定と、さらなる黒色腫-母斑分類器の開発とを支持した。

【0314】

多変量解析

次いで多変量解析を行って、検出された標的のどの部分集合が(統計的な(statistical)見地から)最も強く黒色腫試料タイプと母斑試料タイプとを識別したかを決定した。複数の特徴選択法(RF、LIMMA、t検定、AUC)を使用して、各データセットにおいて試料タイプ間で特定の遺伝子が有意に差次的に発現されたかどうかを評価した。機械学習アルゴリズム(例えば、ロジスティック回帰(LR)、ランダムフォレスト(RF)、サポートベクターマシン(SVM)、K-近傍法(KNN))を使用して、最初の分類器を開発した。特徴選択および分類性能の両方を一つ抜き(leave one out)交差検証法で評価した。遺伝子数の関数としての誤り率と受診者動作特性(ROC)曲線を使用して、分類器の性能を評価した。

20

【0315】

ArrayPlate番号3で検出された遺伝子について、図11は、そのアレイにおいて上位2種(GN=2)から上位40種(GN=40)までの遺伝子に基づく分類器のAUC性能を示す。このタイプの分析について、検定した分類器のより高い感度(すなわちy軸に示される真陽性率)およびより低い偽陽性率(すなわちx軸に示される「1-特異度」)に伴いAUCは増加する。この図は、AUCが0.93を超える場合(1.00は「全部」である)、全てのArrayPlate番号3の分類器が12種よりも多い遺伝子を有することを実証する。当業者であれば、この結果は、本明細書において開示された情報に基づいて12種よりも少ない遺伝子を有する分類器を選択できないということの意味するのではないが、このような分類器は、それほど高い感度および特異度を有さない可能性があることが理解されよう。いくつかの設定では、高い感度または高い特異度が最優先事項ではない場合もあり、その結果として分類器を選択してもよい。例えば、黒色腫-母斑分類器にとって、母斑を黒色腫と誤同定するのではなく黒色腫を母斑と誤同定することのほうがより悪い結果とみなされる可能性がある。その場合、偽陰性を最小化しつつ偽陽性に関してはもう少し厳密でないように(帰無仮説が黒色腫であるかまたは黒色腫ではない場合)、分類器を選択してもよい。

30

40

【0316】

ArrayPlate番号3~6それぞれにおける遺伝子について、図12は、様々な統計的方法(すなわち、AUC、t検定、ランダムフォレスト、LIMMA)によって決定された、分類器中の遺伝子の数の関数としての分類器の誤り率を示す。この図は、黒色腫-母斑分類器が約10種またはそれよりも多い遺伝子を有する場合、その黒色腫-母斑分類器の誤分類のエラーが最小化されることを示す。上記したように、この結果は、本明細

50

書において開示された情報に基づいて約10種よりも少ない遺伝子を有する分類器を選択できないということを意味するのではないが、このような分類器は、誤りを起こす可能性がより高い。

【0317】

Array Plate 番号3~5のそれぞれについて、ランダムフォレストアルゴリズムの一つ抜き交差検証(LOOCV)における最大の出現頻度を示し、さらにAUCで測定した場合に最良の性能を示した遺伝子を選択し、表11に示した遺伝子リストにまとめた。類似のアプローチを使用して、表13に示されるmiRNA、および表14に示される遺伝子の例示的な組合せを選択した。

【0318】

上記の分析に基づいて、表11、13、および14で示されるようなmRNAおよびmiRNAを、当てはまる場合、検定試料を母斑または黒色腫と正確に分類するのに有用なものとして(少なくとも2種、3種、4種、5種、6種、7種、8種、または当てはまる場合9種、10種、11種、12種、13種、14種、15種またはそれより多くの組合せで)選択した。特定の例において、このような分類器は、機械学習(例えば、ランダムフォレストまたはサポートベクターマシン)アルゴリズムを利用する。この実施例において、それぞれの発現生成物を検出するのに使用される代表的なヌクレアーゼ保護プローブも表11および13に示す。いくつかの実施例において、これらの遺伝子の発現レベルを、SDHA、RPS6KB2、RPL37A、および/またはTFRCなど(SDHAおよびRPS6KB2など)の1つまたは複数のハウスキーパーに対して正規化する。

