

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2014-518086

(P2014-518086A)

(43) 公表日 平成26年7月28日(2014.7.28)

(51) Int.Cl. F 1 テーマコード (参考)
C 1 2 Q 1/68 (2006.01) C 1 2 Q 1/68 A 4 B 0 6 3

審査請求 未請求 予備審査請求 未請求 (全 29 頁)

(21) 出願番号	特願2014-518508 (P2014-518508)	(71) 出願人	513255014
(86) (22) 出願日	平成23年6月29日 (2011.6.29)		バイオセラノスティクス インコーポレイテッド
(85) 翻訳文提出日	平成25年12月27日 (2013.12.27)		B I O T H E R A N O S T I C S , I N C
(86) 国際出願番号	PCT/US2011/031870		.
(87) 国際公開番号	W02013/002750		アメリカ合衆国, カリフォルニア州, サン
(87) 国際公開日	平成25年1月3日 (2013.1.3)		ディエゴ, スイート 200, タウン
		(72) 発明者	エルランダー, マーク
			アメリカ合衆国, カリフォルニア州, カールスバッド, カミノ デル プラド
			6770
			最終頁に続く

(54) 【発明の名称】 腫瘍起源の決定

(57) 【要約】

臨床設定の対象から得られたサンプル、例えばホルマリン固定/パラフィン包埋 (FFPE) サンプルにおける、54の癌種類を分類または識別するために遺伝子発現測定を使用する方法を提供する。

【選択図】 図1

【特許請求の範囲】

【請求項 1】

組織のタイプの腫瘍細胞を含有するように細胞を含む試料を分類するための転写配列を増幅する方法であって、前記方法は

ヒト被験者から得られた細胞を含むサンプル中の細胞からの 50 以上の転写された配列の生成 cDNA コピー、

増幅は、cDNA コピーを増幅された分子を生成する工程、

前記比較する式 50 のレベルまたは一つ以上の副腎皮質腫瘍から選択される 1 以上の既知の腫瘍組織、唾液腺の腫瘍、扁平上皮細胞を含む既知の腫瘍タイプの複数の同じ 50 個以上の転写された配列の発現レベルに対してより転写配列癌、神経内分泌膵臓癌、メルケル細胞癌、肺カルチノイド、原始神経外胚葉性腫瘍、性索間質腫瘍、胸腺がん/腺腫、膀胱の腺癌、および膀胱の扁平上皮癌、および

前記複数の腫瘍型又は組織の腫瘍細胞を含むかまたは含まないとして試料を分類する。

【請求項 2】

請求項 1 に記載の方法であって、前記複数の副腎皮質腫瘍、唾液腺腫瘍、扁平上皮癌、神経内分泌膵臓癌、メルケル細胞癌、肺カルチノイド、未分化神経外胚葉性腫瘍、性索間質腫瘍、胸腺癌/胸腺腫を含み、膀胱および膀胱の扁平上皮癌、腺癌、および分類は、副腎皮質腫瘍の腫瘍細胞、唾液腺、扁平上皮細胞癌、神経内分泌膵臓癌、メルケル細胞癌、肺カルチノイド、未分化神経外胚葉性腫瘍、性索間質腫瘍の腫瘍を含むかまたは含まないように、試料である前記、胸腺がん/腺腫、膀胱の腺癌、および膀胱の扁平上皮癌。

【請求項 3】

請求項 2 に記載の方法であって、前記複数の副腎皮質癌、副腎褐色細胞腫、脳の腫瘍、乳房の腺癌、子宮頸部腺癌、子宮頸部扁平上皮癌、胆管癌、子宮内膜腺癌、食道扁平上皮癌、消化管間質腫瘍、腺癌を含む胆嚢、胃食道腺癌、セミノーマ胚細胞腫瘍、非セミノーマ胚細胞腫瘍、唾液腺、扁平上皮癌、結腸直腸腺癌、小腸腺癌、明細胞腎細胞癌、嫌色素性腎細胞癌、乳頭状腎細胞癌の腫瘍、肝細胞癌、肺腺癌、肺扁平上皮癌、リンパ腫、黒色腫、髄膜腫、中皮腫、小/大細胞神経内分泌肺癌、神経内分泌膵臓癌、メルケル細胞癌、消化管カルチノイド、肺カルチノイド、明細胞腺癌、類内膜腺癌、粘液性腺癌、漿液性腺癌、膵臓腺癌、前立腺癌、悪性線維性組織球腫、原始神経外胚葉性腫瘍、平滑筋肉腫、脂肪肉腫、骨肉腫、滑膜肉腫、性索間質腫瘍、基底細胞癌、皮膚扁平上皮癌、胸腺癌/胸腺腫、濾胞/乳頭癌、髄様癌、移行上皮癌、膀胱腺癌、および膀胱の扁平上皮癌、及び、分類は、副腎皮質腫瘍の腫瘍細胞、副腎褐色細胞腫、脳、乳房の腺癌、子宮頸部腺癌、子宮頸部扁平上皮癌、胆管癌、子宮内膜腺癌、食道扁平上皮癌の腫瘍を含むか含まないような試料であり、前記消化管間質腫瘍、胆嚢、胃食道腺癌、セミノーマ胚細胞腫瘍、非セミノーマ胚細胞腫瘍、唾液腺、扁平上皮癌、結腸直腸腺癌、小腸腺癌、明細胞腎細胞癌、嫌色素性腎細胞癌の腫瘍の腺癌、乳頭状腎細胞癌、肝細胞癌、肺腺癌、肺扁平上皮癌、リンパ腫、黒色腫、髄膜腫、中皮腫、小/大細胞神経内分泌肺癌、神経内分泌膵臓癌、メルケル細胞癌、消化管カルチノイド、肺カルチノイド、明細胞腺癌、類内膜腺癌、粘液腺癌、漿液性腺癌、膵臓腺癌、前立腺癌、悪性線維性組織球腫、原始神経外胚葉性腫瘍、平滑筋肉腫、脂肪肉腫、骨肉腫、滑膜肉腫、性索間質腫瘍、基底細胞癌、皮膚扁平上皮癌、胸腺癌/胸腺腫、濾胞/乳頭癌、髄様癌、移行上皮癌、膀胱腺癌、および膀胱の扁平上皮癌。

【請求項 4】

前記 50 個以上の転写配列は、開示された 87 の遺伝子配列を含む、請求項 1、2 又は 3 のいずれかに記載の方法

【請求項 5】

全ての前記発現レベルがマイクロアレイを用いて決定される、請求項 1、2、3 又は 4 のいずれかに記載の方法。

【請求項 6】

分類が 60% 以上の精度である、請求項 1、2、3、4 又は 5 に記載の方法。

【請求項 7】

10

20

30

40

50

増幅は、転写配列のすべてまたは一部の増幅を含むか、または前記転写された配列に対応する転写および標識RNAを逆転装置、請求項1、2、3、4、5又は6に記載の方法。

【請求項8】

前記増幅が線形増幅RNAまたは定量的PCRを含む、請求項7に記載の方法。

【請求項9】

前記増幅は、転写された配列の少なくとも50ヌクレオチドの定量的PCR増幅である、請求項8に記載の方法。

【請求項10】

50～100転写配列のcDNAコピーが生成される、請求項1に記載の方法。

10

【請求項11】

前記試料は、ホルマリン固定パラフィン包埋(FFPET)サンプルである、請求項1、2、3、4、5、6、7、8、9又は10に記載の方法。

【請求項12】

前記複数において50以上の転写配列のうち大部分の発現レベルが重なる、請求項1、2、3、4、5、6、7、8、9、10又は11に記載の方法。

【請求項13】

前記複数において50以上の転写配列のうち30以上の発現レベルが重なる、請求項1、2、3、4、5、6、7、8、9、10又は11に記載の方法。

【請求項14】

前記複数において50以上の転写配列のうち35以上の発現レベルが重なる、請求項1、2、3、4、5、6、7、8、9、10又は11に記載の方法。

20

【請求項15】

前記複数において50以上の転写配列のうち40以上の発現レベルが重なる、請求項1、2、3、4、5、6、7、8、9、10又は11に記載の方法。

【請求項16】

前記複数において50以上の転写配列のうち55以上の発現レベルが重なる、請求項10に記載の方法。

【請求項17】

前記複数において60以上の転写配列のうち60以上の発現レベルが重なる、請求項1、2、3、4、5、6、7、8、9、10又は11に記載の方法。

30

【発明の詳細な説明】

【関連出願の相互参照】

【0001】

この出願は、2005年6月3日出願の米国特許出願第11/145,307号と関連する。この米国特許出願は、2004年6月4日出願の米国仮特許出願第60/577,084号、及び2006年6月2日の出願11/422,056への優先権を主張している。後者の出願は、2005年6月3日出願の米国仮特許出願60/687,174への優先権を主張している。これら四つの出願は、その全体が本書に組み込まれている。

40

【開示の分野】

【0002】

この開示は、人間の腫瘍を分類するために遺伝子発現の使用に関連がある。分類は、実行されるプロファイルまたは、50以上または任意に5以上の、パターンが表した遺伝子発現の、使用するシーケンスが複数の腫瘍種類中の表されるシーケンス。開示は、このように複数の腫瘍種類中の重なる遺伝子発現レベルまたは特定の組織から起こる腫瘍の使用を含む。開示も、提供する使用する50以上または任意に5のまたはより多くの、特定の遺伝子シーケンス、どちらが腫瘍の複数の組織源中の本、発現及びその腫瘍またはガン種

50

類。核酸発現、タンパク質発現または他の発現フォーマット中の実装されるかどうかに関係なく、サンプルが得られた対象の予想と同様にガンのより正確な識別及びこのようにその処置を許すために組織種類の、あるいは、組織起源からの腫瘍細胞を含むにつれて、遺伝子発現プロファイルはサンプルを含んでいる細胞を測定するために使用されてもよい。

【開示の概要】

【0003】

この開示は、関連がある使用する小さいまたは最小限の処置（減少する、非氷点（温度）の単にストレージのような）への何も経なくて及びサンプルを凍らせた新しいサンプルと同様にガンを分類してまたは確認する計測および/またはサンプルを含んでいる細胞中の腫瘍が固定されるホルマリンの場合臨床セッティング中の対象から行われていた、埋められる（FFPE）パラフィンがサンプルをとる遺伝子発現の。開示は、このようにサンプルを臨床FFPEサンプル上にテストを実施する病院及び他研究所に直面する現実の状況に分類する能力を提供する。サンプルは、主要な腫瘍サンプルでもよいか別の腫瘍の転移から生じた腫瘍でもよい。あるいは、もし制限されなくなれば、サンプルは、例えば、細胞学のサンプルであってもよいために、血サンプル内で細胞。分子に、開示はまた、見られてもよいガンまたは腫瘍のために起源の組織を予測することによる未知のガンまたは腫瘍の輪郭を描く。

10

【0004】

腫瘍サンプルのある場合では、腫瘍が従来病理学技術によって経られた分類を有してはならないこと、確認以外は初めに分類してもよい要求されてまたは分類された「未知の予備選挙の癌」（CUP）または「未知の起源の腫瘍」（TUM）にまたは「未知の主要な腫瘍」。確認の必要は、特に標準の技術を使用している10%の誤った分類に、5の推定を考慮して関連する。このように、ガン識別またはCIDのための、腫瘍または腫瘍サンプルの手段を提供するにつれて、複数の可能な腫瘍種類で一つであるにつれて、開示は見られてもよい。可能な腫瘍種類の範囲は、ここで開示されて及び未知のガンまたは腫瘍種類に、以前に割り当てられなかった腫瘍種類を含む。

20

【0005】

開示の第1の態様では、分類が実行されることプロファイルまたは、5以上または任意に50以上の、パターンが表した遺伝子発現の、使用するシーケンス。核酸発現、タンパク質発現または遺伝子発現の他の目印中の実装されるかどうかに関係なく、サンプルが得られた対象の予想と同様にガンのより正確な識別及びこのようにその処置を許すために組織種類の、あるいは、組織起源からの腫瘍細胞を含むにつれて、遺伝子発現プロファイルはサンプルを含んでいる細胞を測定するために使用されてもよい。

