

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7365899号
(P7365899)

(45)発行日 令和5年10月20日(2023.10.20)

(24)登録日 令和5年10月12日(2023.10.12)

(51)国際特許分類		F I	
G 1 6 B	25/00 (2019.01)	G 1 6 B	25/00
C 1 2 Q	1/6886(2018.01)	C 1 2 Q	1/6886 Z
G 0 1 N	33/68 (2006.01)	G 0 1 N	33/68

請求項の数 9 (全105頁)

(21)出願番号	特願2019-518269(P2019-518269)	(73)特許権者	319012152
(86)(22)出願日	平成29年10月5日(2017.10.5)		ユーイーエー エンタープライジズ リミテッド
(65)公表番号	特表2019-537108(P2019-537108 A)		英国 エヌアール4 7ティージェイ ノルウィッチ ノルウィッチ リサーチ パーク ザ レジストリー ユニバーシティ オブ イースト アングリア
(43)公表日	令和1年12月19日(2019.12.19)	(74)代理人	100101281
(86)国際出願番号	PCT/EP2017/075368		弁理士 辻永 和徳
(87)国際公開番号	WO2018/065525	(72)発明者	ルカ ボグダン - アレキサンドル
(87)国際公開日	平成30年4月12日(2018.4.12)		英国 エヌアール4 7ティージェイ ノルウィッチ ノルフォーク ユニバーシティ オブ イースト アングリア シ - ノー
審査請求日	令和2年10月3日(2020.10.3)	(72)発明者	モーション ピンセント
審査番号	不服2022-14628(P2022-14628/J 1)		英国 エヌアール4 7ティージェイ ノ
審査請求日	令和4年9月15日(2022.9.15)		最終頁に続く
(31)優先権主張番号	1616912.0		
(32)優先日	平成28年10月5日(2016.10.5)		
(33)優先権主張国・地域又は機関			

最終頁に続く

(54)【発明の名称】 癌の分類および予後

(57)【特許請求の範囲】

【請求項1】

以下を含む、前立腺癌の進行を予測するための方法：

- a) 患者から得られたサンプル中の表1から選択された少なくとも100の遺伝子の発現ステータスを決定し、患者の発現プロファイルを提供すること、ここで前記の少なくとも100の遺伝子は前立腺癌により発現レベルが変わることが知られている；
- b) ラテントプロセスデコンポジション (LPD) 分析を、患者の発現プロファイルおよび異なる患者からの同じ少なくとも100の遺伝子についての参照データセットについて行うこと；
- c) 任意に複数回、分析ステップb)を繰り返すこと；および
- d) 前立腺癌の進行を予測すること、

ここで、LPD分析が個々の患者の発現プロファイルをグループに分け、
該グループは予後不良 (DESENT) または予後不良ではない (非DESENT) ステータスに分けられ、

それぞれの発現プロファイルについて、LPD分析が個々の患者の発現プロファイルについて、全体の発現プロファイルへの各グループの貢献 (p_i) を決定する、

ここで、

患者の前立腺癌進行が、全体の患者の発現プロファイルへの予後不良 (DESENT) のグループの貢献 (p_i 値) により予測され、全体の患者の発現プロファイルへの予後不良 (DESENT) のグループの貢献 (p_i 値) が高いほど悪い予測とされる；

患者の癌サンプルの予後不良 (D E S N T) のグループに対する p 値が、少なくとも 0.1 である場合、前立腺癌の進行が予測される；または

1 または複数の参照データセットから、患者のサンプルが予後不良 (D E S N T) の癌とグループ化された時、癌進行が予測される。

【請求項 2】

前記ステップ a) が、表 1 から選択された 500 の遺伝子の発現ステータスを決定することを含む、請求項 1 記載の方法。

【請求項 3】

以下のいずれかを含む、請求項 1 または 2 記載の方法：

a . 該患者の発現プロファイルが、全体の発現プロファイルに最も貢献するグループにより個々のグループに分けられる；

10

b . 全体の発現プロファイルへの予後不良 (D E S N T) のグループの貢献が、全体の発現プロファイルへの他のグループの単一での貢献よりも大きい場合に、前立腺癌の進行が予測される；

c . 全体の発現プロファイルへの予後不良 (D E S N T) のグループの貢献、患者の腫瘍のステージ、患者のグリーソンスコアおよび / または患者の P S A スコアにより、前立腺癌の進行が予測される；

d . 1 グループのみが予後不良 (D E S N T) のステータスに分けられる；

e . P L D 分析が複数回行なわれ、P L D 分析の少なくとも 60 % の検定で、1 または複数の参照データセットから予後不良 (D E S N T) の癌とグループ化された時、癌進行が予測される；

20

f . 前記請求項 1 のステップ b) が少なくとも 2 回繰り返される；

g . 異なるランダムなシードが各クラスタリング分析に使用される；

h . 複数の遺伝子の発現ステータスを決定することが、複数の遺伝子の発現レベルを決定することを含む；

i . 統計分析を行なう前に、参照データセットに対して患者の発現プロファイルを正規化することをさらに含む；

j . 統計分析のためにその発現ステータスが決定されている遺伝子の部分集合を選択するステップをさらに含み、任意に該遺伝子の部分集合内のそれぞれの遺伝子の発現ステータスが、癌患者のサンプルで変化することが知られている；

30

k . 統計分析の前に、患者の発現プロファイルにユニークなラベルを付することをさらに含む；

l . 該参照データセット内のそれぞれの発現プロファイルの前立腺癌の予後不良 (D E S N T) ステータスが知られている；

m . 該患者の発現プロファイルが、統計分析の前に少なくとも 2 つの参照データセットと組み合わせられる。

【請求項 4】

以下のいずれかを含む、請求項 1 から 3 のいずれか 1 項記載の方法：

a) サンプルは、尿サンプル、精液サンプル、前立腺の滲出液サンプル、あるいは前立腺に由来する高分子または細胞を含んでいる任意のサンプル、全血サンプル、漿液サンプル、唾液、あるいは生検試料であり、任意にサンプルは前立腺生検、前立腺摘除あるいは T U R P サンプルである；

40

b) 少なくとも 2 のサンプルに対して方法が実行され、任意に方法は複数の患者サンプルに対して同時に行われる；

c) 1 または複数のデータセットが複数の腫瘍、あるいは患者の発現プロファイルを含み、任意のデータセットは各々少なくとも 20 の患者もしくは腫瘍発現プロファイルを含み、さらに任意に、i . 患者または腫瘍の発現プロファイルは、少なくとも 10 の遺伝子の発現ステータスについての情報を含むか、または

i i . 患者または腫瘍の発現プロファイルは、少なくとも 10 の遺伝子の発現レベルについての情報を含む。

50

【請求項 5】

表 1 にリストされたすべての 500 の遺伝子の発現レベルまたは発現ステータスを検知するための手段を含む、前立腺癌の進行を予測するためのキット：

ここで、遺伝子の発現レベルまたは発現ステータスを検知するための手段がマイクロアレイであり、該マイクロアレイは RNA、DNA または蛋白質マイクロアレイであり、該マイクロアレイは発現が検知される遺伝子に特異的な結合分子を含む。

【請求項 6】

さらに以下を含む、請求項 5 記載のキット：

- a. 1 以上のコントロール遺伝子または参照遺伝子の、発現レベルまたは発現ステータスを検知するための手段；および / または、
- b. キットを使用するための指示。

10

【請求項 7】

表 1 にリストされたすべての 500 の遺伝子の発現レベルまたは発現ステータスを検知するように構成されたマイクロアレイ。

【請求項 8】

該マイクロアレイは RNA、DNA または蛋白質マイクロアレイであり、該マイクロアレイは発現が検知される遺伝子に特異的な結合分子を含む、請求項 7 記載のマイクロアレイ。

【請求項 9】

マイクロアレイが、1 以上のコントロール遺伝子または参照遺伝子の、発現レベルまたは発現ステータスを検知するように構成される、請求項 7 または 8 記載のマイクロアレイ。

20

【発明の詳細な説明】

【技術分野】

【0001】

本発明は患者からのサンプルを使用する癌、特に前立腺癌の分類に関する。特に、本発明は潜在的に侵襲性の前立腺癌を同定する方法を提供し、どの癌が侵襲性であるか、または侵襲性になるか（従って治療が必要である）、またどれがインドレント(indolent)のままであるか（従って治療が必要ではない）を決定する。したがって、本発明は患者を予後不良であるかを同定するのに有用である。本発明によって同定された癌の特定の群はここに DESNT 癌と呼ばれる。

【0002】

30

前立腺癌の診断のための一般的な方法は血液中の前立腺特異性抗原 (PSA) の測定である。しかしながら、PSA に検知される前立腺癌の 50 - 80 % は生物学上無関係である。すなわち、治療なしでも、それらは症状を引き起さないだろう。初期の前立腺癌の根治療法は、手術または放射線療法で、顕著な癌の人が理想的にターゲットとされるべきである。その結果、残りの生物学上「無関係の」疾病を有する者は、治療の副作用を免れる。診断時の個々の前立腺癌の振る舞いの正確な予測は、現在可能ではない。また、ほとんどの場合の即時の根治療法は共通のアプローチだった。彼らを悩ませていない「疾病」のための治療の結果、多くの人にインポテンスまたは失禁が起こる。多くの予後のバイオマーカーが前立腺癌のために提案された。重要な質問は、生物学上無関係の疾病を持った者から臨床的に顕著な場合を区別するために、PSA に検知された初期の前立腺癌にこれらのバイオマーカーを適用することができるかどうかである。侵襲性の癌を初期に検知するための有効な方法は、初期の前立腺癌の管理でのパラダイム・シフトに結びつくことがある。

40

【0003】

前立腺癌の臨床の管理での重大問題は、それが高度に多様性(heterogeneous)であるということである。したがって、個々の癌の振る舞いの正確な予測は、診断の時に達成可能ではなく、本質的な過剰治療に結びつく。他の多くの癌タイプとは対照的に、グローバルな表現パターンの管理されていない分析に基づいた前立腺癌の層化が可能だったかどうかは謎のままである：乳癌（例えば ERBB2 過剰発現(overexpressing)）については、基底・管腔のサブグループは同定することができる。