【表11-1】

表11.母斑-黒色腫分類のための遺伝子(mRNA)

記号	代表的なNPP(5'から3')	配列番号
B4GALT1	GTCTTGGAACCTGAGCCCAGGCTGGACCTGGCA AAGGCGCTCAGTGGTAG	125
BAX	CGATGCGCTTGAGACACTCGCTCAGCTTCTTGG TGGACGCATCCTGAGGC	126
MAGEA2	CTCAGGCTCTCCACCTGGATGCTTGGCAGATCC TAGAACCACTGCATCTG	127
NR4A1	CGCCACAGCTGCCACGTGCTCCTTCAGGCAGCT GGCGATGCGGTTCTGCA	128
PDIA4	CACATCAAACCTGCTGGCCAGCACAGACGCTGA GGTTGCATCGATCTTGG	129
PRAME	GTCTGGCTGTGTCTCCCGTCAAAGGCTGCCATG AAGAGTGGCGGGAAGAG	130
RUNX1	GCAGAGTCACACACATGCAAACACGCACTCTTC GGAAGGCAGCCACTGTC	131
SOCS3	GTCTTCTCTACCAGGAGCCTGAGGTGAAAGATG TCCCGTCTCCTCCATCC	132
SAT1	ATTTCAAACATGCAACAACGCCACTGGTAATAA AGCTTTGGAATGGGTGC	133
PDLIM7	CTTCGATGTGTGTGAGGCTACCCGCATTCTCGC CATCGATGCTCAGCACC	134

10

20

30

40

## 【表 1 1 - 2】

表11.母斑-黒色腫分類のための遺伝子(mRNA)

記号	代表的なNPP(5'から3')	配列番号
BIRC5	GCACAGGCTCACAGAAGCCGAGATCCACATCA CCGCCTGGCATGCAAAGG	135
HIF1A	GGCATTCTGTGTGTAAGCATTCTCTCATTTC CTCATGGTCACATGGA	136
MET	CAAAGAAGTTGATGAACCGGTCCTTTACAGATG AAAGGACTTTGGCTCCC	137
MAGEC2	GGACTACTGGGAATGCTCTCGGTAAGATTGGT ATCACACCAGAGGGCAC	138
ERCC1	AGTGGGAAGGCTCTGTGTAGATCGGAATAAGG GCTTGGCCACTCCAGGAG	139
POLR2J3	GAGGTTGCAGTGAGCCAAGATCGCGCCAGCCTG GCGACAGAGTGAGACTC	140
LDHA	TCCATCATCTCTCCCTTCAATTTGTCTTCGATGA CATCAACAAGAGCAAG	141
PICALM	GACAGGCTGGCTGTATATTAAGGTTGGTTGCGT CATTACAGGAACACTTC	142
ZFYVE16	GAAGTTCGCTGTGAGGAAGCCAACCTCTGAAGA AACTGAGCAGTGGTTAGA	143
BEST1	GTTTCTCCAACCTGCTTGTGTTCTGCCGGAGTCAT AAAGCCTGCTTGCACC	144