30

【0006】

表されたシーケンスの発現製品は、知られて複数またはグループ内で複数の腫瘍種類中の見つかってもよい可能な腫瘍種類ここで開示される。シーケンスの発現レベルは、このようにグループ中の複数の腫瘍種類中の起こってもよい。さらに、発現レベルの範囲は、グループ中の既知の腫瘍種類の間で重なってもよい。腫瘍種類を分類するか、確認する開示された方法論は、また、細胞（例えば腫瘍またはガン細胞）の組織源の分類または識別に適用されてもよい。

40

【0007】

分類または識別は遺伝子発現プロファイルの比較によって実行されてもよい、あるいは、50以上または任意に5以上の、パターンは複数の既知の腫瘍種類中の同じ表されたシーケンスの現れに、腫瘍サンプル中のシーケンスを表した。シーケンスのうちの少なくとも1は、既知の腫瘍種類の多数中の、既知の腫瘍種類のうちの1以上中の表される。任意に、1の既知の腫瘍種類中の少なくとも一つのシーケンスの発現レベルの範囲は、複数中の一つ以上の他の既知の腫瘍種類中の、同じシーケンスの発現レベルの範囲と重なる。場合によっては、重複は複数中の5%以上、10%以上、15%以上、20%以上、25%以上、30%以上、35%以上、40%以上、45%以上または他の既知の腫瘍種類の50%以上で起こる。

50

【 0 0 0 8 】

ある実施形態では、2以上、3以上、4以上、5以上、10以上、15以上、20以上、25以上、30以上、35以上、40以上、45以上、50以上、55以上、60以上または大多数の表されたシーケンスは、既知の腫瘍種類のうちの1以上中の表される。任意に、発現の範囲は、水平になるの各複数内で先に述べたように一つ以上の他の既知の腫瘍種類中の同じシーケンスの発現レベルのレンジ付きの1の既知の腫瘍種類重複内で一般に表されたシーケンスの。

【 0 0 0 9 】

ある実施形態では、開示は一群の54の既知の腫瘍またはガンの中で種類を分類するために使用される複数に。分類は、臨床セッティング中のかなりの精度で実行されてもよい。意味がある方法中の54の既知の腫瘍またはガン種類（それらの腫瘍種類の部分集合と同様に）の中で分類することの開示は、ヒトゲノム中の50以上の表されたシーケンスが能力がある不意の及び予想外の発見上に、部分中の基礎を形成されてある。さらに、表されたシーケンスのうちの49への5は、54の既知の腫瘍またはガンの部分集合の中で、種類を分類するために使用されてもよい。

10

【 0 0 1 0 】

それが異なる腫瘍種類で相互関係中の表される遺伝子シーケンスを確認するために監督された学習を使用するのに必要でない発見上に、開示は部分中の基礎を形成される。このように、発現がどんな50以上の表されたシーケンス（表されたシーケンスのランダムなコレクションさえ）でも、平らにする認識上に、開示は部分中の基礎を形成される必要な情報コンテンツを含む分類する、そして、分類する使用してもよい、細胞既知の組織または組織起源の複数またはグループからの腫瘍細胞であること。

20

【 0 0 1 1 】

別の態様では、開示はサンプルを含んでいる細胞を分類するために提供する。そして、5以上または任意に50以上の発現レベルを測定することによって組織種類または起源の腫瘍細胞を含むことはシーケンスを写した及び、発現を比較することはサンプルを含んでいる細胞を分類するために既知の腫瘍組織種類の複数またはグループ中のガン（あるいは、腫瘍）種類の多数の中で種類のガン（あるいは、腫瘍）細胞を含むことと同じ転写配列のそれに水平になる。54の既知のガン種類及びその部分集合の中で分類するわずか、いずれ、5以上または任意に50以上は、シーケンスを表した意味がある方法中の分類のために使用してもよい。開示は、また、表されたシーケンスが他の既知の腫瘍と比較して明らかに、あるいは、高く既知の腫瘍種類のうちの1以上と相関している（直接、あるいは、間接的に表された別のシーケンスによる相互関係を通じて）発現レベル付きのそれらである必要があるだけではないという観察上に、部分中の基礎を形成される。このように、開示、更なる実施形態中の、提供する使用する腫瘍またはガンサンプルに、比較のために既知の腫瘍種類の一つ以上で、強いまたは高い相互関係中の表されない遺伝子の発現レベルの。場合によっては、分類のために使用される遺伝子の全ては *non correlated* であってもよい、あるいは、一部の遺伝子だけは *non correlated* であってもよい。85%ある実施形態で、少なくとも90%、75%、50%または使用される発現レベルの25%は、非既知の腫瘍種類のうちの1以上と相関する。

30

40

【 0 0 1 2 】

開示は、シーケンスが複数の既知のガンまたは腫瘍種類のメンバーと彼らの発現同一水準の相互関係に基づいて、選択される必要はなかった遺伝子シーケンスの発現レベルを評価することによって実施されてもよい。このように、非限定的な例に、遺伝子シーケンスは、相互関係値に基づくガンまたは腫瘍種類またはランキングで彼らの相互関係値に基づいて、選択される必要はない。さらに、開示は必ずしも発現 *level(s)* が分類のために使用した一つ以上の他の遺伝子と、相関しているというわけではない遺伝子発現レベルの使用による実行であってもよい。そう追加の実施形態中の、分類中の機能への1の表されたシーケンスの発現レベルのための能力が冗長でないこと（無所属ある）分類のために使用される少なくとも一つの他の遺伝子発現レベルの能力。

50

【 0 0 1 3 】

開示は、臨床セッティング中のガンの原因の識別を含むがこれに限らず多様なケース中の、患者中のガンの原因を確認するために適用されてもよい。ある実施形態では、識別はガン細胞を含むと知られているサンプルを含んでいる細胞の分類によってなされる、しかし、それらの細胞の起源は知られていない。他の実施形態では、一つ以上のガン細胞を含むことはそれらのガン細胞の起源の識別によってあとに続いたにつれて、識別はサンプルを含んでいる細胞の分類によってなされる。更なる実施形態では、開示がガン及び識別の前の履歴で対象からサンプルで実施されること細胞の分類によって作るガンの前の原因からのガンであることまたは新規起源に。

【 0 0 1 4 】

追加の実施形態は複数のガンが同じ器官または組織中の本それらを含む及び、開示は、ガンが同じ起源のあるかどうかだけでなく、各ガンの原因を測定するために使用される。シーケンスが使用されることができ特定の遺伝子の発現レベルがランダムなグループの遺伝子シーケンスの発現レベルより大きい精度で、腫瘍種類の中で分類する発見上に、開示はまた、部分中の基礎を形成される。ある実施形態では、開示は提供する使用する発現レベルの開示されたセットされたかなりの精度で既知の54のガン種類の中で分類するヒトゲノム中の表されたシーケンスの。開示は、このように識別のために提供して及び表されたシーケンスに基づく遺伝子発現パターン（あるいは、プロファイルまたは「サイン」）の、使用する54のガン種類の起源を確認するために使用されてもよい情報を有する。開示も、発現の使用のために54のガン種類の部分集合の中で、分類するこれらの表されたシーケンスのレベルを提供する。さらに、開示は提供する使用する54のガン種類の部分集合の中で、分類する開示された表されたシーケンスの部分集合（例えば5以上）の発現レベルの。腫瘍種類の数に従い、80%以上から100%まで変動している精度は、成し遂げられてもよい。

【 0 0 1 5 】

開示は、異なる組織からの一組の既知の腫瘍細胞中の遺伝子シーケンスの発現レベルに基づいて及び異なる腫瘍種類である。核酸発現、タンパク質発現または他の発現フォーマット中の実装されるかどうかに関係なく、特定の既知の種類および/または特定の既知の起源または細胞種類の腫瘍を含むにつれて、これらの遺伝子発現プロファイル（異なる既知の腫瘍細胞/種類中の遺伝子シーケンスの）はサンプルを確認するために未知の腫瘍サンプル中の同じシーケンスの発現レベルと比較されてもよい。開示は、提供する例えば臨床セッティング、生き残りを含む予想と同様にガン及びこのようにその処置のより正確な識別の長所および/または、サンプルが得られた対象の、処置の後のガン再発の見込み中の。

【 0 0 1 6 】

開示はまた、発見上に部分中の基礎を形成される。そして、それは5以上または任意に50以上の使用して、表したシーケンス能力があり、ここで説明される、必然的に及び効果的に2以上の既知の腫瘍種類の中で分類することの分類の間の考慮からの一つ以上の既知の腫瘍種類を除去する。これは、分類制度の範囲の中で全ての腫瘍種類と非常に相関している発現レベルで、遺伝子を選択する必要の不足を反映する。様々に定まって、発現がどちらが分類されている腫瘍種類のグループ中の、個々の腫瘍種類または複数の種類のいずれでもと、とても相関していないかについて、平らにする複数の遺伝子で、開示は実施されてもよい。アプローチが選択に基礎をおいて及び非常に相関している遺伝子の使用する、ポジティブな相互関係に基づく腫瘍種類を「支配する」ために反対されるにつれて、たぶん他の腫瘍種類を「外へ規則」にしない他と対照的に、これはある。

【 0 0 1 7 】

腫瘍サンプルの分類一つ以上の他のガン種類を除外してここで説明される可能なガン種類のうちの1であることはもちろん下記のように信頼のレベルに基づくようにされる。信頼のレベルが低い、あるいは、信頼のレベル中の増加が好まれる所で、分類は単にサンプル中のガンのために特定の組織起源または細胞種類の高さになされることことができる。この

10

20

30

40

50

他に、及び腫瘍サンプルがすぐに分類されないところ 単一の腫瘍種類（サンプルの可能な2、3のガン種類の分類がここで一体となって説明した開示許可証）に。これは、サンプルが得られた患者に、療法の選択及び管理のためにどれを考えるべきかについて、可能な組織種類、細胞種類及び腫瘍種類の数を減少する能力のために、有利に提供する。

【0018】

開示は、悩まされた被験者の一つ以上のガンの組織源および/またはガン種類の識別する non subjective な手段を提供する。主観的な解釈が以前に組織源および/またはガン種類（予想および/またはその測定に基づくガンの治療と同様に）を測定するために使用されてもよかった所で、本開示が客観的な遺伝子発現パターンを提供して、そしてそれはガン分類のより正確な識別を提供する主観的な基準と結合して、または単独で使用してもよい。開示は特に有利に第二期であるか転移された腫瘍のサンプルに適用される、しかし、組織源および/または腫瘍種類が望ましくはあるサンプル（主要な腫瘍サンプルを含むこと）を含んでいる細胞が客観的な基準で測定したいずれでもまた、開示で使用されてもよい。もちろん、クラスの最後の測定は、客観的な及び非具象的な（あるいは、主観的な/部分的に主観的な）基準の組合せに基づくようにされてもよい。

10

【0019】

開示は、患者の臨床であるか医学心配の一部として、その使用を含む。このように遺伝子の発現プロファイルを使用することに加えて、組織源を測定するガンおよび/またはガンの腫瘍種類で苦しめられる対象から、サンプルを含んでいる細胞を分析するためにここで説明されるにつれて、プロファイルはまた、対象中のガンの予想を測定するために方法の一部として使用されてもよい。腫瘍/ガンおよび/または予想の分類は、選択してまたは測定してまたは言われて治療的な処置を変更するために使用されてもよい対象。このように、開示の分類方法は病気の治療に向けられてもよい。そして、それは全体中の、あるいは、分類に基づく部分中の診断される。診断、適切な反腫瘍エージェントまたは療法の管理または差し控えを与えるまたは、反腫瘍エージェントまたは療法の交替は、ガンを治療するために使用されてもよい。

20

【0020】

開示の一つ以上の実施形態の詳細は、描くこと及び説明を下で伴うこと中の述べられる。他の特徴及び開示の長所は、描いている及び詳細な説明から及び特許請求の範囲から明らかである。