50

【 0 0 0 4 】

侵襲性の癌（それは治療を要求することがある）および非侵襲性の癌（それは恐らく治療しないでおくことができ、不必要な介入からの患者への副作用も与えないでおくことができる）を同定する際に、前立腺癌のための、よりよいアシストへのもっと信頼できる診断テストの必要がある。本発明は、患者が D E S N T と呼ばれる人間の前立腺癌の予後不良カテゴリーのメンバーであるか否かを予測するための、アルゴリズムに基づいた分子診断アッセイを提供する。ある実施態様では、ある遺伝子（たとえば表 2 あるいは表 3 にリストされたもの）の発現レベルが、癌が D E S N T 癌であるか否かを予測するために単独であるいは組み合わせて使用されることができる。本発明の方法の実施によって提供される、アルゴリズムに基づいた分析および関連する情報は、前立腺癌の最適な治療の意思決定を促進する。例えば、そのような医療ツールは、侵襲性疾患を有する高い危険を有する人、したがって根本的および/または積極的治療を必要とする患者を医師が同定することを可能にするだろう。

【 発明の概要 】

【 発明が解決しようとする課題 】

【 0 0 0 5 】

本発明の発明者は、ラテントプロセスデコンポジション(Latent Process Decomposition)と呼ばれるベイズのクラスタリング手続き(LPD, Simon Rogers, Mark Girolami, Colin Campbell, Rainer Breitling, "The Latent Process Decomposition of cDNA Microarray Data Sets", IEEE/ACM Transactions on Computational Biology and Bioinformatics, vol.2, no. 2, pp. 143-156, April - June 2005) doi:10.1109/TCBB.2005.29)を適用し、4つの独立した前立腺癌トランスクリプトーム・データセット中で、一般的なプロセスを同定して D E S N T を指定した。D E S N T 癌は、遺伝子のコアセットのダウンレギュレーション(down-regulation of a core set of genes)、細胞骨格マシーナリー(cytoskeleton machinery)に含まれる多くのコード化蛋白、イオン輸送および細胞接着により特徴づけられる。P S A 再発データ(PSA failure data)とリンクした3つのデータセットと引き続く D E S N T を持った前立腺摘除患者について、癌は、非 D E S N T 患者と比較して非常に貧弱な結果($p = 2.65 \times 10^{-5}$, $p = 7.74 \times 10^{-9}$ および $p = 4.28 \times 10^{-5}$)を示した。それらが患者について非常に貧弱な結果を結果的にもたらずので、D E S N T 癌は侵襲性の前立腺癌と考えられる。結果は、人間の前立腺癌の新しい予後不良カテゴリーの存在を実証し、および療法のターゲティングでのアシスト、インドレントな疾病を持った人の治療に関連する罹患率を回避するのを支援する。R o g e r s らと異なり、発明者は、生存したグループの相関を決定するために分析を行い、各プロセスのための特徴遺伝子(signature genes)の定義を提供した。発明者は、さらに癌のはるかに大きなセットおよび多数のデータセットの分析を行い、初めて異なるデータセットにわたって共通のプロセスを同定するだけでなく、さらにこの特定のグループと癌の予後不良と不変的に関連させた。発明者は、さらに任意に腫瘍病期、グリーソンスコア(Gleason score)および P S A のような前立腺癌のための他のマーカーと組み合わせて、癌の予後を任意に決定するために使用することができる発現プロフィールを与えるためのこのプロセスの寄与を発見した。顕著には、本発明の方法は単純な階層的クラスタ形成方法ではなく、従来技術と比較して患者サンプルのはるかに詳細で正確な分析を許可する。初めて、発明者は、癌進行の信頼できる予測を許容する方法を提供した。従来技術ではそのような相関を作ることができなかったため、癌進行を検知するために先行技術の方法は使用することができなかった。

【 0 0 0 6 】

発明者は、さらに L P D と異なり、コンピューター処理集約的でなく、広範囲のデータセットおよび個々の癌に適用することができるクラシファイヤーの開発を望んだ。したがって、本発明は、さらに D E S N T 癌を同定するためにランダムフォレスト分類(random forest classification)で使用する遺伝子特徴(gene signature)を同定する方法を提供する。

10

20

30

40

50

【 0 0 0 7 】

発明者は、追加の数学的方法を使用し、予後と診断のさらなる方法を提供し、かつさらに患者を予後不良と同定するのに役立つバイオマーカーおよびバイオマーカー・パネルを提供する。

【 0 0 0 8 】

ここに使用されるように、「D E S N T」癌は予後不良の前立腺癌、および治療を要求する癌を指す。「D E S N Tステータス」は、癌が進行（あるいは、経時データとして進行していた）すると予測されるか否かをいい、D E S N Tステータスを決定するステップは、癌が進行し、従って、治療を要求するか否かを予測することをいう。進行は高いP S A、転移および/または患者の死亡を指すことがある。本発明は、患者を潜在的に予後不良と同定し、それらに治療を推薦するのに役立つ。

10

【課題を解決するための手段】

【 0 0 0 9 】

発明の最初の態様では、患者の癌（たとえば前立腺癌）を分類する方法であって、例えば侵襲性の癌（たとえば侵襲性の前立腺癌）の診断、または患者を癌の予後不良（つまりD E S N T癌の患者）と同定する、以下を含む方法が提供される：

- a) 患者から得られたサンプル中の複数の遺伝子の発現レベルを決定し、患者の発現プロファイルを提供すること；
- b) 統計のベイズのクラスタリング分析あるいは他のクラスタリング分析を、患者の発現プロファイルおよび異なる患者からの同じ複数の遺伝子についての参照データセットについて行うこと；
- c) 任意に複数回分析ステップb)を繰り返すこと；および
- d) 癌を分類し、患者が癌にかかっているかどうか決定する、あるいは患者が予後不良（つまり、患者はD E S N T癌にかかっている）かどうか決定すること。

20

【 0 0 1 0 】

この方法およびその変法は今後方法1と呼ばれる。

【 0 0 1 1 】

発明の第二の態様では、前立腺癌を分類する方法が提供され、例えば患者の侵襲性の前立腺癌の分析方法、または患者を前立腺癌予後不良と同定する方法、つまりD E S N T前立腺癌の患者と同定する方法であって、以下を含む方法が提供される：

30

- a) データセット中の個々の患者のサンプルのD E S N Tステータスが知られている参照データセット（例えばL P D分析によって決定される）を提供すること；
- b) このデータセットから複数の遺伝子を選択すること、ここで複数の遺伝子が、表2にリストされたグループから選ばれた少なくとも5、少なくとも10、少なくとも20、少なくとも30、少なくとも40、あるいは少なくとも45の遺伝子を含むか、または表3にリストされたグループから選ばれた少なくとも5、少なくとも10、少なくとも15あるいは少なくとも20遺伝子を含む；
- c) 任意に：
 - (i) コントロールとして患者のサンプル中の少なくとも1つの異なる遺伝子の発現ステータスをさらに決定すること、ここでコントロール遺伝子は表2あるいは表3にリストされた遺伝子ではない；
 - (i i) 複数の遺伝子およびコントロール遺伝子の発現の相対的なレベルを決定すること；および
- d) これらの選択された遺伝子の発現ステータスを使用し、参照データセットにスーパーバイズされた機械学習アルゴリズム（supervised machine learning algorithm）（例えばランダムフォレスト分析（random forest analysis））を適用し、D E S N T癌のための予測を得ること；
- e) 患者から得られたサンプル中の同じ複数の遺伝子の発現ステータスを決定し、患者の発現プロファイルを提供すること；
- f) 任意に、参照データセットに対して患者の発現プロファイルを正規化すること；お

40

50

よび

g) 患者の発現プロフィールへ予測を適用し、癌を分類し、侵襲性の癌の存在を決定するか、あるいは患者は予後不良かどうか（つまり、患者の癌がDESENTあるいは非DESENTか）を決定すること。

【0012】

この方法およびその変法は今後方法2と呼ばれる。

【0013】

発明の第3の態様では、癌、たとえば前立腺癌を分類する方法が提供され、例えば患者の侵襲性の癌、たとえば前立腺癌の分析方法、または患者の癌を予後不良と同定する方法、つまりDESENT癌の患者と同定する方法であって、以下を含む方法が提供される：

a) データセット中の個々の患者のDESENTステータスが知られている参照データセット（例えばLPD分析によって決定される）を提供すること；

b) このデータセットから複数の遺伝子を選択すること；

c) これらの選択された遺伝子の発現ステータスを使用し、データセットにスーパーバイズされた機械学習アルゴリズム（例えばランダムフォレスト分析）を適用し、DESENT癌のための予測を得ること；

d) 患者から得られたサンプル中の同じ複数の遺伝子の発現ステータスを決定し、患者の発現プロフィールを提供すること；

e) 任意に参照データセットへ患者の発現プロフィールを正規化すること；および

f) 患者の発現プロフィールへ予測を適用し、癌を分類し、侵襲性の癌の存在を決定するか、あるいは患者は予後不良かどうか（つまり、患者の癌がDESENTあるいは非DESENTか）を決定すること。

【0014】

この方法およびその変法は今後方法3と呼ばれる。

【0015】

発明の第4の態様では、前立腺癌を分類する方法が提供され、例えば患者の侵襲性の癌（たとえば侵襲性の前立腺癌）の診断方法、または患者を癌予後不良と同定する方法、つまりDESENT癌の患者と同定することであって、以下を含む方法が提供される：

a) データセット中の個々の患者のサンプルのDESENTステータスが知られている参照データセット（例えばLPD分析によって決定される）を提供すること；

b) このデータセットから、その発現ステータスがDESENTおよび非DESENT癌の間で相違することが知られている複数の遺伝子を選択すること、たとえば表4にリストされた複数の遺伝子、たとえば少なくとも100、少なくとも200、少なくとも300、少なくとも400、少なくとも500あるいは少なくとも1000の遺伝子を選択すること；

c) 選択された遺伝子にLASSOロジスティック回帰モデル分析を適用し、DESENT癌と同定する選択された遺伝子の部分集合を同定すること；

d) 選択された遺伝子のこの部分集合の発現ステータスを使用し、スーパーバイズされた機械学習アルゴリズム（例えばランダムフォレスト分析）をデータセットに適用し、DESENT癌についての予測を得ること；

e) 患者から得られたサンプル中の選択された遺伝子の部分集合の発現ステータスを決定し、患者の発現プロフィールを提供すること；

f) 任意に参照データセットへ患者の発現プロフィールを正規化すること；および

g) 患者の発現プロフィールへ予測を適用し、癌を分類し、侵襲性の癌の存在を決定するか、あるいは患者は予後不良かどうか決定すること（つまり、患者の癌がDESENTあるいは非DESENTかを決定すること）。