10

20

## 【 0 3 1 9 】

上記で列挙した遺伝子のうちいくつかの追加の詳細は表 4 に見出すことができ、残りの遺伝子に関する類似の詳細は以下の表 1 2 に提供される。

## 【表 1 2 - 1】

表12.表4の補足

記号	名称	GenBank参照番号	
BAX	BCL2関連Xタンパク質	NM_138761 (GI: 163659848) (アルファ); NM_004324.3 (GI: 34335114) (ベータ); NM_138763 (GI: 163659849) (デルタ); NM_138764 (GI: 242117892) (シグマ); NR_027882 (GI: 242117894) (イプシロン、非コード)	10
BIRC5	バキュロウイルスのIAP反復含有タンパク質5(別名、アポトーシス阻害因子4;API4サバイビン)	NM_001168 (GI: 59859877) (v1); NM_001012270 (GI: 59859879) (v2); NM_001012271 (GI: 59859881) (v3)	
HIF1A	低酸素誘導因子1、アルファサブユニット(別名、Hif1-アルファ;PASスーパーファミリー1のメンバー;MOP1)	NM_001530 (GI: 194473733) (v1); NM_181054 (GI: 194473734) (v2); NM_001243084 (GI: 340545530) (v3)	20
MET	METがん原遺伝子(別名、肝細胞増殖因子受容体;HGFR)	NM_001127500 (GI: 188595715) (v1); NM_000245 (GI: 42741654) (v2)	
MAGEC2	黒色腫抗原、ファミリーC、2(別名、がん-精巣抗原10;CT10;HCA587;黒色腫抗原、ファミリーE、1;MAGEE1)	NM_016249 (GI:262050676)	
ERCC1	チャイニーズハムスターにおける切除修復、欠損補完、1(別名、チャイニーズハムスター卵巣細胞のUV-20によるDNA修復欠損の補完;UV20)	NM_202001 (GI: 260593723) (v1); NM_001983 (GI: 260593722) (v2); NM_001166049 (GI: 260593724) (v3)	30
POLR2J3	ヒトポリメラーゼ(RNA)II(DNA依存性)ポリペプチドJ3(別名、POLR2J2、RPB11b1、RPB11b2)	NM_001097615 (GI:332634983)	
LDHA	乳酸デヒドロゲナーゼA(別名、LDH、サブユニットM)	NM_005566 (GI: 207028465) (v1); NM_001135239 (GI: 207028493) (v2); NM_001165414 (GI: 260099722) (v3); NM_001165415 (GI: 260099724) (v4); NM_001165416 (GI: 260099726) (v5); NR_028500 (GI: 260099728) (v6, 非コード)	40

## 【表 1 2 - 2】

表12.表4の補足

記号	名称	GenBank参照番号
PICALM	ホスファチジルイノシトール結合ク ラスリン集合タンパク質(別名、ク ラスリン集合リンパ球-骨髄性白 血病遺伝子;CALM;CLTH;LAP、シ ョウジョウバエLAPのホモログ	NM_007166 (GI: 332688229) (v1); NM_001008660 (GI: 332688228) (v2); NM_001206946 (GI: 332688230) (v3); NM_001206947 (GI: 332635086) (v4)
ZFYVE16	ジンクフィンガーFYVEドメイン含有 タンパク質16(別名、エンドソーム 関連 FYVEドメインタンパク質 ;ENDOFIN;KIAA0305)	NM_014733 (GI: 157426863) (v1); NM_001105251 (GI: 157426865)
BEST1	ベストロフィン1(別名、VMD2遺伝 子、TU15B)	NM_004183 (GI: 212720874) (v1); NM_001139443 (GI: 212720888) (v2)