30

【図面の簡単な説明】

【0021】

【図1】54の開示された腫瘍種類のうちの39中の、開示の第1の転写配列の発現レベルの範囲を示す。

【図2】54の開示された腫瘍種類のうちの39中の、開示の第二の転写配列の発現レベルの範囲を示す。

【定義】

【0022】

本明細書で用いられる「遺伝子」は、別個の生成物をコードするポリヌクレオチドであるかどうか、RNAまたはタンパク質性の性質で。これは、複数のポリヌクレオチドが別個の産物をコードすることが可能であり得ることが理解される。用語は、対立遺伝子および多型、同じ製品をコードする遺伝子の、または染色体上の位置と、通常の有糸分裂の際に再結合する能力に基づいて、その（増加、減少、または機能の調節を含む）を機能的に関連するアナログが含まれている。

40

【0023】

「シーケンス」または「遺伝子シーケンス」ここで使用される、組み立てられる核酸分子またはポリヌクレオチドは、ヌクレオチドベースの離散的な順序である。用語は、離散的な製品（すなわち「コーディング領域」）をコード化するベースの配列を含むどうかRNAまたは性質中の proteinaceous。複数のポリヌクレオチドが離散的な製品をコード化することができてよいことは、理解される。ヒト遺伝子シーケンスの対立遺

50

伝子及び多形性が存在してもよくて及び遺伝子シーケンスまたは対立遺伝子の発現 *level(s)* を確認する開示またはその多形性の実行中の使用されてもよいことは、また、理解される。対立遺伝子または多形性の識別は、染色体の位置への部分及び有糸分裂の間、再び結合する能力中の場合による。

【0024】

「シーケンスを表した」、写されるシーケンスは、細胞の範囲内で細胞プロセスのそばにある。表されたシーケンスを検出するために、他の表されたシーケンスと比較してユニークであるシーケンスの領域は、使用されてもよい。表されたシーケンスはポリペプチド製品をコード化してもよい、あるいは、少しの製品もコード化すると知られている。それで、表されたシーケンスは、開いた読んでいるフレームまたは開いた読んでいるフレーム以外を含んでもよい。非限定的な例はおよそ8以上、およそ10以上、およそ12以上、およそ14以上、およそ16以上、およそ18以上、およそ20以上、およそ22以上、およそ24以上、およそ26以上、およそ28以上の領域を含むまたは、表されたシーケンスの範囲内のおよそ30以上の隣接するヌクレオチドは使用されてもよい。「利用できる」期間、中で使用されるにつれて、前の文は定まった数値から1の増加または減少まで参照する。表されたシーケンスの物理的な形態は、RNA分子または対応するcDNA分子であってもよい。

10

【0025】

用語は、「関連させる」または、例えば、「相互関係」またはその等価物は制限されなくて、一つ以上の遺伝子の発現の間の関連及び別のイベントを参照するために（生理的発現型または特性）例えば腫瘍種類。

20

【0026】

「ポリヌクレオチド」は、どんな長さでも、どちらのリボヌクレオチドでもまたはデオキシリボヌクレオチドのヌクレオチドの重合形態である。この用語は、分子の初源的構造体だけを参照する。このように、この用語は、*double* 及び単一の要素をもつDNA及びRNAを含む。それも、ポリヌクレオチドの変更されていない形態と同様に芸術、メチル化、「キャップ」、アナログによる自然に起こっているヌクレオチドの一つ以上の置換及び *internucleotide* 修正（例えば充電してない結合（例えば *phosphorothioates*、*phosphorodithioates*、その他））中の知られているラベルを含む修正の既知の種類を含む。

30

【0027】

用語は、「拡大する」拡大製品を生成することはDNAまたはRNAポリメラーゼで酵素で作られることができることを意味する広義中の使用する。ここで使用されるにつれて、「拡大」は一般に要求されたシーケンス（サンプルの特にそれら）の複数のコピーを生成するプロセスを参照する。「複数のコピー」は、少なくとも2枚の複写を意味する。「コピー」が、必ずしもテンプレートシーケンスにとって完全なシーケンス相補性またはアイデンティティを意味するというわけではない。mRNAを拡大する方法は、技術中の一般に知られているか及び逆のコピーPCR（RT-PCR）及び量的PCR（あるいは、Q-PCR）を含むか、本当にPCRの時間を計る。あるいは、RNAは直接ラベルをつけられてもよい技術中の知られている方法による対応するcDNAに。

40

【0028】

「一致する」ことによって、核酸分子が別の核酸分子でシーケンスアイデンティティの相当な量を共有することを意味される。相当な少なくとも95%量手段、通常少なくとも98%及びより通常少なくとも99%及び、シーケンスアイデンティティは、*Altschul* ほか（1990）中の説明されるにつれて、BLASTアルゴリズムを使用して決定される *J. Mol. Biol.* 215:403-410（発表されたデフォルトのセッティング（すなわちパラメータ $w=4$ 、 $t=17$ ）を使用すること）。

【0029】

「マイクロアレイ」は、領域（各定義済みの領域を有すること）が制限されなくて、例えば強烈な支持の表面上に形をなした独立した装置の線形のまたは二次元のまたは三つ次

50

元の（及び固体の位相）列であるために（ガラス）プラスチックまたは合成の膜。マイクロアレイ上に離散的な領域の密度は、少なくとも単一の強烈な位相支持の表面上に検出される固定されたポリヌクレオチドの総数によって、まわりに決定される $50 / \text{cm}^2$ 少なくともまわりに $100 / \text{cm}^2$ 少なくともまわりにまたは、 $1,000 / \text{cm}^2$ またはそれ以上についてまで、 $500 / \text{cm}^2$ 。配列は、全体でおよそ 500 未満、およそ 1000 、およそ 1500 、およそ 2000 、およそ 2500 またはおよそ 3000 の固定されたポリヌクレオチドを含んでもよい。ここで使用されるにつれて、DNA マイクロアレイはオリゴヌクレオチドの配列である、あるいは、チップまたは他の表面上に置かれる調査が雑種を産んだものであるポリヌクレオチドはサンプルから拡大するか、ポリヌクレオチドのクローンをつくった。各特定のグループの配列内で調査の位置が知られているので、ポリヌクレオチドがそうすることができるサンプルのアイデンティティが彼らがマイクロアレイ中の特定の位置に縛ることに基づいて決定されて。選択肢のためにマイクロアレイの使用する、どんなサイズでもの配列は開示の実行中の使用されてもよい。そして、単一の遺伝子シーケンスの現れを検出するために固体の位相中の二次元のまたは三つ次元の配置の一つ以上の位置の配置を含む。ある実施形態では、本開示用としてのマイクロアレイは、写真平板の技術によって準備されてもよい（表面上に核酸調査の統合のようなから 3' 終端）または固体の表面上に宣誓証言が続く核の統合によって。

10

【0030】

開示が遺伝子発現の識別に頼る所で、特定の遺伝子シーケンスに特有であるポリヌクレオチドへのサンプル細胞の、開示のある実施形態は mRNA の交雑またはその拡大されるかクローンをつくられたバージョンによって発現を決定する。この種類のポリヌクレオチドは少なくともおよそ 16 、少なくともおよそ 18 、少なくともおよそ 20 、少なくともおよそ 22 、少なくともおよそ 24 、少なくともおよそ 26 、少なくともおよそ 28 、少なくともおよそ 30 を含むまたは、遺伝子シーケンスの全くおよそ 32 の連続的 base pairs で、それは他の遺伝子シーケンス中に見つからない。「利用できる」期間、中で使用されるにつれて、前の文は定まった数値から 1 の増加または減少まで参照する。他の実施形態は、ポリヌクレオチドである少なくともまたは、少なくともまたは 100 について少なくともまわりに、 50 についてのまたは、少なくともまたは、少なくともまたは、少なくともまたは、少なくともまたは、少なくともまたは、少なくともまたは他の遺伝子シーケンス中に見つからないシーケンスの 500 の連続的ベースについて、 450 について、 400 について、 350 について、 300 について、 250 について、 200 について、 150 。「利用できる」期間、先立つこと中の使用されるにつれて、文は定まった数値から 10% の増加または減少まで参照する。より長いポリヌクレオチドは、もちろんサンプルの核酸に交雑を変化させないマイナーな不適当な組合せ（例えば突然変異の存在を介して）を含んでもよい。遺伝子のシーケンスまたはそのユニークな部分に雑種を産むことができるあるポリヌクレオチドが深く探るにつれて、そのようなポリヌクレオチドはまた、参照されてもよい。そして、ここで説明される。そのようなポリヌクレオチドは、彼らの検出中の援助するためにラベルをつけられてもよい。シーケンスは、遺伝子（そのような mRNA への対応する cDNA）によってコード化される mRNA のそれらであってもよくておよび / またはそのようなシーケンスのバージョンを拡大した。開示のある実施形態で、ポリヌクレオチド調査は、配列（他の固い支持装置）上に、あるいは、調査を局所化する個々の点中の固定される。

20

30

40

【0031】

開示の他の実施形態で、遺伝子シーケンスの全部または一部が拡大されてもよい及びポリメラーゼ連鎖反応（PCR）のような方法及び、例えば、そのバリエーションによって検出されてもよいが、制限されてはならないことのために（量的 PCR（Q-PCR））コピー PCR（RT-PCR）及びリアルタイム PCR を翻す（を含む mRNA の最初の量を計測する手段が、サンプル中の各シーケンスのためにコピーする）（任意にリアルタイム RT-PCR またはリアルタイム Q-PCR）。そのような方法は遺伝子シーケンスの部分に相補的である 1、2 の下塗りを利用する、そこで、下塗りには主要な核酸合成に使用

50

される。

新しく合成された核酸は、任意にラベルをつけられて及び開示のポリヌクレオチドに交雑によってか直接検出されてもよい。彼らの交雑を可能にする状況の下で、新しく合成された核酸は、開示のポリヌクレオチド（シーケンスを含むこと）で連絡されてもよい。表された核酸の発現を検出する追加の方法は、液体位相交雑を含むRNAse保護分析評価及び細胞の原位置ハイブリッド形成を含む。

【0032】

あるいは、及び開示の更なる実施形態中の、遺伝子発現は、重要な細胞サンプル中の表されたタンパク質の解析によって決定されてもよいによって個々の遺伝子製品（タンパク質）または、言われた細胞サンプル中のまたは対象の身体の流体中の、そのタンパク質分解断片の一つ以上のエピトープのために特定の一個以上の抗体の使用する。細胞サンプルは、上皮細胞が対象の血から豊かにした乳ガンのうちの1であってもよい例えばによって蛍光起動する細胞ソート（FACS）が続く細胞表面目印に対するラベルをつけられた抗体の使用する。そのような抗体は、遺伝子製品に縛った後に彼らの検出を許すためにラベルをつけられてもよい。開示の実行のために、適当な検出方法論は含む、しかし、制限する（サンプルを含んでいる細胞または組織の免疫組織化学）、酵素は組織または血サンプルを含んでいる細胞、質量分光及びimmuno PCRの抗体サンドイッチ分析評価を含んでいる免疫吸着剤分析評価（ELISAs）を結んだ。

10

【0033】

用語は「ラベル」または「ラベルをつけた」は、ラベルをつけられた分子の存在を示す見つけられる信号を生成することができる構成を指す。適当なラベルは、放射性同位元素、ヌクレオチド発色団、酵素、サブストレート、蛍光性の分子、chemiluminescentな半分、磁気小片、bioluminescentな半分、などを含む。このため、ラベルはそばに見つけられるどんな構成でもある分光器の（光化学物質）生化学、免疫化学的、電気的、光学のまたは化学的手段。

20

【0034】

用語「支持」は、ビーズ、小片、計量棒、繊維、フィルタ、膜及びシランまたはケイ酸塩支持（例えばガラスのスライド）のような従来支持を参照する。

「発現」及び「遺伝子発現」は、核酸材のコピーおよび/または翻訳を含む。表されたシーケンスの発現レベルは参照によって任意に標準化されてもよい、あるいは、一つの制御の発現レベルへの比較は遺伝子を表した。これらの「正常化遺伝子」は、複数の全てのメンバーまたは既知の腫瘍種類のグループ中の比較的一定である発現レベルを有する。