【0016】

この方法およびその変法は今後方法4と呼ばれる。

【0017】

発明の第5の態様では、癌の進行の予測、あるいは癌の分類指標として表2にリストさ

10

20

30

40

50

れた遺伝子を含むバイオマーカー・パネルが提供される。特に、表 2 にリストされた遺伝子は癌（たとえば前立腺癌）の進行を予測するために使用することができる。これらの遺伝子のダウンレギュレーションは癌進行の予測である。一般に、発明の実施態様では、表 2 からの遺伝子の、少なくとも 5、少なくとも 10、少なくとも 20、少なくとも 30 あるいは少なくとも 40 が使用される。ある実施態様では、表 2 からの 45 の遺伝子のすべてが使用されるだろう。したがって、同じバイオマーカー・パネルを使用して、他の癌タイプの進行を予測することができるが、このパネルは患者の中の侵襲性の癌の診断、特に侵襲性の前立腺癌の診断に役立つ。

【 0 0 1 8 】

発明の第 6 の態様では、癌の進行の予測、あるいは癌の分類指標として表 3 にリストされた遺伝子を含むバイオマーカー・パネルが提供される。特に、表 3 にリストされた遺伝子は癌の進行を予測するために使用することができる。一般に、発明の実施態様では、表 3 からの遺伝子の、少なくとも 5 つ、少なくとも 10、あるいは少なくとも 15 が使用される。ある実施態様では、表 3 からの 20 の遺伝子はすべて使用されるだろう。このパネルは特に前立腺癌に関連し、したがって、患者の中の前立腺癌進行を予測するのに役立つ。

【 0 0 1 9 】

発明の第 7 の態様では、癌の進行の予測、あるいは癌の分類指標として表 1 にリストされた遺伝子を含むバイオマーカー・パネルが提供される。特に、表 1 にリストされた遺伝子は癌の進行を予測するために使用することができる。一般に、発明の実施態様の中で、表 1 からの遺伝子の少なくとも 5、少なくとも 10、あるいは少なくとも 15、少なくとも 20、少なくとも 50、少なくとも 100、少なくとも 200、少なくとも 300 あるいは少なくとも 400 が使用されるだろう。ある実施態様では、表 1 からの 500 の遺伝子はすべて使用されるだろう。このパネルは前立腺癌へ特に関連し、したがって、患者の中の前立腺癌進行を予測するのに役立つ。表 1 から使用される遺伝子の選択はここに記述されるような方法を使用して決定されることができる。発明のある実施態様では、バイオマーカー・パネルは、癌のための予測の決定に関する本発明の方法によって生成される。そのような分析は、遺伝子の任意のセットで行うことができる。好ましくは、バイオマーカー・パネルが選ばれる遺伝子のセットは、少なくとも 1000 のランダムに選択された遺伝子を含む。ある実施態様では、遺伝子は、ハウスキーピング遺伝子（housekeeping genes）ではない（例えば表 6 にリストされた遺伝子ではない）。

【 0 0 2 0 】

上に定義されたパネルは総体として本明細書で「バイオマーカー・パネル」と呼ばれることがある。

【 0 0 2 1 】

発明の一層の態様では、癌（たとえば前立腺癌）について診断するか、スクリーンするか、検査する方法が提供され、特に侵襲性か D E S N T 癌か（たとえば侵襲性かまたは D E S N T 前立腺癌か）について提供され、バイオマーカー・パネルからの遺伝子のすべてまたは選択された遺伝子のサンプル内の発現のレベルを検査することを含む。ある実施態様では、生体試料は、前立腺組織生検（たとえば疑わしい腫物サンプル）、唾液、血液サンプルあるいは尿サンプルである。好ましくは、サンプルは、前立腺生検、前立腺摘除標本（取り出された前立腺）あるいは T U R P（前立腺切除術）標本からの組織試料である。

【 0 0 2 2 】

癌（たとえば前立腺癌）、特に侵襲性の癌（たとえば侵襲性の前立腺癌）の診断で使用するためのバイオマーカー・パネル中の 1 つ以上の遺伝子がさらに提供される。そのような癌を検知するか診断する方法におけるバイオマーカー・パネルの中の 1 つ以上の遺伝子の使用、およびバイオマーカー・パネルの中の 1 つ以上の遺伝子を使用してそのような癌を検知するか診断する方法を提供する。

【 0 0 2 3 】

癌（たとえば前立腺癌）、特に侵襲性の癌（たとえば侵襲性の前立腺癌）の進行の予

10

20

30

40

50

測での使用のための、バイオマーカー・パネル中の1つ以上の遺伝子がさらに提供される。癌の進行を予測する方法でのバイオマーカー・パネル中の1つ以上の遺伝子の使用、およびバイオマーカー・パネル中の1つ以上の遺伝子を使用して、癌の進行を予測する方法がさらに提供される。

【0024】

癌（たとえば前立腺癌）の分類で使用するためのバイオマーカー・パネル中の1つ以上の遺伝子がさらに提供される。癌を分類する方法におけるバイオマーカー・パネル中の1つ以上の遺伝子の使用、およびバイオマーカー・パネル中の1つ以上の遺伝子の使用する癌の分類方法がさらに提供される。

【0025】

生体試料中のバイオマーカー・パネルの1つ以上の遺伝子の発現のレベルを検知するための手段を含む、前立腺癌を検査するためのキットの一部がさらに提供される。さらにバイオマーカー・パネルの中にない1つ以上のコントロール遺伝子の発現のレベルを検知するための手段を含むことができる。

【0026】

生体試料中のバイオマーカー・パネルの1つ以上の遺伝子の発現のレベルを検知することを含む、侵襲性および非侵襲性の前立腺癌を同定する方法がさらに提供される。任意に、測定された遺伝子の各々の発現レベルは、参照と比較される。参照はコントロールまたはハウスキーピング遺伝子であることができる。ある実施態様では、コントロール遺伝子は表6あるいは表7にリストされた遺伝子から選ばれる。表7のコントロール遺伝子は前立腺癌へ特に関連する。表6のコントロール遺伝子はより広く有用である。

【0027】

さらに、侵襲性の癌の診断方法、癌の分類方法、癌の予後予測方法、癌進行の予測方法であって、生体試料中のバイオマーカー・パネル中の1つ以上の遺伝子の発現のレベルを検知を含む方法が提供される。任意に、方法は定量化された遺伝子の各々の発現レベルを参照と比較することをさらに含む。

【0028】

発明の一層の態様では、患者の前立腺癌を治療する方法であって、侵襲性の前立腺癌あるいは予後不良の癌が診断されるか疑われる場合に、前立腺癌の治療を行う方法が提供される。発明では、発明の方法のうちの1つを使用して、患者は侵襲性の前立腺癌を持っているか、予後不良であると診断される。ある実施態様では、患者の癌（たとえば前立腺癌）を診断するか、分類するか、予後不良か、癌の進行を予測する方法、あるいは患者を前立腺癌の予後不良と患者と同定（つまり、患者をDESNT前立腺癌と同定）する方法により、治療方法が進められる。

【図面の簡単な説明】

【0029】

【図1】ラテントプロセスデコンポジション(Latent Process Decomposition:LPD)、遺伝子相関および臨床の結果。

【図2】DESNT予後不良前立腺癌(DESNT poor prognosis prostate cancer)で共通してダウンレギュレートされた遺伝子

【図3】癌ゲノムアトラスデータセット(The Cancer Genome Atlas dataset)中のRF-DESNTおよび非-RF-DESNT癌の比較。

【図4】コンピュータ機器の例

【図5】ログ尤度プロット(Log-likelihood plots)

【図6】トランスクリプトーム・データセットのラテントプロセスデコンポジション(LPDP)分析。

【図7】LPDによって同定されたDESNT癌についての結果の分析。

【図8】LPD分類によって同定されたDESNT癌の遺伝子発現の相関。

【図9】20の遺伝子シグネチャー(gene signature)を使用するRF分類によるDESNT癌の検知。

10

20

30

40

50

【図 10】RF 分類によって同定された DESNT 癌のための結果の分析。

【図 11】RF 分類によって同定された DESNT 癌の遺伝子発現の相関。

【図 12】LPD 結果の分布

【図 13】MSKCC データセットの LPD。

【図 14】DESNT 癌存在のパーセンテージに基づいた前立腺癌の層化(stratification)。

【図 15】LPD について 1、3、5 および 7 年での PSA のない生存を予測するために開発されたノモグラムモデル

【図 16】LPD のためのコックス・モデル。

【0030】

表 1: 前立腺癌の大部分にわたり変わる 500 の遺伝子プロンプ。 (隅付きかっこ、表 2 - 1 から表 2 - 4 に示される。)

表 2: Cancer Map、Stephenson、MSKCC および Klein データセットからの 80 / 100 の 45 の共通してダウンレギュレートされた遺伝子。 (隅付きかっこ、表 3 に示される。)

表 3: 20 のランダムフォレスト遺伝子。 (隅付きかっこ、表 4 に示される。)

表 4: DESNT および非 DESNT 癌の間で変化する 1669 の遺伝子。 (隅付きかっこ、表 5 - 1 から表 5 - 6 に示される。)

表 5: CamCap、Stephenson、MSKSS および Klein データセットからの 67 / 100 の 35 の共通してダウンレギュレートされた遺伝子。 (隅付きかっこ、表 6 に示される。)

表 6: 一般コントロール / ハウスキーピング遺伝子。 (隅付きかっこ、表 7 に示される。)

表 7: 前立腺癌のためのコントロール / ハウスキーピング遺伝子。 (隅付きかっこ、表 8 に示される。)

【発明を実施するための形態】

【0031】

詳細な発明の説明

本発明は癌進行予測に役立つ方法、バイオマーカー・パネルおよびキットを提供する。

発明の 1 つの実施態様では、癌を分類する方法、侵襲性の癌を診断する方法、予後不良の癌 (つまり DESNT 癌の患者) を同定する方法であって以下を含む方法が提供される:

a) 患者から得られたサンプル中の複数の遺伝子の発現レベルを決定し、患者の発現プロファイルを提供すること;

b) 統計的ベジアンクラスタリング分析あるいは異なるクラスタリング分析を、患者の発現プロファイルと異なる患者からの同じ複数の遺伝子についての参照データセットについて行うこと;

c) 任意に、複数回分析ステップ b) を繰り返すこと; および

d) 癌を分類し、患者が侵襲性の癌にかかっているかどうか決定するか、あるいは患者が予後不良 (つまり、患者は DESNT 癌にかかっている) かどうか決定すること。