10

20

【表 1 3】

表13.母斑-黒色腫分類に関するmiRNA

識別子	代表的なNPP(5'から3')	配列 番号	miRBase 受託番号
hsa.miR.122	CAAACACCATTGTCACACTCCA	145	MI0000442
hsa.miR.1291	ACTGCTGGTCTTCAGTCAGGGCCA	146	MI0006353
hsa.miR.191	CAGCTGCTTTTGGGATTCCGTTG	147	MI0004941
hsa.miR.19b	TCAGTTTTGCATGGATTTGCACA	148	MI0000074
hsa.miR.200a	ACATCGTTACCAGACAGTGTTA	149	MI0000737
hsa.miR.200c	TCCATCATTACCCGGCAGTATTA	150	MI0000650
hsa.miR.203	CTAGTGGTCCTAAACATTTAC	151	MI0000283
hsa.miR.205	CAGACTCCGGTGAATGAAGGA	152	MI0000285
hsa.miR.21	TCAACATCAGTCTGATAAGCTA	153	MI0000077
hsa.miR.23b	GGTAATCCCTGGCAATGTGAT	154	MI0000439
hsa.miR.29c	TAACCGATTTCAAATGGTGCTA	155	MI0000735
hsa.miR.342.3p	ACGGGTGCGATTTCTGTGTGAGA	156	MI0000805
hsa.miR.375	TCACGCGAGCCGAACGAACAAA	157	MI0000783
hsa.miR.665	AGGGGCCTCAGCCTCCTGGT	158	MI0005563
hsa.miR.1304	CACATCTCACTGTAGCCTCAA	159	MI0006371
hsa.miR.142.5p	AGTAGTGCTTTCTACTTTATG	160	MI0000458
hsa.miR.1254	ACTGCAGGCTCCAGCTTCCAGGCT	161	MI0006388
hsa.let.7a	AACTATACAACCTACTACCTCA	162	MI0000060
hsa.miR.140.5p	CTACCATAGGGTAAAACCACTG	163	MI0000456
hsa.miR.183	AGTGAATTCTACCAGTGCCATA	164	MI0000273

10

20

30

## 【表 1 4】

表14.例示的な遺伝子の組合せ

実施形態	遺伝子の組合せ
C1	B4GALT1, BAX, MAGEA2, NR4A1, PDIA4, PRAME, RUNX1, SOCS3, SAT1, PDLIM7, BIRC5, MET, MAGEC2, POLR2J3, ZFYVE16, BEST1
C2	NR4A1, SOCS3, PRAME, POLR2J3, BEST1, RUNX1, BIRC5, MET, PDLIM7, ZFYVE16, HIF1A, PICALM
C3*	MAGEA2, PRAME, PDIA4, NR4A1, PDLIM7, B4GALT1, SAT1, RUNX1, SOCS3
C4	hsa.miRNA.342.3p, hsa.miRNA.191, hsa.miRNA.29c, hsa.miRNA.183, hsa.miRNA.182, hsa.miRNA.19b, hsa.miRNA.23b, hsa.miRNA.205, hsa.miRNA.122, hsa.miRNA.200a, hsa.miRNA.200c, hsa.miRNA.203

10

\*表4および表11のそれぞれに見出される組合せ

## 【 0 3 2 0】

まとめると、この実施例は、試料を母斑または黒色腫（例えば原発性黒色腫）として特徴付けるために例えば機械学習（例えば、ランダムフォレストまたはサポートベクターマシン）モデルで使用される場合の特定のmRNAおよびmiRNAの有用性を実証する。

20

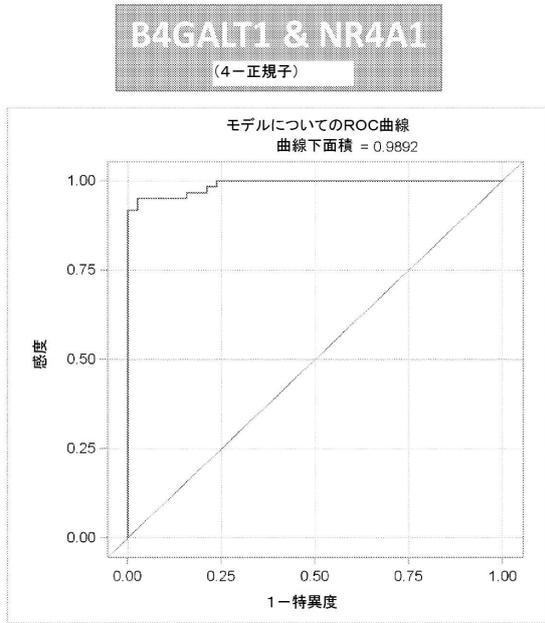
## 【 0 3 2 1】

本開示の原理が適用可能な多くの可能性のある実施形態を考慮すれば、例示された実施形態は本開示の単なる例に過ぎず、本発明の範囲を制限するものとして解釈されないことが理解されよう。そうではなく、本開示の範囲は、以下の特許請求の範囲によって定義される。それゆえに本発明者らは、本発明者らの発明はいずれもこれらの特許請求の範囲および趣旨に当てはまるものと主張する。