30

【0035】

ここで使用されるにつれて、「含む」こと及びその親族が中で使用される期間、彼らの含んでいる意味；それがあって、用語に等価物「含む」及びその対応する親族。

【0036】

イベントが起こることを「可能にする」状況または、交雑のような、起こるためにイベントに「ふさわしい」状況は拡張、などを立ち往生させる、あるいは、「適当な」状況はそのようなイベントが起こるのを防がない状況である。このように、これらの状況許可証は、高まる、容易にしておよび/または貢献するイベント。たとえば、そのような状況（技術中の知られていて及びここで説明される）は、ヌクレオチド配列、温度及びバッファ状況の性質に依存する。これらの状況も、交雑、裂け目、岸拡張またはコピーのような、どんなイベントが要求されるかについて次第である。

40

【0037】

ここで使用されるにつれて、シーケンス「突然変異」は参照シーケンスに比較中の関心を遺伝子のシーケンス内でシーケンス変更がここで開示したいずれにでも参照する。代用、削除または挿入のようなメカニズムのために、シーケンス突然変異は、シーケンス中の複数のヌクレオチドの単一のヌクレオチド変化または変更を含む。ここで使用されるにつれて、単一のヌクレオチド多形性（SNP）はまた、シーケンス突然変異である。本開示

50

が遺伝子発現の相対的なレベルに基づくので、ここで開示されるにつれて、遺伝子の領域を non coding して中で突然変異はまた、開示の実行中の分析されてもよい。

【0038】

「検出」または「検出する」は検出するどんな手段でも含む。そして、遺伝子発現及び変化のレベルの直接の及び間接的な測定をその中で含む。

【0039】

さもなければ定義されない限り、一般に、この開示がふさわしい当業者に理解されるにつれて、ここで使用される全ての技術及び科学的な用語は同じ意味を有する。

【実施形態の詳細な説明】

【0040】

この開示は、従来の病理学技術で可能であるより、客観的な方法中の、ガンおよび/または腫瘍を分類するために遺伝子発現情報の使用のための方法を提供する。開示は、複数の腫瘍種類（例えば米国特許公報 US 2006/0094035 及び US 2007/0020655 の下で及び中の説明される腫瘍種類）の、一体となって腫瘍サンプルを分類するために使用される遺伝子シーケンスの数をランダムに減少する結果上に、部分中の基礎を形成される。合計は 16,948 の遺伝子の数える。そして、ダウンを濾過されたより大きいここで説明されるにつれて、使用されるサンプル内で表示または不変の信号がクロス確認及び予言精度のために低く使用された遺伝子の除去に基づいてセットする。

【0041】

このように第 1 の態様中の、開示がサンプルを含んでいる細胞を分類するガンまたは腫瘍細胞を含む方法を提供することの（またはから）一種の組織（あるいは、組織起源であるにつれて）。方法は 5 以上または任意に 50 以上の発現レベルを決定するか、計測することを含む。そして、発現に比較中のサンプルの細胞中の言われたシーケンスの発現レベルに基づく複数の腫瘍種類から一種の組織の腫瘍細胞を含むことは既知の腫瘍中の水平になるにつれて、細胞中の細胞からの転写配列が対象から得られるサンプルを含んで及びサンプルを分類する。ここで使用されるにつれて、「複数」は 2 以上の状態を参照する。

【0042】

分類する既知の腫瘍サンプル内で彼らの発現レベルへのサンプルおよび/または既知の non tumor サンプルの細胞中の分析された転写配列の発現レベルの比較に基づく。この他に、分類する同じサンプル内で参照シーケンスの現れへの分析された転写配列の発現レベルの比較に基づいてまたは基づく、既知の腫瘍サンプルおよび/または既知の non tumor サンプル内で同じ比較。非限定的な例（シーケンスがレベルが検出してまたは対象からサンプルを含んでいる細胞中の、決定した発現が比較されるデータベースを提供するために既知の腫瘍サンプルの 1 組決定されてもよい遺伝子の発現レベル）。サンプル内でも遺伝子 sequence(s) の発現 level(s) は、同じサンプルまたは対象から望ましくは通常のまたは non cancerous な細胞中の言われた sequence(s) の発現 level(s) と比較されてもよい。Q-PCR またはリアルタイム Q-PCR を利用している開示の実施形態の下で及び中の説明されるにつれて、発現レベルは同じサンプル中の参照遺伝子の発現レベルと比較されてもよい、あるいは、発現レベルの比率は使用されてもよい。

【0043】

あるいは、種々の基準に基づく選択によって、使用する表されたシーケンスの選択は、ランダムでもよい。1 の非限定的な例に、遺伝子は配列する集まっている技術を含む管理されない学習に基づいて選択してもよい。非限定的な別の例に、選択は減少することになっていてもよいまたは、彼らの能力に関する移転冗長は、分類する腫瘍種類。例えば、遺伝子シーケンスは、彼らの発現間の相互関係の不足及び分類するために使用される一つ以上の他の遺伝子シーケンスの現れに基づいて選択される。これは、相関係数の相互関係行列を生成するデータセット中の各他の遺伝子の発現レベル付きのサンプルの多数にまたがって、相互関係のために発現データセット中の各遺伝子シーケンスの発現レベルを評価することによって達成される。遺伝子シーケンスの各ペアの、あるいは、間接的に遺伝子シ

10

20

30

40

50

ーケンスの各ペアの発現値の直接の比較のない発現の間で、これらの相互関係測定は、直接実行されてもよい。

【 0 0 4 4 】

様々な相互関係方法論は、データセットの範囲内で個々の遺伝子シーケンスの発現データの相互関係中の使用されてもよい。非限定的な例は、相互の情報及び非線形のアプローチに基づく方法論と同様にパラメータの及び非母数方法を含む。パラメータのアプローチの非限定的な例は、ピアソン相互関係を含む（または、ピアソン r は、また、参照したために線形または $product\ moment$ 相互関係）及びコサイン相互関係。非母数方法の非限定的な例は、Spearman の R （あるいは、ランク-順序）相互関係、ケンドールの Tau 相互関係及び $Gamma$ 統計を含む。各相互関係方法論は、データセット中の個々の遺伝子シーケンスの現れ間の相互関係のレベルを決定するために使用されることができ、他の全てのシーケンスによる全てのシーケンスの相互関係は、最もすぐに考えられる行列に。ピアソンの相互関係を使用する非限定的な例に、方法中の相関係数 r は、使用される相互関係のレベルのインジケータに。他の相互関係方法が使用されるとき、 0.5 で、あるいは、についてあることに 0.25 で、あるいは、についてある r と一致している相互関係の等価レベルの認識と一緒に、 r に類似している相関係数は使用されてもよい。

10

【 0 0 4 5 】

種々の数に相関している遺伝子シーケンスの数を減少するのを要求されるにつれて、相関係数は選択されてもよい。 r を使用している開示のある実施形態で、選択された係数値は、およそ 0.25 でもよいか、 0.3 または 0.4 またはそれ以上（およそ 0.45 またはそれ以上）についての、あるいは、 0.5 またはそれ以上についてのそれ以上（およそ 0.35 またはそれ以上）について、より高くてもよい。データセット中の遺伝子シーケンスの間の発現があるところがその値またはそれ以上で相関させた、彼らがおそらく部分集合中の含まれない共同作用の値手段の選択開示。このようにある実施形態では、方法は、腫瘍種類データセット中の別の遺伝子シーケンスによる要求された相関係数より上に、相互関係中の表される 1 以上の（分類のために使用していない）遺伝子シーケンスを除外するか、削除することを含む。他のどの遺伝子シーケンスもと相関していない遺伝子シーケンスの状況が彼らが必要しも移動されるというわけではなくて、どのケースであるかについてあることがありえることは、指摘されるしかし分類中の使用する。

20

30

【 0 0 4 6 】

このように、発現は遺伝子シーケンスの水平になるどこでおよそ 10% 、およそ 20% 以上、およそ 30% 以上、およそ 40% 以上、およそ 50% 以上、およそ 60% 以上、およそ 70% 以上より、レベルのおよそ 80% またはおよそ 90% 以上がシーケンスが使用した遺伝子の別の一つのそれと、相関していないより、より、開示の実行中の使用されてもよくしなさい。発現レベル間の相互関係は、およそ 0.9 、およそ 0.8 、およそ 0.7 、およそ 0.6 、およそ 0.5 、およそ 0.4 、およそ 0.3 またはおよそ 0.2 の下で値に基づいてもよい。部分集合中の遺伝子シーケンスの現れが部分集合から締め出される遺伝子シーケンスの現れと相関しているので、ある遺伝子シーケンスの発現レベルの除外で、クラスの間で分類する能力は本。それで、除外された遺伝子シーケンスの現れに基づく情報が部分集合中の保持されるシーケンスによってさらに表されるので、情報は消失しなかった。したがって、部分集合の遺伝子シーケンスの現れは、プロパティに関連する情報コンテンツおよび/または電池の特性（あるいは、発現型）を有する。これは、出願及び分類されることができる本来の遺伝子発現データセットの一部として、含まれない追加の腫瘍種類クラスの分類への関連を有する開示の部分集合の使用する追加のそれら中の表される部分集合及びシーケンス内でシーケンスの現れ間の情報の冗長が属してに基づいて。このように、既知のクラスの多数を越えた腫瘍種類であることは本来の遺伝子発現データセットを生成したものであるにつれて、開示は細胞を分類するために使用されてもよい。

40

【 0 0 4 7 】

50

特定の腫瘍種類に発現の相互関係を減少することに基づく遺伝子シーケンスの選択は、また、使用されてもよい。これも本開示の発見を反映する。そして、とても一つ以上の腫瘍種類と最も相関した発現レベルが異なる腫瘍種類の内の一つ必ずしも分類中の最も大きい値であるというわけではなかったという観察に基づく。これは分類のためにランダムに選択された遺伝子シーケンスを使用する能力によって、両方とも反映される使用する特定のシーケンスここで説明されるにつれて、の、一つ以上の腫瘍種類で最上位の相互関係で表されない。このように、開示は最もかなりのP値に基づく遺伝子シーケンスまたは遺伝子発現及び一つ以上の腫瘍種類の相互関係に基づくランキングの選択なしで実施されてもよい。このように、開示は第一流のベースの方法論（例えばKruskal-ウォリスH-テスト）を使わずに実施されてもよい。

10

【0048】

開示の実行中の使用される遺伝子シーケンスは特定の既知の腫瘍種類（例えばエストロゲンレセプターの現れ）で、相互関係中の表されるのを見られたそれらを含んでもよい。そして、それはある胸及び卵巣のガンで相互関係中の表されるのを見られた。開示のしかしある実施形態で開示が実施されること発現レベルが既知の腫瘍種類のうちの2以上中の重なる複数の遺伝子シーケンスの発現レベルの使用する。場合によっては、転写配列のうちの1以上は、既知の腫瘍種類、既知の腫瘍種類のうちの少なくとも50、既知の腫瘍種類のうちの少なくとも45、既知の腫瘍種類のうちの少なくとも40、既知の腫瘍種類のうちの少なくとも35、既知の腫瘍種類のうちの少なくとも30、既知の腫瘍種類のうちの少なくとも25、既知の腫瘍種類のうちの少なくとも20、既知の腫瘍種類のうちの少なくとも10または開示されたグループの既知の腫瘍種類のうちの少なくとも5の全て中の重なる発現レベルを有する。