【0032】

この方法は前立腺癌へ特に関連するが、他の癌にも適用することができる。

【0033】

方法 1 の分析ステップ b) が繰り返される実施態様では、分析が実行されるごとに、異なる最初のランダムシードが使用されることができる。

【0034】

複数の遺伝子の発現のレベルを決定する方法 1 のステップ a) は、その発現レベルが癌により変わることが知られている遺伝子に対して実行されることができる。例えば、発現のレベルは、少なくとも 50、少なくとも 100、少なくとも 200、あるいは最も好ましくは少なくとも 500 の癌により変わることが知られている遺伝子について決定されることがある。当業者は、たとえば癌患者について以前に公表されたデータセットを使用し

10

20

30

40

50

、その発現レベルが異なる癌サンプルで変わる一群の遺伝子を選んで、どの遺伝子を測定しなければならないかを定めることができる。特に、遺伝子の選択は、それらの発現レベルが癌の種類により変わることが知られている量に基づいて決定される。

【 0 0 3 5 】

癌による変異は、同じ組織起源（例えば前立腺、胸、肺など）を有する癌について見られる発現における変異を指す。例えば、発現における変異は、同じ組織起源の癌にかかっている異なる患者から採取されたサンプル間で測定することができる発現の差である。遺伝子の選択を見る時、いくらかまたはすべてのサンプルにわたって同じか同様の発現を持つだろう。これらはほとんど変異がないか、あるいは少ない変異を持つと言われている。他のものはハイ・レベルの変異（いくつかのサンプルでは高い発現があり、他のものにおいては低い）を有する。

10

【 0 0 3 6 】

発現レベルが前立腺癌間でどれくらい変わるかの測定は、特には統計分析で、当業者に知られている多くの方法で決定することができる。例えば、当業者は、各々の複数の癌サンプル中の複数の遺伝子を考慮し、複数のサンプルにわたる発現レベルの標準偏差または相互四分位の範囲が、前もって定義した閾値を越える遺伝子を選択することができる。遺伝子はサンプルまたは患者にわたるそれらの変化によってオーダーすることができる。また、変わる遺伝子の選択を行うことができる。例えば、変化が最大の遺伝子を使用することができる、たとえば500の遺伝子が最大の変化を示す。もちろん、最大に変化する遺伝子が必ずしも使用されることが重大だとは限らない。例えば、トップの500から1000の遺伝子を使用することができることがある。一般に、変化によるものである場合、選ばれた遺伝子はすべて、遺伝子のトップ50%であるだろう。重要なことは、発現レベルが参照データセットにわたり変わるということである。遺伝子の選択は臨床の攻撃性(clinical aggression)に関係なく行われる。これはスーパーバイズされていない分析として知られている。当業者は、この目的のための遺伝子の選び方に気づいている。

20

【 0 0 3 7 】

ステップb)は、1つ以上の参照データセットの使用を要求する。参照データセットが当業者に利用可能であるので、参照データセットの調製は一般に方法の一部ではない。方法1の以前に得られた参照データセット（あるいはステップb)の中で新たに得られた参照データセットでさえ）を使用する場合、参照データセットへの患者のサンプル中の複数の遺伝子についての発現のレベルの正規化は、患者のサンプルについて得られた情報が参照データセットと比較可能だったことを保証するために要求されることがある。正規化技術は当業者に知られており、例えば完全なマイクロアレイ・データセットが利用可能な場合に、Robust Multi-Array Average、Froze Robust Multi-Array Average または Probe Logarithmic Intensity Errorが知られている。Quantile正規化も使用することができる。最初の発現プロファイルが参照データセットと組み合わせられた後に正規化を行い、正規化された組み合わせられたデータセットを提供することができる。

30

【 0 0 3 8 】

正規化の方法は、たとえば分析されたRNAの量の差および、使用されるRNAの品質の変異等を説明するために、測定されたレベルの補正を含み、すべての分析された遺伝子を比較可能なスケールにする。コントロール遺伝子（ハウスキーピング遺伝子とも呼ばれる）は、適切な条件（例えばDESNT癌）下で発現ステータスにおいて異ならないと知られているので有用である。ハウスキーピング遺伝子の例は当業者に知られており、それらはRPLP2、GAPDH、PGK1、Alas1、TBP1、HPRT、K-Alpha1およびCLTCを含んでいる。ある実施態様では、ハウスキーピング遺伝子は表6あるいは表7にリストされたものである。表7は前立腺癌へ特に関連する。発明の好ましい実施態様は、少なくとも2つのハウスキーピング遺伝子を正規化に使用する。

40

【 0 0 3 9 】

方法1のステップa)は、一人の患者からの単一の発現プロファイルを含むことができる。あるいは、診断を行う異なる患者からの2つ以上の発現プロファイルを使用すること

50

ができる。多くの患者を同時に診断する場合、そのようなアプローチは有用である。方法は、それらの発現プロファイルが分析ステップでより容易に同定されることを可能にするために、患者の発現プロファイルの各々のユニークなラベルを付するステップを含むことがある。

【 0 0 4 0 】

ある実施態様では、特には前立腺癌に関係するものでは、発現のレベルは表 1 のリストから選ばれた複数の遺伝子のために決定される。

【 0 0 4 1 】

ある実施態様では、方法 1 のステップ a) は表 1 の中のリストから選ばれた少なくとも 2 0、少なくとも 5 0、少なくとも 1 0 0、少なくとも 2 0 0 あるいは少なくとも 5 0 0 の遺伝子の発現レベルを決定することを含んでいる。遺伝子の数が増加するにつれて、テストの正確さはさらに増加する。好ましい実施態様では、5 0 0 の遺伝子は少なくともすべて表 1 の中のリストから選ばれる。しかしながら、表 1 の遺伝子に制限する必要はない。

【 0 0 4 2 】

ある場合には、さらに多くの遺伝子の発現のレベルについての情報が、方法 1 のステップ a) の中で得られることがある。たとえば非常に多くの遺伝子の発現のレベルを決定するマイクロアレイの使用によって得られる。全体のトランスクリプトームを得ることは可能である。しかしながら、その発現レベルが前立腺癌にわたって変わることが知られている遺伝子の部分集合に、引き続く分析ステップを行なうことのみが必要である。好ましくは、もし遺伝子の部分集合が癌（たとえば前立腺癌）により異なる発現レベルを伴っていれば、これは厳密に必要ではないが、使用される遺伝子は、前立腺癌によりその発現レベルが大きく変わる（つまり、発現は癌攻撃性によって変わる）ものになるだろう。

【 0 0 4 3 】

分析が行なわれる実際の遺伝子は、利用可能な発現レベル情報に依存するだろう。また、それはデータセットごとに変わることがある。この方法のステップは、遺伝子の特定のリストに制限される必要はない。しかしながら、表 1 にリストされた遺伝子は使用することができる。

【 0 0 4 4 】

したがって、方法 1 のステップ a) は、方法の残りに必要とされる遺伝子の非常に多数の決定を含んでいることがある。したがって、方法は、患者のサンプルについて、発現プロファイルから、発現レベルが前立腺癌により変わることが知られている遺伝子の部分集合を選択するステップをさらに含むことがある。部分集合は表 1 から選ばれた少なくとも 2 0、少なくとも 5 0、少なくとも 1 0 0、少なくとも 2 0 0、あるいは少なくとも 5 0 0 の遺伝子であることがある。

【 0 0 4 5 】

好ましい実施態様では、ベイズのクラスタリング分析はラテントプロセスデコンポジション (LPD) 分析である。そのような数理モデルは当業者に知られており、例えば Simon Rogers, Mark Girolami, Colin Campbell, Rainer Breitling, "The Latent Process Decomposition of cDNA Microarray Data Sets", IEEE/ACM Transactions on Computational Biology and Bioinformatics, vol.2, no. 2, pp. 143-156, April-June 2005, doi:10.1109/TCBB.2005.29、に述べられている。使用することができる代替のベイズのクラスターアルゴリズムは次のものを含んでいる：デリクレットのプロセス混合モデル (Dirichlet Process Mixture Models)、ベイズの階層的クラスタリング (Bayesian Hierarchical Clustering)、ベイズの超パラメータ再評価による多重トピック・マイクロアレイ分析 (Bayesian Multi-topic Microarray Analysis with Hyperparameter Reestimation)、ベイズの混合モデル (Bayesian Mixture Models)、LPD へのマルコフ連鎖モンテカルロアプローチ (a Markov Chain Monte Carlo approach to LPD)、マージナライズドベイズのアプローチ (a marginalized variational Bayesian approach)。

【 0 0 4 6 】

LPD分析が参照データセット（それは複数の患者についての、その発現レベルが前立腺癌により著しく変わる多くの遺伝子の発現レベルについての情報を含んでいる）上で行なわれる場合、DESENT（侵襲性が予後不良）癌をすべて示す患者の個体群を同定することは可能である。LPD分析は「プロセス」へ患者をグループ化する。本発明の発明者は、その発現レベルが前立腺癌により変わると知られている遺伝子を使用してLPD分析が行なわれる場合、患者について否定的な結果（つまりDESENT / 侵襲性の癌）を常に本質的に伴う、1つの特別の患者集団（あるいはプロセス）が同定されることを驚くべきことに発見した。さらにより驚いたことに、このプロセスは多数の異なるデータセットでも存在する。

【0047】

本発明では、発明者は、log尤度プロット（log-likelihood plots）（図1b、5）から示されるように、全体の発現プロフィールに寄与する、3から8の根本的なプロセス間に使用して、LPD分析を行なった。各データセットのデコンポジションに続いて、それらの最も高い p_i 値に基づいて個々のプロセスに割り当てられた癌は、図1aおよび図6の中で示される結果を与えた。 p_i は、個々の癌の発現プロフィールへの各プロセス i の貢献値である：すべてのプロセスでの p_i の合計 = 1である。しかしながら、必ずしも最も高い p_i 値を使用する必要があるとは限らず、 p_i は違う様に定義することができる。また、当業者は可能な変化を理解できるだろう。例えば、 p_i は、少なくとも0.1、少なくとも0.2、少なくとも0.3、少なくとも0.4あるいは好ましくは少なくとも0.5であることができる。