【 図 4 】

FIG. 4



【 図 5 】

感度 レベル	正		誤		ハートゼージ	
	対象 なし	対象 あり	対象 なし	対象 あり	正	特異度 偽陽性 偽陰性
0.640	57	36	2	4	53.9	53.4 94.7 3.4 10.0
0.660	56	36	2	5	52.9	51.8 94.7 3.4 12.2
0.680	56	36	2	5	52.9	51.8 94.7 3.4 12.2
0.700	56	37	1	5	53.9	51.8 97.4 1.8 11.9
0.720	56	37	1	5	53.9	51.8 97.4 1.8 11.9
0.740	56	37	1	5	53.9	51.8 97.4 1.8 11.9
0.760	56	37	1	5	53.9	51.8 97.4 1.8 11.9
0.780	56	37	1	5	53.9	51.8 97.4 1.8 11.9
0.800	55	37	1	6	52.9	50.2 97.4 1.8 14.0
0.820	53	37	1	8	50.9	55.0 97.4 1.9 17.8
0.840	53	36	0	8	51.9	56.9 100.0 0.0 17.4
0.860	53	38	0	8	51.9	56.9 100.0 0.0 17.4
0.880	53	38	0	8	51.9	56.9 100.0 0.0 17.4
0.900	53	38	0	8	51.9	56.9 100.0 0.0 17.4
0.920	53	38	0	8	51.9	56.9 100.0 0.0 17.4
0.940	53	38	0	9	50.9	55.2 100.0 0.0 19.1
0.960	48	38	0	13	46.9	78.7 100.0 0.0 25.5
1.000	0	38	0	61	38.4	0.0 100.0 61.6

FIG. 5

B4GALT1 & NR4A1  
(4-正規子)

- 感度への影響を比較的小さくしつつ高い特異度を得るためのモデルを使用できる

【 図 6 】

B4GALT1 & NR4A1  
(4-正規子)

- 特異度への影響を比較的小さくしつつ高い感度を得ることができる
- 80%より高い特異性と共に高い感度を維持する。試料を原発性黒色腫とコールするための可能な閾値の範囲は極めて大きい
- 一頑健なモデル

FIG. 6

感度 レベル	正		誤		ハートゼージ	
	対象 なし	対象 あり	対象 なし	対象 あり	正	特異度 偽陽性 偽陰性
0.000	61	0	38	0	51.6	100.0 0.0 38.4
0.020	61	22	15	0	63.6	100.0 57.5 20.0 0.0
0.040	61	26	12	0	67.3	100.0 68.4 16.4 0.0
0.060	60	25	9	1	69.5	95.4 76.3 13.0 3.3
0.080	60	25	9	1	69.5	95.4 76.3 13.0 3.3
0.100	60	25	9	1	69.5	95.4 76.3 13.0 3.3
0.120	60	25	9	1	69.5	95.4 76.3 13.0 3.3
0.140	59	25	9	3	67.3	95.1 76.3 13.4 5.4
0.160	59	25	9	3	67.3	95.1 76.3 13.4 5.4
0.180	59	25	9	3	67.3	95.1 76.3 13.4 5.4
0.200	59	25	9	3	67.3	95.1 76.3 13.4 5.4
0.220	59	25	9	3	67.3	95.1 76.3 13.4 5.4
0.240	59	25	9	3	67.3	95.1 76.3 13.4 5.4
0.260	59	25	9	3	67.3	95.1 76.3 13.4 5.4
0.280	59	25	9	3	67.3	95.1 76.3 13.4 5.4
0.300	59	25	9	3	67.3	95.1 76.3 13.4 5.4
0.320	59	25	9	3	67.3	95.1 76.3 13.4 5.4
0.340	59	25	9	3	67.3	95.1 76.3 13.4 5.4
0.360	59	25	9	3	67.3	95.1 76.3 13.4 5.4
0.380	59	25	9	3	67.3	95.1 76.3 13.4 5.4
0.400	59	25	9	3	67.3	95.1 76.3 13.4 5.4
0.420	59	25	9	3	67.3	95.1 76.3 13.4 5.4
0.440	59	25	9	3	67.3	95.1 76.3 13.4 5.4
0.460	59	25	9	3	67.3	95.1 76.3 13.4 5.4
0.480	59	25	9	3	67.3	95.1 76.3 13.4 5.4
0.500	59	25	9	3	67.3	95.1 76.3 13.4 5.4
0.520	59	25	9	3	67.3	95.1 76.3 13.4 5.4
0.540	59	25	9	3	67.3	95.1 76.3 13.4 5.4
0.560	59	25	9	3	67.3	95.1 76.3 13.4 5.4
0.580	59	25	9	3	67.3	95.1 76.3 13.4 5.4
0.600	59	25	9	3	67.3	95.1 76.3 13.4 5.4
0.620	59	25	9	3	67.3	95.1 76.3 13.4 5.4