20

【0049】

2以上、3以上、4以上、5以上、6以上、7以上、8以上、9以上、10以上、11以上、12以上、13以上、14以上、15以上、16以上、17以上、18以上、19以上、20以上、22以上、24以上、26以上、28以上は、開示された方法の実行、2中の重なり合うレベルの範囲で表される転写配列の数または開示されたグループのより既知の腫瘍種類中の使用されてもよい、30以上、32以上、34以上、36以上、38以上、40以上、42以上、44以上、46以上、48以上、50以上、52以上、54以上、56以上、58以上、60以上、62以上、64以上、66以上、68以上、70以上、72以上、74以上、76以上、78以上、80以上、92以上、94以上、96以上、98以上、100以上、105以上、110以上、120以上、130以上、140以上または150以上。もちろん、上記の値は、2の間で重ならない発現レベル付きの一つ以上の転写配列で使用されてもよいが、より多くであるあってもよい、一群の既知の腫瘍種類のメンバー。複数中の既知の腫瘍種類の数に基づいて、熟練した人は、サンプルの現れを重ねることによる多くの腫瘍種類の適切な組合せが発現を重ねて、シーケンス及び多くの転写配列を写したいずれでも2以上の腫瘍種類中の平らにする開示を実施してもよい。

30

【0050】

開示が人間の対象に関して主に説明される間、他の対象からのサンプルはまた、使用されてもよい。必要である全ては、発現が未知数またはテストサンプル中の水平になるようなものはそうしてもよい複数の既知の腫瘍サンプル中の、遺伝子シーケンスの発現レベルを評価する能力である比較する。このように、開示は複数の表されたシーケンス及び複数の既知の腫瘍サンプルが利用できるどんな有機体からでも、サンプルまで印加されてもよい。1の非限定的な例はマウスサンプルへの開示の出願である。そして、表されたネズミシーケンスの検出及び既知のマウス腫瘍サンプルまたは能力の入手可能性が既知のサンプルを得るのを許すためにマウスゲノムの入手可能性に基づく。このように、開示は哺乳類、霊長類及び非限定的に、臨床テスト（例えばネズミ、マウス、ウサギ、犬、猫及びチンパンジー）中の使用される動物のそれらを含む他のサンプルの用途に、考えられる例。

40

【0051】

50

開示がサンプルを含んでいる細胞を使ってすぐに実施される間、遺伝子発現のために分析されてもよいサンプルを含んでいる核酸が平らにするいずれでも開示の実行中の使用されてもよい。開示を制限することなく、開示のサンプルは、疑われるか、腫瘍細胞を含むと知られているものであってもよい。あるいは、開示のサンプルは「腫瘍サンプル」または「サンプルを含んでいる腫瘍」であってもよい、あるいは、組織の「サンプルを含んでいる腫瘍細胞」または個人から分離される流体は、あるいは、危険な状態に、発達することでガンに悩むことで疑いをかけた。例えば、開示用としてのサンプルの非限定的な例は、制限されなくて、臨床サンプルを含むために（固定サンプル）新しいサンプルまたは凍ったサンプル。サンプルは気音、細胞学のサンプル（血または、腹水または肋膜の空洞から流体を含んで、他の身体の流体を含む）または組織標本であってもよい。そして、適切な細胞または核酸が遺伝子発現レベルの測定に利用できる限り、それは見本中の細胞の元の位置の関係に関して少なくともある情報を含む。開示は、極寒の組織地域で得られる結果が固定組織または細胞サンプルで、正当に状況に印加されることができ発見上に部分中の基礎を形成されて及び新しいサンプルに延長される。

10

20

30

40

50

【0052】

固定サンプルの非限定的な例は、ブーダンのもの、glutaldehyde、アセトン、アルコールまたは、免疫組織化学（IHC）のために細胞または組織サンプルを固定するために使用されるそれらのような、他のどの色留め剤でも、ホルマリンまたはホルムアルデヒド（FFPEサンプルを含むこと）を据えられるそれらを含む。他の例は、急な細胞が核酸及びタンパク質を関連付けた色留め剤を含む。凍った組織標本（例えばその凍った状態を維持する必要）を取り扱うこと中の可能な複雑化を与えられて、開示は血または他の身体の流体からの細胞または組織を含むnon frozenサンプル（例えば固定サンプル、新しいサンプル）で実施されてもよくて及び最小限にサンプルを扱った。開示のある出願では、サンプルは、例えば、制限されなくて、標準の病理学技術を使用することを分類されなかった、免疫組織化学ベースの分析評価。

【0053】

開示のある実施形態で、54に続くことから選択される種類の腫瘍細胞及びその部分集合を含むにつれて、サンプルは分類される：adrenal corticalな腫瘍、副腎のクロム親和性細胞腫、脳の腫瘍、胸の腺癌、子宮頸部の腺癌、子宮頸部の扁平上皮細胞癌、cholangiocarcinoma、子宮内膜腺癌、食道の扁平上皮細胞癌、胃腸内間質腫瘍、胆嚢の腺癌、胃食道の腺癌、seminomatousな性細胞腫瘍、nonseminomatousな性細胞腫瘍、唾液の腺、扁平上皮細胞癌、結腸直腸腺癌、小腸腺癌の腫瘍（はっきりした細胞腎臓細胞癌、非染色性細胞腎臓細胞癌、乳頭腎臓細胞癌、肝細胞性癌、肺腺癌、肺扁平上皮細胞癌、リンパ腫、黒色腫、髄膜腫、中皮腫、小さい/大きい細胞神経内分泌系肺ガン、神経内分泌系すい臓ガン、merkel細胞癌、胃腸内カルチノイド、肺カルチノイド、はっきりした細胞adenocarcinoma、endometrioid腺癌、mucinousな腺癌、漿液の腺癌、膵臓の腺癌、前立腺）、腺癌、悪性の繊維のhistiocytoma、最初のneuroectodermalな腫瘍、leiomyosarcoma、liposarcoma、骨肉腫、滑液肉腫、性的コード間質腫瘍、基底細胞癌、皮膚扁平上皮細胞癌、thymicな癌/thymoma、follicular/papillary癌、骨髄の癌、移り変わる細胞癌、膀胱の膀胱及び扁平上皮細胞癌の腺癌。

【0054】

これらの54の腫瘍種類は、以下の組織種類と一致する：副腎の組織（adrenal corticalな腫瘍及び副腎のクロム親和性細胞腫）、脳組織（脳の腫瘍）、胸組織（胸の腺癌）、子宮頸部の組織（子宮頸部の腺癌及び子宮頸部の扁平上皮細胞癌）、胆管組織（cholangiocarcinoma）、子宮内膜組織（子宮内膜腺癌）、食道の組織（食道の扁平上皮細胞癌）、胃腸内組織（胃腸内間質腫瘍またはGIST）、ずうずうしさ膀胱組織（ずうずうしさ膀胱の腺癌）、胃食道の組織（胃食道の腺癌）、性細胞組織（seminomatousな性細胞腫瘍及びnonseminomatousな

性細胞腫瘍)、頭及び首組織(唾液の腺及び扁平上皮細胞癌の腫瘍)、腸の組織(結腸直腸腺癌及び小腸腺癌)、腎臓組織(はっきりした細胞腎臓細胞癌、非染色性細胞腎臓細胞癌及び乳頭腎臓細胞癌)、肝臓組織(肝細胞性癌)、肺組織(肺腺癌及び肺扁平上皮細胞癌)、癌)(リンパ球(リンパ腫))メラニン形成細胞(黒色腫)、髄膜の組織(髄膜腫)、中皮(中皮腫)の組織、神経内分泌系組織(小さい/大きい細胞神経内分泌系肺ガン、神経内分泌系すい臓ガン、merkel細胞癌、胃腸内カルチノイド及び肺カルチノイド)、卵巣組織(はっきりした細胞adenocarcinoma、endometrioid腺癌、mucinousな腺癌及び漿液の腺癌)、すい臓(膵臓の腺癌)の組織、前立腺組織(前立腺腺癌)、肉腫の組織(悪性の繊維のhistiocytomaまたはleiomyosarcomaな、liposarcomaなMFH(最初のneuroectodermalな腫瘍またはPNET)(骨肉腫)及び滑液を分泌する肉腫)(性的コード間質組織(性的コード間質腫瘍))皮膚組織(基底細胞癌及び皮膚扁平上皮細胞癌)、胸腺組織(thymicな癌/thymoma)、甲状腺の組織(follicular/papillary癌及び骨髄の癌)及び膀胱組織(移り変わる細胞癌(膀胱の腺癌)及び鱗状である膀胱の細胞癌)。

上にリストした種類のいずれでもの部分集合の腫瘍の腫瘍細胞を含むにつれて、開示の方法はまた、サンプルを含んでいる細胞を分類するために印加されてもよい。

【0055】

部分集合のサイズは少なくともよい。そして、2、3、4、5、6、7、8、9または上で説明される腫瘍種類のうちの10から成る。あるいは、部分集合のサイズは、セットするもののいっばいのサイズまでのどんな整数でもあってもよい。このように、開示の実施形態は、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51、52、53または上記の種類の中の54の中に分類を含む。ある実施形態では、部分集合は同じ組織または器官種類のある腫瘍種類から成る。あるいは、部分集合は異なる組織または器官の腫瘍種類から成る。開示が腫瘍種類のいかなる特定のコンビネーションにも、基づかないことは、及び上記の54の既知の種類全ての可能な組合せが明白に考えられることは、過度に強調されることができない。開示の実施形態に、54の腫瘍種類の全ての組合せを明示的に書く有限の可能性は、発見及び開示の要旨の上に、任意の形態を必要とすることになっている。

【0056】

上記の腫瘍種類のある部分集合の中の分類がMaほか(Arch. Pathol. Lab. Med., 130:465-473, 2006)と同様に米国特許公報US2006/0094035及びUS2007/0020655中で報告されているが、即座の開示が少なくともadrenal corticalな腫瘍、唾液の腺の腫瘍、扁平上皮細胞癌、neuroendocrine pancreasガン、merkel細胞癌、肺カルチノイド、最初のneuroectodermalな腫瘍、性的コード間質腫瘍、thymicな癌/thymoma、膀胱の腺癌及び膀胱の扁平上皮細胞癌の第1の注意している成功した識別であると思われる。したがって、及びある実施形態で、一群の既知の腫瘍種類は、このリストから一つ以上の種類を含む。

【0057】

開示は、およそ10以上、およそ15以上、およそ20以上、およそ25以上、およそ30以上、およそ35以上の発現レベルで実施されてもよい、人間の「transcriptome」(ゲノムの部分を写した)中の見つかるにつれて、およそ40以上、およそ45以上またはおよそ50以上はシーケンスを写した。開示のある実施形態で、写された遺伝子がランダムにつつかれてもよくてまたは全てを含んでもよいことまたはあるここで開示される特定の遺伝子シーケンスの。およそ55%、およそ60%、およそ65%、およそ70%、およそ75%、およそ80%、およそ85%、およそ90%またはおよそ95%のまたはより精度による分類は、実行されることができ即座の開示の使用する。

10

20

30

40

50

【 0 0 5 8 】

他の実施形態では、他の遺伝子シーケンスの遺伝子発現レベルは、分類のために、発現レベルの上記の説明された測定とともに決定されてもよい。これの1の非限定的な例は遺伝子発現を決定するためにマイクロアレイベースのプラットフォームの場合見られる、そこで、他の遺伝子シーケンスの現れはまた、計測される。それらの他の発現レベルが既知の腫瘍種類中の発現に比較中の使用されない所で、彼らが考えられてもよくて、「過剰な」転写配列の及び開示の実行に重要でない結果。この他に、及びそれらの他の発現レベルが分類中の使用されるところ、彼らは開示の範囲内である。そして、事項を使用する説明がシーケンスの計算するところが必ずしも除外するというわけではない使用する追加のシーケンスの発現レベルの。ある実施形態では、開示は含む使用する、開示の方法中の使用される一つ以上の他の遺伝子シーケンスに、冗長な情報を提供してもよいそれらのような、一つ以上の「過剰な」遺伝子シーケンスからの発現レベルの。