【0048】

確かに、実施例2に実証されるように、 p_i は連続変数で、与えられたサンプルの発現プロフィールへの与えられたプロセスの貢献の指標である。DESENTプロセスの貢献が高いほど（与えられたサンプルについて発現プロフィールへのDESENTプロセスに関する p_i 値が高いほど）、癌が不良な結果を持つという可能性は大きい。実施例2に実証され示されるように、与えられたサンプルについては、多くの異なるプロセスが発現プロフィールに寄与することができる。

DESENTプロセスが予測される不良な結果について最も優性な（つまり発現プロフィールに寄与するすべてのプロセスの中で最も高い p_i 値を有すること）ものであることは必ずしも必要としない。しかしながら、 p_i 値がより高いと、患者の結果はより悪くなり；PSAへの参照だけではなく、さらに転移と死亡はより起こりそうである。ある実施態様では与えられた癌についての全体的な発現プロフィールへのDESENTプロセスの貢献は、DESENTである癌の可能性を評価する場合に、決定されることができる。ある実施態様では、癌進行の予測は、腫瘍のステージ、グリーソンスコア（Gleason score）および/またはPSAスコアの1つ以上との組み合わせで行われることがある。したがって、ある実施態様では、癌予後を決定するステップは、発現プロフィールへのDESENTプロセスについての p_i 値を決定するステップを含むことができ、任意に、さらに、患者の腫瘍のステージ、患者のグリーソンスコア、または患者のPSAスコアの決定を含むことがある。

【0049】

ある実施態様では、個々の患者の発現プロフィールをグループ化するステップは、各発現プロフィールについてLPD分析を使用して、個々の患者の発現プロフィールについて全体的な発現プロフィールへの各グループの貢献（ p_i ）を決定することを含む（そこでは与えられた発現プロフィールに対するすべての p_i 値の合計は1である）。患者の発現プロフィールは、全体的な発現プロフィールに最大に寄付するグループによって個々のグループに割り当てられることができる（言い換えれば、患者の発現プロフィールは最も高い p_i 値を備えたグループに割り当てられる）。ある実施態様では、グループはそれぞれDESENTあるいは非DESENTステータスのいずれかに割り当てられる。患者の癌進行は、全体的な発現プロフィールへのDESENTプロセスの貢献（ p_i 値）によって予測することができる。ある実施態様では、患者の癌サンプルのDESENTプロセスに対する p

10

20

30

40

50

i 値が、少なくとも 0.1、少なくとも 0.2、少なくとも 0.3、少なくとも 0.4、あるいは少なくとも 0.5 である場合、DESENT 癌が予測される。

【0050】

図 1a では、「 p_i 」値は縦軸に示される。カラムはそれぞれ単一の患者を示す。各データセットの LPD に続いて、癌は、この図の中で示される結果を与えるそれらの最も高い p_i 値に基づいて個々のプロセスに割り当てられた。 p_i は、個々の癌の発現プロフィールへの各プロセス i の寄与である：すべてのプロセスのパイ値の合計 = 1 である。

【0051】

参照データセットは先に得られることができる。一般に、これらのデータセットを得ることは要求される本発明の方法の一部ではない。しかしながら、ある実施態様では、方法は LPD 分析へ包含される追加のデータセットを得ることをさらに含むことがある。参照データセットは、患者のサンプルの中で測定された同じ遺伝子を含む、複数の発現プロフィールの形である。

【0052】

ここに記述された方法のすべてにおいては、DESENT 癌あるいは DESENT 癌グループを同定することが可能な、4 つの主な方法がある：

(i) 別の 1 または複数のデータセットの DESENT 癌グループとの遺伝子発現レベルの相関；

(ii) コアダウンレギュレート遺伝子セット(core down-regulated gene set)と DESENT および非 DESENT 癌の間の差別的に発現した遺伝子とのオーバーラップのデモンストレーション；

(iii) より不良な臨床の結果；および

(iv) 一緒にされた参照の LPD、および患者のデータセットが既知の DESENT グループでクラスターにされる場合に割り当てられた DESENT ステータスの未知の患者のデータセットに LPD を行うこと。

【0053】

最初の (i) の方法では、LPD 分析が行なわれた後、患者のサンプルは異なるプロセスと一緒にグループ化される。各プロセスでの患者について、各遺伝子の発現のレベルは平均される。その後、DESENT 癌の存在が知られている他のデータセットからのデータと、平均された発現レベルを関連づける。従って、新しいデータセット中のプロセスが、DESENT 癌に関連しているかを決定することができる。その後、新しいデータセットは、本発明の方法の中で使用することができる。新しいデータセットが 1 つ以上の未知の患者サンプルを含んでいる場合、未知の患者のサンプル・グループが DESENT プロセスを備えているか否かの決定をなすことができる（つまり、DESENT プロセス/グループと同じプロセス/グループに割り当てられる）からである。あるいは、DESENT グループ中の平均遺伝子発現レベルとサンプル中の遺伝子の発現を関連付けすることは可能である。このように、単一の未知の標本の相関付けを実行することが可能である。

【0054】

第 2 の (ii) の方法では、DESENT 癌で異なる発現をすることが知られている（例えば、ダウンレギュレートされる）、遺伝子の参照セットを持っていることが必要である。これは、複数のデータセットについての LPD 分析を行い、各データセット中のプロセスを決定することにより、先に得られることができる。その方法では、共通のプロセスは上に記述された方法を使用して、遺伝子相関に基づいて同定することができる。各データセットの非 DESENT プロセスと、DESENT プロセスの各発現プロフィールと比較して、一貫してダウンレギュレートされる遺伝子の部分集合が同定される。本発明の発明者によって行なわれた分析では、分析された 4 つのデータセットについての LPD 分析の 100 回の検定で、少なくとも 80 回で 45 の遺伝子が最も共通してダウンレギュレートされたと分かった（表 2）。しかしながら、異なるデータセットが使用される場合、異なる遺伝子が同定されることがある。異なるデータセットが使用される時、同定された遺伝子間に相当なオーバーラップがあるだろう。例えば、データセットの異なるコンビネーション

10

20

30

40

50

を使用して、研究者によって行なわれた第2の分析では、35遺伝子が、各データセットのLPD分析の100回の内の少なくとも67回でダウンレギュレートしたと分かった(表5)。最初の分析で同定された、45の共通してダウンレギュレートされた遺伝子と、27の遺伝子にオーバーラップがあった。

【0055】

一旦コアダウンレギュレートされた遺伝子セットが得られれば、方法(i i)は実行することができる。特に、DESENT癌は、テスト・データセットからの1つのグループにおける差別的に発現した遺伝子と、コアダウンレギュレートした遺伝子のセットの間でのオーバーラップを示すことにより同定することができる。「オーバーラップ」は、50%、60%、70%、80%、90%あるいは100%のオーバーラップであることができる。好ましくは、オーバーラップは少なくとも67%である。コアダウンレギュレートされた遺伝子セットは、少なくとも5、少なくとも10、少なくとも20、少なくとも30、あるいは少なくとも40の遺伝子を含むことができる。例えば、コアダウンレギュレートされた遺伝子セットは、表1の45の遺伝子を含むことができる。

10

【0056】

発明の1つの実施態様では、したがって、癌を分類する方法であって、表2からの少なくとも5、少なくとも10、少なくとも20、少なくとも30、少なくとも40、あるいはすべての45の遺伝子の患者サンプルの発現のレベルと、健康な者または侵襲性もしくはDESENT癌にかかっていない者の同じ遺伝子の発現とレベルを比較することを含む方法が提供される。あるいは、方法は表5からの少なくとも5、少なくとも10、少なくとも20、少なくとも30、あるいは35の遺伝子すべての患者サンプルの発現のレベルを比較することを含むことがある。患者のサンプル中の遺伝子の少なくとも50%、60%、70%あるいは80%の発現のレベルがコントロールまたは参照の遺伝子よりも低い場合、DESENT癌は存在し、癌進行が予測される。

20

【0057】

新しい(テスト)データセットが1以上の未知の患者サンプルを含んでいる場合、未知の患者のサンプル・グループがDESENTプロセスを備えているか否かの決定をなすことができる。

【0058】

第3(i i i)の方法では、LPDによって同定されたDESENT癌プロセスはより不良な臨床結果、たとえば、非DESENT癌と比較された時、患者の死亡あるいは癌再発に関係する。再び、新しい(テスト)データセットが1以上の未知の患者サンプルを含んでいる場合、第3(i i i)の方法を使用して未知の患者のサンプル・グループがDESENTプロセスを備えているか否かの決定をなすことができる。

30

【0059】

第4の(i v)の方法では、単一の患者サンプルからの複数の遺伝子についての単一の発現プロファイルへLPD分析を実行し、その個々の患者がDESENT癌にかかっているかどうか決定することは可能ではない。発明の1つの方法では、患者のサンプルからの発現プロファイルがより大きなデータセットの分析に含められることが必要である。したがって、例えば、方法1のステップb)(LPD分析ステップ)は、患者の発現プロファイルおよび参照データセットに同時に行うことができる。言い換えれば、患者の発現プロファイルはLPD分析に先立って、参照データセットと結合することができる。追加の患者のサンプルがDESENT癌プロセスにグループ化される場合、患者はDESENT癌にかかっている。

40

【0060】

したがって、本発明の1つの実施態様では、方法は、

- a) 患者から得られたサンプル中の複数の遺伝子の発現のレベルを決定し、最初の発現プロファイルを提供すること、
- b) 参照データセットと最初の発現プロファイルを組み合わせて結合データセットを得ること、ここで該参照データセットは異なる患者から得た同じ複数の遺伝子のための発現

50

プロフィールを含み、任意に参照データセット中の患者の臨床の結果は知られている；

c) 該結合データセットについてのLPD分析を行なうこと；

d) DESNT癌に関係しているLPD分析からプロセス（患者群）を同定すること；
および

e) 患者のサンプルがDESNT癌を伴うプロセス（患者群）中かどうか決定することにより、癌を分類するか、患者の中のDESNT癌の存在又は非存在を決定することを含む。

【0061】

既に記載されたように、本発明の方法のうちのいくつかは、多数の患者サンプルに対して同時に実行することができる。例えば、サンプル中の複数の遺伝子の発現のレベルは、少なくとも2人の異なる患者から得られた少なくとも2つのサンプルの中で決定され、各患者についての発現プロフィールを提供する。