FIG. 7

B4GALT1およびNR4A1(4-正規子)モデルを検証する他の方法:

- 経験的な共分散「サンディッチ」推定量を用いたモデルの検定
  - SASのproc GLIMMIXを用いて推定した
  - サンディッチ推定量は以下の通り:
    - 共分散モデルの選択には感度が低い
    - 不均一分散によるバイアスを最小化すること

ファーストのバイアス減少のペナルティ付き尤度

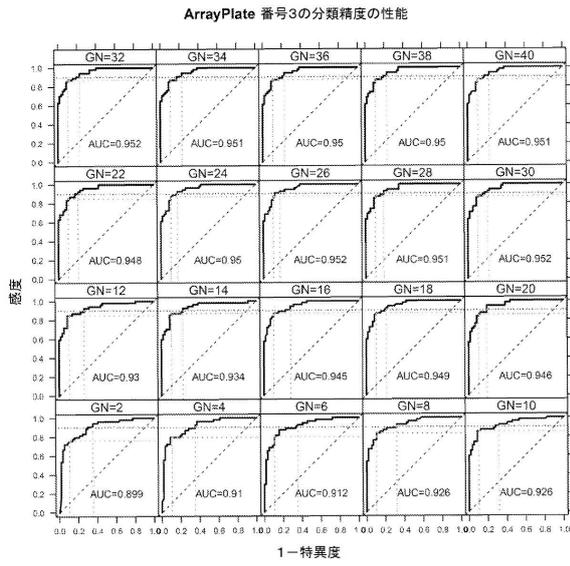
収束判定基準 (GOINV = E-8) を満たした			
モデル適合統計量			
AIC	424.780		
SC	427.395		
-2*LogL	422.760		
判定基準: 切片のみ 共変量			
AIC	424.780		
SC	427.395		
-2*LogL	422.760		
全体帰無仮説の検定: ベータ=0			
検定	カイ二乗	DF	P>カイ二乗
尤度比	98.2081	2	<.0001
スコア	73.1213	2	<.0001
ワルド	20.1588	2	<.0001

適合統計量				
-2log尤度	25.91			
AIC (小さい方がよい)	31.91			
BIC (小さい方がよい)	32.16			
CAIC (小さい方がよい)	39.69			
HQIC (小さい方がよい)	42.69			
FIQIC (小さい方がよい)	35.06			
ヒアンソカイ二乗	31.16			
ヒアンソカイ二乗 / DF	0.32			
固定効果のタイプIIIの検定				
効果	分子	分母	F値	P>F
B4GALT1	1	96	10.87	0.0014
NR4A1	1	96	14.73	0.0002