10

【 0 0 5 9 】

サンプルの分類本質的に上記の腫瘍種類のうちの1もの細胞を含むことはサンプルの組織または器官サイト起源を分類する、サンプルが得られた対象の特定の組織または器官サイトであるにつれて、開示の方法は腫瘍サンプルの分類に印加されてもよい。サンプルが別の腫瘍による転移の結果である腫瘍であるケース中の、開示のこの出願は、特に役に立つ。開示のある実施形態で、以下の30の既知の組織種類で一つであるにつれて、腫瘍サンプルは分類される：副腎組織、脳組織、胸組織、子宮頸部の組織、胆管組織、子宮内膜組織、食道の組織、胃腸内組織、ずうずうしさ膀胱組織、胃食道の組織、性細胞組織、頭及び首組織、腸の組織、腎臓組織、肝臓組織（肺組織、リンパ球、メラニン形成細胞、髄膜の組織、中皮の組織、神経内分泌系組織、卵巣組織、すい臓の組織、前立腺組織、肉腫、性的コード間質組織、皮膚組織、胸腺組織、甲状腺の組織及び膀胱組織の組織）。

20

【 0 0 6 0 】

サンプルを含んでいる細胞の分類本質的に上の開示された54の腫瘍種類のうちの1の腫瘍細胞を有することはサンプルの組織または器官サイト起源を分類する。例えば、サンプルの識別必然的に子宮頸部の扁平上皮細胞癌であることは腫瘍を分類する、子宮頸部の起源（扁平上皮細胞種類（及びこのように起源中の *non epithelial* であるというよりはむしろ上皮に関する））の。それも、腫瘍が必ずしも起源内で性細胞であるというわけではなかったことを意味する。このように、対象または患者の特定の組織または器官サイトであるにつれて、開示の方法は腫瘍サンプルの分類に印加されてもよい。サンプルが別の腫瘍による転移の結果である腫瘍であるケース中の、開示のこの出願は、特に役に立つ。

30

【 0 0 6 1 】

サンプルを含んでいる細胞を分類する開示の実行、1を受け入れるために監督された学習を利用する適切な分類アルゴリズムの、上記の種類の中の1の細胞がある腫瘍を有することは、使用する）複数の既知の腫瘍種類中の遺伝子シーケンスの現れのレベルセットされるトレーニング及び2に）腫瘍種類のうちの1の細胞を有することと同じサンプルを分類するサンプルの一つ以上の細胞中の遺伝子の発現のレベル。そのようなアルゴリズムは、熟練した人に知られていて及び他の場所で説明された。発現のレベルはどんなフォーマット中でもの信号に基づいて提供されてもよい。そして、ここで説明されるにつれて、核酸発現またはタンパク質発現を含む。

40

【 0 0 6 2 】

開示の実施形態は、方法の使用及び患者からガンの原因を確認するためにここで説明される材料を含む。このように腫瘍細胞を含んでいるサンプルを与えられて、腫瘍細胞の組織起源は、本開示の使用によって確認される。1の非限定的な例は、ガン細胞を含んでいる燃やされたリンパ節で、対象の場合ある。細胞がティッシュからあってもよい、あるいは、リンパ節またはそれへの流れ出る器官が別の組織源からあってもよい。特定の腫瘍または組織種類（あるいは、起源）であるにつれて、本開示は細胞を分類するために使用されてもよい、そしてそれは、ガン細胞の源の識別は可能にする。他の非限定的な例では、

50

サンプル（例えばリンパ節からのそれ）が細胞を含むこと、そしてそれは第1である分析する少なくとも一つの細胞を分類する開示の使用する組織種類または起源の腫瘍細胞であること。これは、それからサンプル中のガン細胞の源を確認するために使用される。これらの両方とも、他のガンの使用中の時間、労力及びコストに診断上のテストを取っておく開示の有利な使用の例である。

【0063】

更なる実施形態では、開示はガンの前の履歴で対象からサンプルで実施される。非限定的な例に、サンプルを含んでいる細胞（リンパ節からまたは他の場所で）、細胞がそれから同じものまたは前のガンで異なるティッシュからそうであるかどうかに関係なく、本開示が決定するために使用されてもよいようなものが、ガン細胞を含むとわかって、対象のあってもよい。開示のこの出願は、また、新規主要な腫瘍（例えば新規ガン細胞が以前に乳ガンを有した対象の肝臓中の見つかるケース）を確認するために使用されてもよい。開示は、新規ガン細胞を確認するために使用されてもよい前の乳ガン（あるいは、別の腫瘍種類から以前に、確認されるかどうかに関係なく）からの転移の結果であることまたは肝臓ガンの新規主要な発生に。開示は、また、ティッシュのサンプルまたは複数のガンが、ガンが同じ起源のあるかどうかだけでなく、各ガンの原因を決定するとわかる器官に印加されてもよい。

10

【0064】

開示がランダムなグループの表された遺伝子シーケンスの発現レベルを使って実施されてもよい間、開示も開示の実行のために、典型的な遺伝子シーケンスを提供する。開示は、5以上が開示の実行中の使用されてもよい一群の87の遺伝子シーケンスを含む。遺伝子シーケンスは、追加のシーケンスの87でセットするものからのシーケンスが使用される遺伝子の、発現が水平になる限り発現レベルの測定とともに使用されてもよい分類する。開示のそのような実施形態の非限定的な例は、あるどこで 発現87の遺伝子シーケンスのうち5以上からの複数の他のシーケンスの発現レベルとともに計測するによってベースのプラットフォームが開示された方法を実行するために使用したマイクロアレイの、使用する。それらの他の発現レベルが分類中の使用されない所で、彼らが考えられてもよくて、「過剰な」転写配列の及び開示の実行にオプションの結果。この他に、及びそれらの他の発現レベルが分類中の使用されるところ、彼らは開示の範囲内であるどこで 使用する説明されたシーケンスがそうしない上記の必ずしも、除外する使用する追加のシーケンスの発現レベルの。

20

30

【0065】

代表的な及び非限定的に、開示の実行のために、一組の87の遺伝子シーケンスと一致しているmRNAシーケンスは、米国特許公報US2006/0094035及びUS2007/0020655中の以前に報告された。ATFが腹水腫瘍流体を示す以下のテーブルのそばで、登録番号、ジーンSymbols及びDescriptionを含む、情報を確認することのリストは提供される。そして、EGFは上皮細胞増殖因子である及び、C1は慢性リンパの白血病である。

表

【0066】

熟練した人によって理解されるにつれて、上記の確認されたシーケンスのいずれでも現れの検出はどんな適切な部分でも現れまたはこれらのシーケンスの断片の検出によって実行されてもよい。望ましくは、部分はサンプルを含んでいる細胞中の表される他のシーケンスと比較して、ユニークなシーケンスを含むために十分に大きい。更に、熟練した人は、それを認識する開示されたシーケンスが、倍の座礁された分子の1の岸を表す及びどちらの岸でも、検出されてもよい開示されたシーケンスの現れのインジケーターに。これは、分子をcDNAに望ましくは変換される細胞中のRNA分子が操作及び検出の軽くなるにつれて、開示されたシーケンスが表されるからである。結果として生じるcDNA分子は、それに相補的な岸のそれらと同様に表されたRNAのシーケンスを有してもよい。このように、RNAシーケンス岸または相補的な岸は、検出されてもよい。もちろん

40

50

、cDNAへの転換なしで表されたRNAを検出することが可能である。

【0067】

開示のある実施形態で、発現が遺伝子シーケンスの水平になること提供される登録番号によって示されるにつれて、開示された遺伝子シーケンスのオリゴヌクレオチドに雑種を産むにつれて、サンプルを含んでいる細胞中の表されたシーケンスの検出によって計測されてある。

【0068】

追加の実施形態では、開示は開示の方法中の任意の数である遺伝子の使用のために87でセットするもののシーケンスを提供する。このように、1から87の遺伝子シーケンスの全てへのどんな整数でも、開示の実行中の使用されてもよい。

10

【0069】

ここで使用される、「腫瘍サンプル」または「サンプルを含んでいる腫瘍」または、「サンプルを含んでいる腫瘍細胞」またはそのパリエーションは、または危険な状態に、発達することでガンに悩むと疑われる個人から分離される組織または流体のサンプルを含んでいる細胞を参照する。サンプルは、熟練開業医によって望ましいと考えられるにつれて、既知の方法または他の適切な方法によって孤立してもよい腫瘍細胞を含んでもよい。これらは含む、しかし、制限する(顕微解剖)、レーザーは顕微解剖(LCM)またはレーザー顕微解剖(LMD)を捕らえる即座の開示中の使用する。あるいは、組織の「地域」の中の精査されてない細胞は、使用されてもよい。そのようなサンプルの非限定的な例は、予備選挙を含む孤立させ(洗練された細胞と対照的に)て及び、管洗浄を含むがこれに限らず、非侵入性のまたは最小限に侵入する手段が針抱負、針生検、装置及び認識される米国特許第6,328,709号または他のどの適当な手段中もの説明される方法に罰金を科すいづれにでもよって、集められてもよい技術。あるいは、外科の生検を含むがこれに限らず、サンプルは侵入する方法によって集められてもよい。

20

【0070】

転写配列の検出及び計測は、技術中の知られている様々な手段によって達成されてもよいまたは熟練開業医によって適切であると考えられるにつれて。本質的に、量的に、あるいは、質的に、分析評価が反映する限り、どんな分析評価方法でも使用されてもよい。そして、転写配列の現れのレベルが検出される。

【0071】

腫瘍サンプルを分類する能力は遺伝子シーケンスの現れのレベルの関連の認識によって提供される(選択されてランダムにまたは特定である)及び、発現の実際のレベルは分析法の形態によって決定しりはしでない。ここで開示されるにつれて、量的に、あるいは、質的に、分析評価が「transcriptome」(ゲノム中の遺伝子の写された分数)または「proteome」(ゲノム中の表された遺伝子の翻訳された分数)中の遺伝子の発現を反映する限り、開示の分析評価は個々の遺伝子シーケンスの特徴を確認しているいづれでも利用してもよい。追加の分析評価は、関連したメンバーのポリペプチド断片またはproteomeのメンバーの検出に基づくそれらを含む。後者の非限定的な例は、生物学的流体(例えば血または血清)中の見つかるタンパク質分解断片の検出を含む。

30

40

特徴を確認することは、含む、しかし、制限する(シーケンスがコード化したものである(DNA)ユニークな核酸)または急行(RNA)、言われた遺伝子または特定のエピトープ ためにまたは活動の、遺伝子シーケンスによってコード化されるタンパク質。

【0072】

追加の手段は、核酸拡大の増やされた発現レベル及び核酸不活化(削除)を示し検出を含むまたはメチル化減少された発現レベルを示し。開示がDNA template(s)の一つ以上の態様を分析することによって実施されてもよいことを様々に述べる基礎をなす、各遺伝子シーケンスの、RNAの発現は、使用した、またはシーケンス(そのような製品のタンパク質分解断片と同様に)によって表されるproteinaceousな製品の、シーケンスを表す中間物に。このため存在の検出の、達する、安定性のまたは退廃

50

(率を含むこと)の(そのようなDNA)、RNA及びproteinaceousな分子は、開示の実行中の使用されてもよい。

【0073】

ある実施形態では、遺伝子シーケンスの全部または一部は、拡大されてもよい及びポリメラーゼ連鎖反応(PCR)のような方法及び、例えば、そのバリエーションによって検出されてもよいが、制限されてはならないために(量的PCR(Q-PCR))コピーPCR(RT-PCR)及びリアルタイムPCRを翻す(を含むmRNAの最初の量を計測する手段が、サンプル中の各シーケンスのためにコピーする)(任意にリアルタイムRT-PCRまたはリアルタイムQ-PCR)。そのような方法は遺伝子シーケンスの部分に相補的である1、2の下塗りを利用する、そこで、下塗りには主要な核酸合成に使用される。新しく合成された核酸は、任意にラベルをつけられて及び開示のポリヌクレオチドに交雑によってか直接検出されてもよい。彼らの交雑を可能にする状況の下で、新しく合成された核酸は、開示のポリヌクレオチド(遺伝子シーケンスを含むこと)で連絡されてもよい。表された核酸の発現を検出する追加の方法は、液体位相交雑を含むRNA-seq分析評価及び細胞の原位置ハイブリッド形成を含む。