10

【0062】

本発明の方法はさらに、1つ以上の患者の発現プロフィールにユニークなラベルを割り当てることを含み、分析ステップの間により容易にそれらを同定することができるようにされる。

【0063】

本発明の方法で、DESNT癌に関係するプロセス/患者群の同定を、上記の最初の3つの方法のうちの1つの使用により行うことができる。具体的には、(i) 別のデータセットのDESNT癌グループと遺伝子発現レベルの相関；(ii) コアダウンレギュレート遺伝子セットを有するDESNTおよび非DESNT癌の間の差別的に発現した遺伝子のオーバーラップのデモンストレーション；

20

(iii) より不良な臨床の結果；

【0064】

DESNTとしての個々の癌の割付は、方法(iv)の患者のデータセットが既知のDESNTグループか否かを決定するために、結合した参照データセットと患者のデータセットにLPDを実行することにより達成される。方法(iii)は、既知とされる参照データセット中の患者に臨床結果を要求する。

【0065】

「臨床結果」は、参照データセット中の各患者について、癌が進行したか否かを意味する。例えば、初期評価の一部として、患者は前立腺特異性抗原(PSA)レベルをモニターされることがある。それが特定のレベルを上回る場合、これは再発、そしてその結果として疾患の進行を示す。病理組織学的診断も使用されることがある。リンパ節への広がり、転移にも使用される。さらに癌による患者の死亡(あるいは単に一般に患者の死亡)は臨床の終了である。グリーソン・スコアリング(Gleason scoring)、癌進度、および多数の生検(サンプルを得るために中空針を含むコアリング法(coring method))を使用して得られたもののような)が使用されることがある。臨床結果は前立腺癌の治療の後に評価されることがある。これは長期での患者の身に起こることである。通常、患者は、有効に前立腺を切除するか殺すために根本的に(前立腺摘除、放射線療法)治療されるだろう。再発あるいは引き続くPSAレベルの上昇(PSA再発として知られている)の存在は、進行した癌を示す。本発明の方法を使用して同定されたDESNT癌群は、より速く進行する癌の部分母集団を含む。そのような方法(i)、(ii)、(iii)および(iv)の組み合わせが使用されることがある。また、当業者は、参照データセット中の患者についての患者の結果を決定する方法に精通している。

30

40

【0066】

従って、発明の方法のすべては、DESNT癌が疑われる患者の中で実行されることができる。重要なことには、癌の治療が提供される前に、本発明は、癌進行を予測することを許容する。癌が治療なしでも進行しない時にも、多くの患者が前立腺癌の不必要な治療を受けるので、これは前立腺癌にとって特に重要である。

【0067】

50

さらに、同じデータが分析されている場合さえ、分析が実行されるごとに、LPDの結果がわずかに異なることがあるので、診断の正確さは分析の繰り返しにより増加される場合がある。多くの場合、その変異は、LPDプロセスの各実行の中で使用されている乱数発生機（シード値として使用される）の異なる出発点による。また、したがって、同じデータセット上の繰り返しされた実行についてさえ、多数の異なる結果が発生する場合がある。したがって、分析を複数の回行ない、モード値（最頻値）、または平均値を参照することが有益である。ある実施態様では、LPD分析は少なくとも2、3、5あるいは少なくとも20回行なわれる。ある実施態様では、分析は少なくとも50回行なわれる。好ましい実施態様では、分析は少なくとも100回（つまり、少なくとも99回繰り返される）行なわれる。

10

【0068】

分析ステップが繰り返される実施態様では、癌がDESENTかどうか決定するステップは、癌がDESENTとして示される数と、癌が非DESENT（つまりインドレントか、非侵襲性の前立腺癌）として示される数の比較を要求することがある。例えば、患者が侵襲性の癌にかかっているという決定は、試みられた分析ステップの少なくとも50%で癌がDESENTであることを要求することがある。好ましい実施態様では、癌は少なくとも60%、あるいはもっと好ましい実施態様では少なくとも70%でDESENTでなければならない。最も好ましい実施態様では、癌は分析の少なくとも67%でDESENTである。

【0069】

LPD分析が試みられる場合、それは多くのプロセス（グループ）へ分析されるデータセット中の患者を分割する。発明のある実施態様では、特定の患者（その臨床の結果は知られていない）がDESENT癌にかかっているかどうか決定するステップは、プロセス（例えば侵襲性の癌に関連した患者群）が知られていることを要求する。その後、参照データセットに加えられた患者のサンプルは、侵襲性の癌（DESENT）グループに存在するか、あるいは存在しないことができ、患者が侵襲性の癌にかかっているかどうか決定される。

20

【0070】

しかしながら、上に記載されたように、参照データセット中の患者の臨床の結果を前もって知ることは必ずしも必要とは限らない。DESENT癌の存在を決定するこれらの2つの方法のどちらかあるいは両方を使用することができる：

30

- (i) 他のデータセット中のDESENT癌グループと遺伝子発現レベルとの相関；または
- (ii) コアをダウンレギュレートされた遺伝子セットのDESENTおよび非DESENT癌の間の差別的に発現した遺伝子のオーバーラップのデモンストレーション。

【0071】

DESENTとしての個々の癌の割付は、結合した参照データセットと患者のデータセットにLPDを実行し、患者のデータセットクラスターが既知のDESENTグループのクラスターに分けられるかどうか決定することにより達成することができる。

【0072】

理想的には、参照データセット中のDESENT癌の存在または不存在は、これらの方法の3つまでを使用して決定される：

40

- i. 別の1または複数のデータセット中のDESENT癌グループの遺伝子発現レベルの相関、
- ii. コアダウンレギュレートされた遺伝子セットのDESENT癌および非DESENT癌の間の差別的に発現した遺伝子のオーバーラップのデモンストレーション、および
- iii. 臨床の結果との相関。

【0073】

患者のサンプル中の複数の遺伝子の発現のレベルを決定するステップは、ここに議論されたもののような当業者に公知の方法、たとえばProkopec SD, Watson JD, Waggott DM, Smith AB, Wu AH, Okey AB et al. Systematic evaluation of medium-throughput mRNA abundance platforms. RNA 2013; 19: 51-62; Chatterjee A, Leichter

50

AL, Fan V, Tsai P, Purcell RV, Sullivan MJ et al. A cross comparison of technologies for the detection of microRNAs in clinical FFPE samples of hepatoblastoma patients. Sci Rep 2015; 5: 10438; Pollock JD. Gene expression profiling: methodological challenges, results, and prospects for addiction research. Chem Phys Lipids 2002; 121: 241-256; Mantione KJ, Kream RM, Kuzelova H, Ptacek R, Raboch J, Samuel JM et al. Comparing bioinformatic gene expression profiling methods: microarray and RNA-Seq. Med Sci Monit Basic Res 2014; 20: 138-142; Casassola A, Brammer SP, Chaves MS, Ant J. Gene expression: A review on methods for the study of defense-related gene differential expression in plants. American Journal of Plant Research 2013; 4, 64-73; Oszlak F, Milos PM. RNA sequencing: advances, challenges and opportunities. Nat Rev Genet 2011; 12: 87-98. A cross comparison of technologies for the detection of microRNAs in clinical FFPE samples of hepatoblastoma patients. Sci Rep 2015; 5: 10438; Pollock JD. Gene expression profiling: methodological challenges, results, and prospects for addiction research. Chem Phys Lipids 2002; 121: 241-256; Mantione KJ, Kream RM, Kuzelova H, Ptacek R, Raboch J, Samuel JM et al. Comparing bioinformatic gene expression profiling methods: microarray and RNA-Seq. Med Sci Monit Basic Res 2014; 20: 138-142; Casassola A, Brammer SP, Chaves MS, Ant J. Gene expression: A review on methods for the study of defense-related gene differential expression in plants. American Journal of Plant Research 2013; 4, 64-73; Oszlak F, Milos PM. RNA sequencing: advances, challenges and opportunities. Nat Rev Genet 2011; 12: 87-98、によることができる。

【 0 0 7 4 】

発明の実施態様では、すべての方法の分析ステップはコンピューターで実行できる。発明は、さらに本発明の方法の任意の方法を実行するようにプログラムされたコンピューター読み取り可能なメディアを提供する。

【 0 0 7 5 】

発明の一層の実施態様では、前立腺癌を分類する方法、例えば患者の中の侵襲性の前立腺癌を診断するか、または患者の前立腺癌について予後不良（つまり D E S N T 前立腺癌の患者）と同定する方法であって、以下を含む方法が提供される。

- a) データセット中の個々の患者のサンプルの D E S N T ステータスが知られている（例えば L P D 分析によって決定される）参照データセットを提供すること；
- b) このデータセットから複数の遺伝子を選択すること、ここで該複数の遺伝子は、表 2 にリストされたグループから選ばれた少なくとも 5、少なくとも 10、少なくとも 20、少なくとも 30、少なくとも 40、あるいは少なくとも 45 の遺伝子、あるいは表 3 にリストされたグループから選ばれた少なくとも 5、少なくとも 10、少なくとも 15、あるいは少なくとも 20 の遺伝子を含む；
- c) 任意に：
 - (i) コントロールとして、少なくとも 1 つの患者のサンプル中の遺伝子の発現ステータスをさらに決定すること、ここで該コントロール遺伝子は表 2 あるいは表 3 にリストされた遺伝子でない；
 - (i i) 複数の遺伝子およびコントロール遺伝子の発現の相対的なレベルを決定すること；および
- d) 選択された遺伝子の発現ステータスを使用して、データセットにスーパーバイズされた機械学習アルゴリズム（例えばランダムフォレスト分析）を適用し、D E S N T 癌についての予測を得ること；
- e) 患者から得られたサンプル中の同じ複数の遺伝子の発現ステータスを決定し、患者の発現プロファイルを提供すること；
- f) 任意に、参照データセットへ患者の発現プロファイルを正規化すること；および

10

20

30

40

50

g) 患者の発現プロフィールへ予測を適用して癌を分類し、侵襲性の癌の存在を決定するか、あるいは患者は予後不良かどうか決定すること（つまり、患者の癌がDESENTあるいは非DESENTかどうか決定すること）。

【0076】

この方法およびその変種は今後方法2と呼ばれる。表2の遺伝子は、多数のデータセットについてLPD分析を行ないDESENTグループの中で共通してダウンレギュレートされる遺伝子を決定することにより、発明者によって同定された。表3の遺伝子は、方法4に述べられているようなLASSO分析を行なうことにより発明者によって同定された。