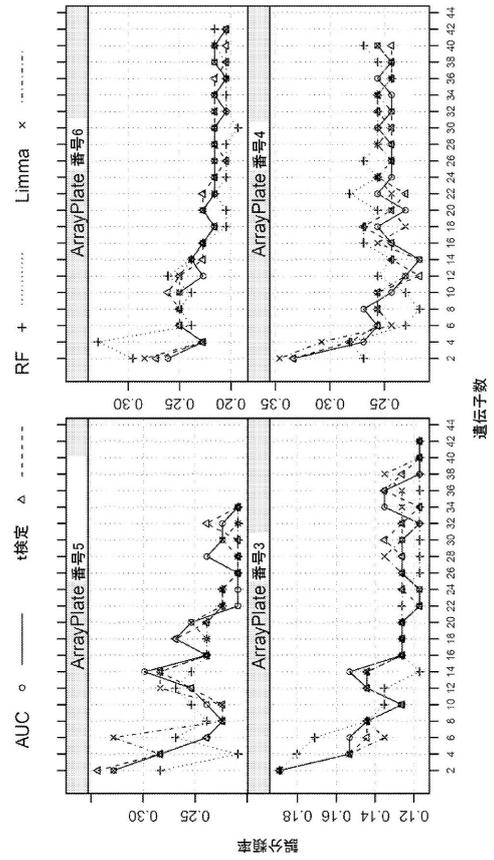
【図 1 1】

FIG. 11



【図 1 2】

FIG. 12 分類性能:CV誤差率



【配列表】

000636658000001.app

## フロントページの続き

- (74)代理人 100113413  
弁理士 森下 夏樹
- (74)代理人 100181674  
弁理士 飯田 貴敏
- (74)代理人 100181641  
弁理士 石川 大輔
- (74)代理人 230113332  
弁護士 山本 健策
- (72)発明者 ワン, ファイ  
アメリカ合衆国 アリゾナ 8 5 7 0 6 , ツーソン, イー. グローバル ループ 3 4 3 0
- (72)発明者 ロバーツ, クリストファー  
アメリカ合衆国 アリゾナ 8 5 7 0 6 , ツーソン, イー. グローバル ループ 3 4 3 0
- (72)発明者 マドゥラ, クリシュナ  
アメリカ合衆国 アリゾナ 8 5 7 0 6 , ツーソン, イー. グローバル ループ 3 4 3 0
- (72)発明者 ルー, ゼンチアン  
アメリカ合衆国 アリゾナ 8 5 7 0 6 , ツーソン, イー. グローバル ループ 3 4 3 0
- (72)発明者 ブラウン, ジャスティン  
アメリカ合衆国 アリゾナ 8 5 7 0 6 , ツーソン, イー. グローバル ループ 3 4 3 0
- (72)発明者 バシチェック, トム  
アメリカ合衆国 アリゾナ 8 5 7 0 6 , ツーソン, イー. グローバル ループ 3 4 3 0
- (72)発明者 カーンズ, ビー.ジェイ.  
アメリカ合衆国 アリゾナ 8 5 7 0 6 , ツーソン, イー. グローバル ループ 3 4 3 0
- (72)発明者 セリグマン, ブルース  
アメリカ合衆国 アリゾナ 8 5 7 0 6 , ツーソン, イー. グローバル ループ 3 4 3 0
- (72)発明者 フーン, デイブ エス.ピー.  
アメリカ合衆国 カリフォルニア 9 0 4 0 4 , サンタモニカ, サンタモニカ ブールバード  
2 2 0 0

審査官 吉田 知美

- (56)参考文献 特表2011-520451(JP,A)  
特表2008-504034(JP,A)  
特表2010-522748(JP,A)  
C. De Smet, Immunogenetics, 1994年, Vol.39, pp.121-129  
Aaron G. Smith, Pigment Cell Melanoma Res., 2011年, Vol.24, pp.551-563  
Chava Perry, Oncogene, 2002年, Vol.21, pp.8428-8441  
Gregory B Lesinski, BMC Cancer, 2010年, Vol.10, pp.1-10

## (58)調査した分野(Int.Cl., DB名)

C12Q 1/68-1/6897

CAplus/MEDLINE/WPIDS/BIOSIS(STN)