10

【0074】

あるいは、米国特許7364846B2(完全に、まるで述べられるように、どちらが参照によってこれによって組み込まれる)中の開示されるにつれて、FFPEサンプル内で遺伝子シーケンスの現れは検出されてもよい。一時的に、表された遺伝子シーケンスの全部または一部またはコピーの現れは、検出されてもよいによって交雑調停された検出の使用(例えば制限されない、ために、マイクロアレイ、ビーズまたは小片は、基礎を形成した技術)または量的PCR調停された検出(例えば制限されない、本当にPCR及び逆転写酵素PCR時間を計る)非限定的な例に。表されたポリペプチドの全部または一部の現れは、免疫組織化学技術または他の抗体調停された検出の使用によって検出されてもよい(例えば制限されない、ために(特に他のポリペプチドと比較してポリペプチドの少なくとも分かれるために縛るラベルをつけられた抗体の使用))非限定的な例に。遺伝子発現の解析する追加の手段は、分析評価の範囲内で発現の検出を含む利用できる全世界のまたは近い全世界の、サンプル(例えばマイクロアレイ上に例えば解析の輪郭を描いている遺伝子発現の一部として)内で遺伝子発現。

20

【0075】

発現を決定するベースの分析評価が強烈な支持上に一つ以上の遺伝子シーケンスの固定することを含む核酸を使用している実施形態では、しかし、制限する、固いサブストレート配列にまたは、技術中の知られているにつれて、技術はビーズまたはビーズに基づいた。あるいは、技術中の知られている解決ベースの発現分析評価は、また、使用されてもよい。ユニークでまたはさもなければgene(s)に特定であるポリヌクレオチドの形で、sequence(s)がそうしてもよい固定された遺伝子は、あるポリヌクレオチドが言われたgene(s)のDNAまたはRNAに雑種を産むことができるようである。これらのポリヌクレオチドはgene(s)の全長であってもよくてまたは遺伝子の短いシーケンスであってもよい(削除によって技術中の知られている全長シーケンスより短い1のヌクレオチドまで上がる5'または3'シーケンスの終端)、任意に最小限にそのようなその交雑の一致しているDNAまたはRNAで腰を折られる(例えば不適当な組合せまたは挿入されたnon-complementary base pairsによって)遺伝子変化させる。ある実施形態では、使用されるポリヌクレオチドは、そうであるから3'遺伝子の、例えばおよそ350、およそ300、およそ250、およそ200、およそ150、およそ100の範囲内でまたは遺伝子または表されたシーケンスのポリアデニル化信号またはポリアデニル化サイトからの50のヌクレオチドについて終端。突然変異の存在がまだ交雑が見つけられる信号を生成することを可能にする限り、開示された遺伝子のシーケンスと比較して、突然変異を含んでいるポリヌクレオチドはまた、使用されてもよい。

30

40

このように、本開示の実行は、開示されたシーケンスの間のマイナーな不適当な組合せ及

50

び対象のサンプルの細胞によって表されるそれらの存在に影響されない。そのような不適当な組合せの存在の非限定的な例が、種（例えばヒト *sapiens* の範囲内の個々の人間の患者）の個人の間で、シーケンス多形性の場合見られてある。

【0076】

当業者に知られているにつれて、ある遺伝子シーケンスは3'を含む、開示されたシーケンスのユニークさに貢献するポリA（あるいは、相補的な岸上にポリエステル繊維T）は、伸びない。開示は、このように不足している遺伝子シーケンスで実施されてもよい3'、ポリA（あるいは、ポリエステル繊維T）は引き伸ばす。開示されたシーケンスのユニークさは、見つかるユニークなシーケンスを含む核酸だけ中の見つかるシーケンスの部分または全部を参照する3'その翻訳されていない部分。開示の実行のためのユニークなあるシーケンスは、ユニークなシーケンスが多形性のために特定のよりむしろ様々な個人中の、発現を検出すること中の役に立つようなものはある個人中の提示する遺伝子のために、共通配列に貢献するそれらである。あるいは、個人または部分母集団に独特のシーケンスは、使用されてもよい。ここで説明されるにつれて、ユニークなシーケンスは開示のポリヌクレオチドの長さであってもよい。

10

【0077】

開示の追加の実施形態では、シーケンスを有しているポリヌクレオチドは、中で提示する3'、遺伝子シーケンスの翻訳されていないおよび/または non coding している領域は、開示のサンプルを含んでいる細胞中の、発現レベルを検出するために使用されてある。そのようなポリヌクレオチドは、任意に見つかるシーケンスを含んでもよい3'遺伝子シーケンスのコーディング領域の部分。コーディング及び3'からシーケンスの組合せを含んでいるポリヌクレオチド非コーディング領域は、望ましくは、heterologousなsequence(s)に介入することなしで、隣接して手配されるシーケンスを有する。

20

【0078】

あるいは、開示は本シーケンスを有することをポリヌクレオチドで実施されてもよい5'開示の細胞及びサンプル中の発現のレベルを検出する遺伝子シーケンスの翻訳されていないおよび/または non coding している領域。そのようなポリヌクレオチドは、任意に見つかるシーケンスを含んでもよい5'コーディング領域の部分。コーディング及び5'からシーケンスの組合せを含んでいるポリヌクレオチド非コーディング領域は、heterologousなsequence(s)に介入することなしで、隣接して手配されるシーケンスを有してもよい。開示は、また、遺伝子シーケンスのコーディング領域中の、本シーケンスで実施されてもよい。

30

【0079】

ある実施形態のポリヌクレオチドは、3'からシーケンスを含むまたは5'およそ16、少なくともおよそ18、少なくともおよそ20、少なくともおよそ22、少なくともおよそ24、少なくともおよそ26、少なくともおよそ28、少なくともおよそ30、少なくともおよそ32、少なくともおよそ34、少なくともおよそ36、少なくともおよそ38、少なくともおよそ40、少なくともおよそ42、少なくともおよそ44または少なくともおよそ46の連続的ヌクレオチドの少なくとも翻訳されていないおよび/または non coding している領域。「利用できる」期間、中で使用されるにつれて、前の文は定まった数値から1の増加または減少まで参照する。実施形態が少なくともシーケンスを含んでいるポリヌクレオチドを使用する他または、少なくともまたは100について少なくともまわりに、およそ50または、少なくともまたは、少なくともまたは、少なくともまたは、少なくともまたは、または少なくともまたは400の連続的ヌクレオチドについて、350について、300について、250について、200について、150。「利用できる」期間、先立つこと中の使用されるにつれて、文は定まった数値から10%の増加または減少まで参照する。

40

【0080】

シーケンスから3'または5'遺伝子コーディング領域の終端開示のポリヌクレオチド

50

中の見つかる、彼らが当然コーディング領域の長さによって制限されること以外は、それらの説明された上記と同じ長さである。3'、コーディング領域の終端は、シーケンスを含んでもよい3'コーディング領域の半分。逆に5'、コーディング領域の終端は、シーケンスを上を含んでもよい5'コーディング領域の半分。もちろん、上記はシーケンスを説明した、あるいは、その部分を含んでいるコーディング領域及びポリヌクレオチドは全体として使用されてもよい。

【0081】

開示の別の実施形態でヌクレオチドの削除を含んでいるポリヌクレオチドから5'および/または3'、遺伝子シーケンスの終端は、使用されてもよい。削除は、望ましくは15、5 10、10 15、15 20、20 25、25 30、30 35、35 40、40 45、45 50、50 60、60 70、70 80、80 90、90 100、100 125、125 150、150 175または175 200ヌクレオチドであるから5'および/または3'削除の範囲が当然発現レベルの検出のためにポリヌクレオチドを使用することができるシーケンス及び必要の長さによって制限されるが終端。

10

【0082】

開示の他のポリヌクレオチドから3'遺伝子シーケンスの終端が、量的PCRのために下塗り及びオプションの調査のそれらを含む。好ましくは、プライマーおよびプローブは、約750より少ない領域を増幅するもので、約未満、約700未満、約650未満、約600未満、約550未満、約500未満、約450未満、ある400未満、約350未満、約300未満、約250未満、約200未満のポリアデニル化シグナルまたは遺伝子のポリアデニル化部位から、または発現さ約50ヌクレオチド未満、約100未満、または、約150未満シーケンス。少なくともあるいは、100について少なくともまわりに、開示のampliconがどんなサイズでもの50についてか少なくとも含んでもよい、あるいは、少なくともあるいは、200について少なくとも、あるいは、少なくともあるいは、300について少なくとも、あるいは、あるいは、少なくともあるいは、400の連続的ヌクレオチドについて全くPCR下塗りに相補的な部分の包含で、350について、250について、150が使用したPCRのサイズ。

20

【0083】

開示の実行に用いられる他のポリヌクレオチドは、彼らの発現を検出するシーケンスが交雑技術の使用する遺伝子に、十分な異体同形を有するものを含む。そのようなポリヌクレオチドは、望ましくはまわりに有する95%または利用できるまたは96%利用できる97%またはおよそ98%またはまたはおよそまたは使用される遺伝子シーケンスによる99%のアイデンティティ。アイデンティティは、先に述べたように、BLASTアルゴリズムを使用して決定される。開示の実行に用いられる他のポリヌクレオチドは、また、およそ50%のホルムアミドに及び交雑のための約0.01Mから約0.15M塩への及び約55 から約65 での洗浄状況のためのおよそ0.15 Mの塩までおよそ0.01 Mからおよそ30%のv/vの厳しい条件の下で、開示のポリヌクレオチドに雑種を産む能力に基づいて説明されてもよい。それにそれ以上または等しい状況をまたは。

30

40

【0084】

開示の更なる実施形態では、少なくとも人前の言われた集団が一方または両方に交雑してもよいようなものは量的に核酸分子の立ち往生させる調査が細胞のRNAまたは開示のサンプルから詳細に述べたにつれて、ヒト遺伝子シーケンスの一方または両方の岸を含んでいる単一の立ち往生する核酸分子の集団は提供される。集団はヒト遺伝子シーケンスのアンチセンスの岸だけであってもよいそのようなその感覚分子の座礁してまたは詳細に述べた、細胞は一部の言われた集団に雑種を作られてもよい。集団は、相補的な遺伝子シーケンスを含んでいる表された(あるいは、拡大された)核酸分子の量に、比較中の望ましくはヒト遺伝子シーケンスの言われた一方または両方の岸の十分に過剰な量を含む。

【0085】

50

開示はさらに5以上の発現レベルを検出することによって人間の腫瘍サンプルを分類する方法を提供する。そして、一つ以上の他の人間の腫瘍種類を除外して、人間中の見つかる腫瘍種類の腫瘍細胞を含むにつれて、50以上、核酸中の転写配列または細胞が任意に人間の対象から得られるサンプルを含んで及びサンプルを分類する。ある実施形態では、方法はサンプルを分類するために使用されてもよいまたは細胞を有するの、他の一つ以上を除外してより上にリストされる54の腫瘍種類のうちの1。

【0086】

発現を検出することによってここで説明される可能な腫瘍種類の部分集合のうちの1であることは5以上または任意に50以上の水平になるにつれて、開示も腫瘍サンプルを分類する方法を提供する。そして、一つ以上の他の人間の腫瘍種類を除外して、人間中の見つかる多くの腫瘍種類の一つであるにつれて、核酸中の転写配列が人間の対象から得られる腫瘍サンプルを含んで及びサンプルを分類する。開示のある実施形態で、他の腫瘍種類の数は、1からおよそ3まで、より望ましくは1から、およそ7への1から、およそ5まで、あるいは、1からおよそ9またはおよそ10までである。他の実施形態では、腫瘍種類は、上でリストされるそれらのような同じ組織または器官起源の全てである。