【0077】

好ましい実施態様では、ステップ(i)の中で使用されるコントロール遺伝子は、表6あるいは表7にリストされたハウスキーピング遺伝子から選ばれる。

表7は、前立腺癌に特に関連する。発明のある実施態様では、少なくとも1、少なくとも2、少なくとも5あるいは少なくとも10のハウスキーピング遺伝子を使用する。好ましい実施態様では少なくとも2つのハウスキーピング遺伝子を使用する。上記のステップ(ii)はテスト遺伝子とハウスキーピング遺伝子の間の比率を決定することを含むことがある。

【0078】

発明の一層の方法では、患者の中の侵襲性の癌（たとえば侵襲性の前立腺癌）を診断するか、または患者を予後不良の癌（つまりDESENT癌の患者）と同定する、以下を含む方法が提供される：

a) データセット中の個々の患者のサンプルのDESENTステータスが知られている（例えばLPD分析によって決定される）参照データセットを提供すること；

b) このデータセットから複数の遺伝子を選ぶこと；

c) これらの選択された遺伝子の発現ステータスを使用して、データセットに、スーパーバイズされた機械学習アルゴリズム（例えばランダムフォレスト分析）を適用し、DESENT癌のための予測を得ること；

d) 患者から得られたサンプル中の同じ複数の遺伝子の発現ステータスを決定し、患者の発現プロフィールを提供すること；

e) 任意に、参照データセットへ患者の発現プロフィールを正規化すること；および

f) 患者の発現プロフィールに予測を適用し、患者の癌がDESENTあるいは非DESENTかどうか決定すること。

【0079】

この方法およびその変種は今後方法3と呼ばれる。発明のさらなる方法では、患者の中の侵襲性の癌（たとえば侵襲性の前立腺癌）を診断するか、または患者を予後不良の癌（つまりDESENT癌の患者）と同定する、以下を含む方法が提供される：

a) データセット中の個々の患者のサンプルのDESENTステータスが知られている（例えばLPD分析によって決定される）1又は複数の参照データセットを提供すること；

b) このデータセットから、その発現ステータスがDESENTおよび非DESENT癌の間で変化することが知られている複数の遺伝子を選ぶこと（例えば表4表にリストされた複数の遺伝子、たとえば表4にリストされた少なくとも100、少なくとも200、少なくとも300、少なくとも400、少なくとも500、あるいは少なくとも1000の遺伝子）；

c) 選択された遺伝子にLASSOロジスティック回帰モデル分析を適用し、DESENT癌を[最良に]同定する選択された遺伝子の部分集合を特定する；

d) 選択された遺伝子の部分集合の発現ステータスを使用して、データセットにスーパーバイズされた機械学習アルゴリズム（例えばランダムフォレスト分析）を適用し、DESENT癌のための予測を得ること；

e) 患者から得られたサンプル中の選択された遺伝子の部分集合の発現ステータスを決定し、患者の発現プロフィールを提供すること；

f) 参照データセットへ患者の発現プロフィールを正規化すること；および

10

20

30

40

50

g) 患者の発現プロフィールに最適な予測を適用し、患者の癌がDESNTあるいは非DESNTかどうか決定すること。

【0080】

この方法およびその変種は今後方法4と呼ばれる。

【0081】

ランダムフォレスト分析を含む方法を利用して同定されたDESNT患者集団は、「RF-DESNT」と呼ばれる。LPD分析を含む方法を利用して同定されたDESNT患者集団は、「LPD-DESNT」と呼ばれる。

【0082】

本発明の発明者は、LPDと異なり、コンピューター処理集約的でなく、広範囲のデータセットに適用可能なクラシファイヤーを開発することを望んでいる。

方法2から4はそのような解決策を提供し、癌進行を予測するために使用することができる。したがって、本発明は、RF-DESNT癌を同定し、かつ癌進行を予測するためにランダムフォレスト分類の中で使用することができる遺伝子シグネチャーを同定する方法を提供する。

【0083】

スーパーバイズされた機械学習アルゴリズムあるいは一般的な線形モデルが、DESNTステータスの予測を生成するために使用される。好ましいアプローチはランダムフォレスト分析であるが、支援ベクター機械(support vector machines)、ニューラル・ネットワーク(neural networks)、ネイティブベイズクラシファイヤー(naive Bayes classifier)あるいはニアレストネイバーアルゴリズム(nearest neighbour algorithms)のような選択肢を使用することができる。そのような方法は当業者によって知られており理解されている。

【0084】

ランダムフォレスト分析は癌がDESNTか否かを予測するために使用することができる。上記の方法2から4は、方法1より相当に少ない計算能力を要求し、従ってより容易に実行することができる。

【0085】

ランダムフォレスト分析は分類、回帰および他のタスクのための全体学習法であり、トレーニング中にディシジョンツリーの強度を構成することにより作動し、クラスのモードであるクラスを出力し(分類)、個々のディシジョンツリーの平均の予測(回帰)を出力する。従って、ランダムフォレストはデータの過剰適合を修正し、任意の1つのディシジョンツリーにする。

【0086】

ディシジョンツリーは、木のようなグラフが決定のモデルを含み、チャンスイベント結果(chance event outcomes)を含むそれらの可能な結論を含む。ディシジョンツリーのそれぞれの中間ノードは、典型的には1または複数の属性のテスト(たとえば、癌サンプル中の遺伝子の発現レベルが前もって定義した閾値より上にあるか否か)を表し、ディシジョンツリーのそれぞれの枝は、典型的にはテストの結果を表わす。また、ディシジョンツリーのそれぞれのリーフノードはそれぞれ典型的にはクラス(分類)ラベルを表わす。

【0087】

ランダムフォレスト分析では、全体クラシファイヤーは、典型的にはトレーニング・データセット(参照データセットとも呼ばれる)でトレーニングされ、ここで、たとえばLPDによって決定されるように、DESNTグループか否かが知られている。トレーニングは、DESNTあるいは非DESNTのメンバーシップについての予測であるモデルを生成する。RFによって同定されたグループはRF-DESNTおよびRF-非-DESNTと呼ぶことができる。一旦トレーニングされたら、ランダムフォレストクラシファイヤーは、未知のサンプルからのデータセットに適用することができる。このステップは決定論的である、つまり、クラシファイヤーが同じデータセットに続いて繰り返し適用されれば、それは一貫して毎回同じクラスの中へ新しいデータセットの各癌をソートするだろ

10

20

30

40

50

う。

【 0 0 8 8 】

全体クラシファイヤーは、R F - D E S N T 癌あるいは R F - 非 - D E S N T 癌のいずれかとして新しいデータセット中の各癌サンプルを分類するように作用する。従って、ランダムフォレスト分析が試みられる場合、全体クラシファイヤーは多くのクラスへ分析されているデータセット中の癌を分割する。クラスの数、2 であることができる。つまり、全体の分類は、D E S N T 癌を有する D E S N T クラスまたは D E S N T グループ、および他の癌を有する非 D E S N T クラスまたは非 D E S N T グループに、データセット中の患者を分類することができる。

【 0 0 8 9 】

ランダムフォレストのディシジョンツリーはそれぞれ、与えられた癌サンプルについて、それが認識するように訓練されたクラスのうちの 1 つ、つまり D E S N T / 非 D E S N T にそれを帰する独立予測因子である。各ディシジョンツリーのノードはそれぞれ、患者から得られた癌サンプル中の同じ複数遺伝子の 1 つ以上の遺伝子に関するテストを含む。いくつかの遺伝子はノードでテストされることができる。例えば、テストは、複数の遺伝子の 1 つ以上の遺伝子の発現レベルが前もって定義した閾値より上にあるかどうか尋ねることがある。

【 0 0 9 0 】

ディシジョンツリー間の変異は、異なる方法でクラスにサンプルを分けるように各ディシジョンツリーに導くだろう。全体の分類は、すべての独立したディシジョンツリーによって生成された分類を行い、最多のディシジョンツリーが導くクラスにサンプルを分ける。

【 0 0 9 1 】

方法 2、3 あるいは 4 (ランダムフォレスト分析の決定に基づく) のステップ b) の中で発現のレベルが決定される複数の遺伝子は、任意の適切な方法を使用して選ぶことができる。可能な 1 つの方法は、トレーニング・データセットに L P D 分析あるいは他のベイズの統計分析を適用し、D E S N T グループ / プロセスに割り当てられる癌を決定することである。その後、非 D E S N T 癌と比較して、D E S N T 癌で一貫してダウンレギュレートされる遺伝子を選択する。このダウンレギュレーションは、L P D 分析が行なわれたいくつかの異なるデータセットにわたり一貫していることがある。ある実施態様では、方法 2、3 および 4 のステップ b) の中で使用される複数の遺伝子は、少なくとも 5、少なくとも 10、少なくとも 15、少なくとも 20、少なくとも 30、少なくとも 40 あるいは少なくとも 45 の遺伝子を含む。特に、方法 2 および方法 3 のステップ b) の中で使用される複数の遺伝子は、表 2 にリストされた遺伝子の少なくとも 5、少なくとも 10、少なくとも 15、少なくとも 20、少なくとも 30、少なくとも 40 あるいは 45 の遺伝子すべてを含む。

【 0 0 9 2 】

別の可能な方法 (方法 4 のような) では、ア リースト アブソリュート シュリンケージ アンド セレクション オペレーター (a least absolute shrinkage and selection operator; L A S S O) 分析を行ない、かつ最も良く D E S N T メンバーシップを特徴づける遺伝子を選択する。モデル係数の絶対値の合計がある閾値未満であるように、ロジスティック回帰モデルは係数の制約を受けて実行される。これは、D E S N T メンバーシップを予測する能力を持っていないか、既にモデル内で遺伝子の発現と関連づけられている遺伝子を除去する効果がある。L A S S O は、トレーニングまたは参照データセット中の D E S N T と非 D E S N T サンプルを同定する可能性が最もありそうな遺伝子を見つける数学的な方法である。遺伝子の部分集合は方法 4 のステップ e) であり、少なくとも 5、少なくとも 10、少なくとも 15、あるいは少なくとも 20 の遺伝子を含むことがある。特別の実施態様では、方法 4 のステップ a) および b) は必要ではない。また、代わりに、方法は、表 3 に示された遺伝子の部分集合の、少なくとも 5、少なくとも 10、少なくとも 15 あるいは少なくとも 20 を使用して、方法 4 のステップ c) から始めることができる。