【0087】

追加の実施形態では、孤立するか、浄化されて違っているか離れている開示が単細胞からの遺伝子発現または切開された同種の細胞集団を解析することによって実施されてあってもよい。そして、単純な生検中の本、サンプルの細胞を汚染する。これらの実施形態によって提供される1の利点はそれである汚染する、そうに、non tumor細胞（例えば浸透しているリンパ球または他の免疫性システム細胞）は除去されてもよい、ここで提供されるにつれて、確認される遺伝子または遺伝子発現レベルの以降の解析を変化させることを休んでいなさい。生検が遺伝子発現プロファイルを生成するために使用される所で、そのような汚染は本。

【0088】

Q PCRまたは逆転写酵素Q PCRを利用している開示の更なる実施形態では分析評価プラットフォームに、開示の遺伝子シーケンスの発現レベルは同じサンプル中の参照遺伝子の発現レベルと比較されてもよいまたは、発現レベルの比率は使用されてもよい。これは、複数の既知の腫瘍種類上にデータ及び分析されるサンプルを含んでいる細胞の比較のために、発現データを「標準化する」ために手段を提供する。更に、複合的なフォーマットの使用で全体または一部のQ PCRは、実行されてもよい。

【0089】

追加の態様では、本開示によって提供される全体または一部の方法は、また、自動化されてもよい。これは、ソフトウェア中の開示の実施形態を含む。非限定的な例は、一つ以上のコンピュータ読取り可能な記憶装置上に実行可能な命令がそこで、命令がここで説明されるにつれて、遺伝子発現に基づく腫瘍サンプルの分類が平らにすると指令すると言ったプロセッサを含む。一つ以上のコンピュータ読取り可能な記憶装置上に追加のプロセッサ実行可能な命令は、考えられたそこで言われた命令原因発現および/または、分類方法のプロセスまたは結果のコンピュータ出力装置を介して、操作である。

【0090】

開示は、複数の既知の腫瘍種類中の一組の遺伝子シーケンスの遺伝子発現データが実装されるソフトウェア及びハードウェア実施形態を含むデータセットに。ある実施形態では、遺伝子発現データセットは、開示の方法の実行のために使用される。開示も、ここで開示される方法を実行するためにコンピュータ関連した手段及びシステムを提供する。ある実施形態では、サンプルを含んでいる細胞を分類することのための装置は、提供される。そのような装置はここで説明されるにつれて、遺伝子発現データセットを格納するように構成される質問ストレージを受けるように構成される質問入力を含んでもよい。そして、質問入力から受け取られる；及び分類アルゴリズム中のアクセスして及びストレージからデータを使用するためのモジュールここで説明される。装置は、分類アルゴリズムの結果のために、ストリングストレージを更に含んでもよい任意に出力アルゴリズム中のアク

10

20

30

40

50

セスして及びストリングストレージからデータを使用するためのモジュールで ここで説明される。

【0091】

プロセッサで実行されるソフトウェアモジュール中の、あるいは、二つの組合せ中の、方法、プロセスまたはここで開示される実施形態と関連して説明されるアルゴリズムのステップは、ハードウェア中の直接実装されてもよい。方法またはプロセス中の種々のステップまたは行動は、示される順序中の実行されてもよいが、別の順序中の実行されてもよい。さらに、一つ以上の方法ステップは省略されてもよい、あるいは、一つ以上の方法ステップは方法及びプロセスに加えられてもよい。

【0092】

追加のステップ、ブロックまたは動作は、始め、終端または方法及びプロセスの介入している既存の要素中の加えられてもよい。開示の更なる態様は、本開示の使用のために、臨床活動に対する関係性を提供することを含む。ある実施形態、測定または遺伝子発現の計測でここで説明される、治療を提供することを支持して診断上のサービスを提供することを含む患者に、治療を提供することの一部として実行する。このように、開示は患者の治療中の方法を含む。そして、ここで説明されるにつれて、サンプルを含んでいる細胞中の決定するか、遺伝子シーケンスの発現レベルを計測することを含んでいる方法が患者から得られる。方法はサンプルを分類することを更に含んでもよい。そして、ここで説明されるにつれて、方法中の腫瘍種類または組織起源の腫瘍細胞を含むにつれて、測定/計測に基づく

10

20

【0093】

ここで説明されるにつれて、測定および/または分類がどんな態様にでも対する関係または開示の実施形態のために、あってもよい。発現レベルの測定または計測は、様々な関連した動作によって先行されてもよい。ある実施形態では、計測は言われた計測の必要の場合のように人間の対象の測定または診断の後にある。計測は、計測、例えばとてもそばに医者、看護婦または他の健康管理プロバイダーの必要のまたはプロの測定によって先行されてもよいまたは計測のパフォーマンスを承認すること中の彼らの命令または健康保険またはメンテナンス組織の人員の下で働いているそれら 返済を要請するための根拠またはパフォーマンスに対する報酬に。

【0094】

計測は、また、計測しているドキュメンタリーに必要な予備の行動によって先行されてもよい。非限定的な例は、人間の対象からサンプルを含んでいる細胞を実際に得ることを含む；あるいは、サンプルを含んでいる細胞の受領；あるいは、サンプルを含んでいる細胞を区分すること；あるいは、細胞をサンプルを含んでいる細胞から分離すること；あるいは、サンプルを含んでいる細胞の細胞からRNAを得ること；あるいは、サンプルを含んでいる細胞の細胞からRNAを写すことを逆にしなさい。開示の実行のためにここで説明されるにつれて、サンプルはいずれでもあってもよい。開示は測定のためにさらにキットを提供する、あるいは、ここで説明されるにつれて、遺伝子発現の計測はサンプルを含んでいる細胞中の水平になる。

30

【0095】

本開示の実行のために、ここで説明されるにつれて、キットは遺伝子発現を検出するために通常、一つ以上の試薬を含む。非限定的な例は発現レベルの検出のためにポリヌクレオチド調査または下塗りを含む。そして、一つ以上の酵素が開示の実行のために、開示及び一つ以上のチューブの方法中の使用される。ある実施形態では、ここで説明されるにつれて、キットは配列または配列への集められることができる、遺伝子発現の検出のための固い媒体を含む。他の実施形態では、キットは遺伝子シーケンスの現れを示すポリペプチド上に、本エピトープで免疫反応性である一つ以上の抗体を含んでもよい。ある実施形態では、抗体は抗体断片である。開示のキットはまた、教材を含んでもよい開示してまたは説明する使用するキットまたは下塗りのまたは、ここで提供されるにつれて、開示の方法中の本開示の深く探りなさい。キットは、また、キットが設計される特定の出願を容易に

40

50

するために追加の部品を含んでもよい。このように、例えば、キットはさらにラベル（例えば酵素のラベルのための酵素サブストレート、蛍光性のラベルを検出するフィルタセット、羊 anti mouse HRP のような適切な第二のラベル、またはその種の他のもの）を検出する手段を含んでもよい。キットは、さらに開示の方法のために、認識されるバッファ及び他試薬を含んでもよい。

【0096】

現在一般に開示を提供して、同じものは指定されない限り、実例として提供されて及び開示の制限している婚約者でない以下の例の参照を通して、より容易に理解される。

【実施例】

【0097】

例1：PANX3 (pannexin3) の発現レベル

人間の対象からの複数の39の既知の腫瘍種類の複数のサンプル内でPANX3 (pannexin 3) の発現レベルは決定された及び、結果は図1中の示される。各腫瘍種類中の発現レベルの範囲は、39の腫瘍種類のうちの38中の、かなり重なった。

【0098】

例2：BATF の発現レベル

例1の39の既知の腫瘍種類の多数の複数のサンプル内でBATF (基本的ロイシンジッパーコピー要因 (ATF like)) の発現レベルは決定された及び、結果は図2中の示される。各腫瘍種類中の発現レベルの範囲は、全ての39の腫瘍種類のメンバーの間で重なった。

【0099】

例3：追加の転写配列の発現レベル

ここで開示されるにつれて、発現は追加の85の転写配列の水平になる同じ39の腫瘍種類にまたがって例1及び2と同じ方法中の決定した。各転写配列のための発現レベルの範囲は、複数の腫瘍種類の間で重なるのを見られた。

【0100】

例4：追加の腫瘍種類中の転写配列の発現レベル

発現は、水平になる各87の転写配列のここで開示される、開示された54の腫瘍種類にまたがって例1及び2と同じ方法中の決定した。各転写配列のための発現レベルの範囲は、複数の腫瘍種類の間で重なるのを見られた。

【0101】

特許、特許出願及び出版を含む、ここで引用される全ての参照は、以前に、特に組み込まれるかどうかに関係なく、彼らの全部中の参照によってこれによって組み込まれる。

【0102】

現在完全に発明の内容を説明して、同じものが趣旨及び開示の及び不当な実験のない範囲から出発することなく広範囲にわたる等価パラメータ、集中及び状況の範囲内で実行されることができるとは、当業者によって理解される。

【0103】

この開示がその特定の実施形態と関連して説明される間、それが更なる修正ができると理解される。この出願はどんなバリエーションでもカバーするために意味される使用するまたは、一般に、開示の適合があとに続いて、開示が付随する技術の範囲内で、開示の及び知られて来るような本開示からの離脱または慣習法を含む原則は実施する及び、要点に印加されてもよいにつれて、特徴は先に出発した。



10

20

30

40

【 国際調査報告 】

INTERNATIONAL SEARCH REPORT		International application No. PCT/US2011/031870
A. CLASSIFICATION OF SUBJECT MATTER		
<i>C12Q 1/68(2006.01)i, C12N 15/11(2006.01)i</i>		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED		
Minimum documentation searched (classification system followed by classification symbols) C12Q 1/68; G06F 19/00; G06Q 10/00		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Korean utility models and applications for utility models Japanese utility models and applications for utility models		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) cKOMPASS(KIPO internal), PubMed, Google		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2006-0094035 A1 (ERLANDER, MARK G. et al.) 04 May 2006 See abstract; paragraphs [0023],[0026],[0027],[0059]-[0061],[0080],[0176]; claims 1,2,6-13.	1,4,10,16 2,3
A		
X	US 2007-0020655 A1 (ERLANDER, MARK G. et al.) 25 January 2007 See abstract; paragraphs [0061]-[0063],[0080],[0177]; claims 1,2,8,9.	1,4,10,16 2,3
A		
A	WO 2010-108638 A1 (BRASMUS UNIVERSITY MEDICAL CENTER ROTTERDAM et al.) 30 September 2010 See abstract; claims.	1-4,10,16
A	WO 2007-137366 A1 (TELETHON INSTITUTE FOR CHILD HEALTH RESEARCH et al.) 06 December 2007 See abstract; claims.	1-4,10,16
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed		"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family
Date of the actual completion of the international search 25 JANUARY 2013 (25.01.2013)		Date of mailing of the international search report 28 JANUARY 2013 (28.01.2013)
Name and mailing address of the ISA/KR  Korean Intellectual Property Office 189 Cheongsu-ro, Seo-gu, Daejeon Metropolitan City, 302-701, Republic of Korea Facsimile No. 82-42-472-7140		Authorized officer PARK, JEONG UNG Telephone No. 82-42-481-3478 

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2011/031870

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. Claims Nos.: 8,9
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

Claims 8,9 are unclear, since they refer to claims which are not searchable due to not being drafted in accordance with the third sentences of PCT Rule 6.4(a).

3. Claims Nos.: 5-7,11-15
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

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

1. As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

- Remark on Protest**
- The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/US2011/031870

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2006-0094035 A1	04.05.2006	None	
US 2007-0020655 A1	25.01.2007	None	
WO 2010-108638 A1	30.09.2010	GB 0904957 D0 WO 2010-108638 A9	06.05.2009 30.09.2010
WO 2007-137366 A1	06.12.2007	None	

フロントページの続き

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

(72)発明者 マ, シャオ - ジュン

アメリカ合衆国, カリフォルニア州, サン ディエゴ, カーレ マー デ アーモニア 4
4 8 2

Fターム(参考) 4B063 QA01 QA13 QA18 QA19 QQ02 QQ08 QQ28 QQ42 QQ53 QR08
QR36 QR42 QR50 QR55 QR62 QR72 QR77 QR82 QS25 QS28
QS34 QS36 QS39 QX02