10

20

30

40

50

【 0 0 9 3 】

方法 4 を実施する時、発明者は次のステップを行う。出発点として、発明者によって分析された 5 つ（合計の遺伝子）のデータセットの少なくとも 2 つにおいて一般に存在した D E S N T 癌の中の、（非 D E S N T 癌と比較して）著しく変化する発現を示す遺伝子はすべて同定された。L A S S O を実行するための代表的な D E S N T L P D 分類は、分析された M S K C C データセットから選ばれた。実際には、使用される D E S N T 分類は、たとえばいくつかの統計的検定の平均の p 値、あるいはいくつかの記述のすべての合計から選択される代表的な方法を使用することができ、たとえば、サンプルの少なくとも 5 0 % が D E S N T ステータスに該当する場合、D E S N T として分類される。

【 0 0 9 4 】

L A S S O ロジスティック回帰モデルは、M S K C C データセットの内の D E S N T メンバーシップを予測し、D E S N T メンバーシップを特徴づける 2 0 のセットの遺伝子の選択に結びつけるために使用された。これらの遺伝子は表 3 にリストされる。1 6 6 9 の遺伝子からのこれらの 2 0 の遺伝子の除去およびこの手続きの反復は、D E S N T メンバーシップを特徴づけることができる遺伝子のさらなるセットを同定した。出発点として L P D によって分析された他のデータセットを使用して、同じ分析を行なうことにより、遺伝子の追加のセットを得ることができる。

【 0 0 9 5 】

本発明は、D E S N T 癌または癌進行と関連するか、それを予測する遺伝子の一層のリストを提供する。例えば、1 つの実施態様では、L A S S O 分析が D E S N T 癌、特に D E S N T 前立腺癌を示すかまたは予測する発現シグネチャーを提供するために使用することができる。発現シグネチャーもバイオマーカー・パネルと考えられることができ、表 3 にリストされた遺伝子から成るグループから選ばれる少なくとも 5、少なくとも 1 0、少なくとも 1 2、少なくとも 1 5 あるいはすべての 2 0 の遺伝子を含む。

【 0 0 9 6 】

発明の任意の方法では、統計分析を分析される遺伝子の発現のレベルについて行なうことができるか、または統計分析を、遺伝子発現および任意のコントロール遺伝子の発現の相対的なレベルにしたがって計算された比率に行なうことができることに留意すべきである。

【 0 0 9 7 】

例えば、方法 1 に関して、方法は、次のステップを含むことができる：

- a) 患者から得られたサンプル中の複数の遺伝子の発現のレベルを決定し、患者の発現プロファイルを提供すること；
- b) コントロールとして患者のサンプル中の、少なくとも 1 つの異なる遺伝子の発現ステータスをさらに決定すること、ここでコントロール遺伝子は発現のレベルがステップ a) で決定された遺伝子ではない；
- c) 複数の遺伝子および 1 または複数のコントロール遺伝子の発現の相対的なレベルを決定すること；
- d) 異なる患者からの発現プロファイルを含む参照データセットを提供し、ステップ a) で使用されたものと同じ複数の遺伝子、およびステップ b) で使用したものと同じ 1 または複数のコントロール遺伝子の発現の相対的なレベルを決定すること；
- e) 統計的ベイズクラスタリング分析あるいは他のクラスタリング分析を、ステップ c) および d) で得られた相対的な発現レベルについて行うこと；
- f) 任意に、複数回分析ステップ e) を繰り返すこと；および
- g) 癌を分類し、患者が癌にかかっているかどうか、あるいは患者が予後不良（つまり、患者は D E S N T 癌にかかっている）かを決定すること。

【 0 0 9 8 】

方法 2 に関して、方法は、以下のステップを含むことができる：

- a) データセット中の個々の患者のサンプルの D E S N T ステータスが知られている（例えば L P D 分析によって決定されたように）、参照データセットを提供すること；

b) このデータセットから複数の遺伝子を選択すること、ここで複数の遺伝子が、表 2 にリストされたグループから選ばれた少なくとも 5、少なくとも 10、少なくとも 20、少なくとも 30、少なくとも 40 あるいは少なくとも 45 の遺伝子を含むか、あるいは表 3 にリストされたグループから選ばれた少なくとも 5、少なくとも 10、少なくとも 15 あるいは少なくとも 20 の遺伝子を含む；

c) コントロールとして少なくとも 1 つのさらに異なる患者のサンプル中の遺伝子の発現ステータスを決定すること；

d) 複数の遺伝子および 1 または複数のコントロール遺伝子の発現の相対的なレベルを決定すること；

e) 発現の相対的なレベルを使用し、スーパーバイズされた機械学習アルゴリズム（例えばランダムフォレスト分析）を参照データセット適用し、D E S N T 癌についての予測を得ること；

10

f) 患者から得られたサンプル中の同じ複数の遺伝子およびコントロール遺伝子の発現の相対的なレベルを決定し、患者の発現プロフィールを提供すること；

g) 任意に参照データセットへ患者の発現プロフィールを正規化すること；および

h) 患者の発現プロフィールへ予測を適用し、癌を分類し、侵襲性の癌の存在、または患者が予後不良（つまり、患者の癌が D E S N T あるいは非 D E S N T かどうか決定する）かどうかを決定すること。

【0099】

方法 3 に関して、方法は、以下のステップを含むことができる：

20

a) データセット中の個々の患者のサンプルの D E S N T ステータス（つまり癌分類）が知られている（例えば L P D 分析によって決定されている）参照データセットを提供すること；

b) 複数の遺伝子をこのデータセットから選択すること；

c) コントロールとして少なくとも 1 つさらに異なる患者のサンプル中の遺伝子の発現ステータスを決定すること；

d) 複数の遺伝子および 1 または複数のコントロール遺伝子の発現の相対的なレベルを決定すること；

e) これらの選択された遺伝子の発現の相対的なレベルを使用し、スーパーバイズされた機械学習アルゴリズム（例えばランダムフォレスト分析）をデータセットに適用し、D E S N T 癌についての予測を得ること；

30

f) 患者から得られたサンプル中の発現の相対的なレベルを含む患者の発現プロフィールを提供すること、ここで発現の相対的なレベルはステップ b) で選択されたものと同じ複数の遺伝子、およびステップ d) で使用されたものと同じコントロール遺伝子を使用して得られる；

g) 任意に参照データセットへ患者の発現プロフィールを正規化すること；および

h) 患者の発現プロフィールへ予測を適用し、癌を分類し、侵襲性の癌の存在、または患者が予後不良（つまり、患者の癌が D E S N T あるいは非 D E S N T かどうか決定する）かどうかを決定すること。

【0100】

40

方法 4 に関して、方法は、以下のステップを含むことができる：

a) データセット中の個々の患者のサンプルの D E S N T ステータス（つまり癌分類）が知られている（例えば L P D 分析によって決定されている）1 以上の参照データセットを提供すること；

b) このデータセットから、その発現ステータスが D E S N T および非 D E S N T 癌の間で変わることが知られている複数の遺伝子を選ぶこと（例えば表 4 にリストされた複数の遺伝子、例えば表 4 にリストされた遺伝子の少なくとも 100、少なくとも 200、少なくとも 300、少なくとも 400、少なくとも 500 あるいは少なくとも 1000 の遺伝子）；

c) 選択された遺伝子上に L A S S O ロジスティック回帰モデル分析を適用し、D E S N

50

T 癌を同定する選択された遺伝子の部分集合を同定すること；

d) コントロールとして少なくとも1つさらに異なる患者のサンプル中の遺伝子の発現ステータスを決定すること；

e) 遺伝子、および1または複数のコントロール遺伝子の部分集合の発現の相対的なレベルを決定すること；

f) 相対的な発現レベルを使用し、スーパーバイズされた機械学習アルゴリズム（例えばランダムフォレスト分析）をデータセットに適用し、D E S N T 癌についての予測を得ること；

g) 患者から得られたサンプル中の発現の相対的なレベルを含む患者の発現プロフィールを提供すること、ここで発現の相対的なレベルはステップc) で選択されたものと遺伝子の同じ部分集合、およびステップe) で使用されたものと同じコントロール遺伝子を使用して得られる；

h) 任意に参照データセットへ患者の発現プロフィールを正規化すること；および

i) 患者の発現プロフィールへ予測を適用し癌を分類し、侵襲性の癌の存在を決定するか、あるいは患者は予後不良を決定すること、（つまり、患者の癌がD E S N T あるいは非D E S N T かどうか決定する）。

【0101】

上記の方法のうちのどれでも、1または複数のコントロール遺伝子は表6あるいは表7にリストされた遺伝子から選ばれることができる。

【0102】

データセット

本発明者は分析において、M S K C C、C a n c e r M a p、S t e p h e n s o n、C a m C a p およびT C G A データセットを使用した。しかしながら、他の適切なデータセットも利用可能である。一般に、データセットは、患者または腫瘍のサンプルからの複数の発現プロフィールを含む。データセットのサイズは変わることができる。例えば、データセットは、患者または腫瘍サンプルからの少なくとも20、任意に少なくとも50、少なくとも100、少なくとも200、少なくとも300、少なくとも400あるいは少なくとも500の発現プロフィールを含むことがある。好ましくは、データセットは、少なくとも500の患者あるいは腫瘍からの発現プロフィールを含む。

【0103】

ある実施態様では、発明の方法は、複数のデータセットからの発現プロフィールを使用する。例えば、ある実施態様では、方法は、少なくとも2つのデータセットからの発現プロフィールを使用し、それぞれのデータセットは、少なくとも250の患者あるいは腫瘍からの発現プロフィールを含む。

【0104】

患者または腫瘍からの発現プロフィールは、遺伝子の部分集合の発現のレベルについての情報を含み、たとえば少なくとも10、少なくとも40、少なくとも100、少なくとも500、少なくとも1000、少なくとも1500、少なくとも2000、少なくとも5000あるいは少なくとも10000遺伝子についての情報を含む。好ましくは、患者の発現プロフィールは、少なくとも500の遺伝子の発現データを含む。発明の様々な方法の分析ステップで、遺伝子の部分集合のどんな選択もデータセットの中にある遺伝子から得られるだろう。

【0105】

癌の分類

ここに示された方法とバイオマーカーは、進行の可能性によって癌を分類するのに役立つ（従って癌の予後に役立つ）。本発明は、特に前立腺癌に焦点が当てられるが、方法は他の癌にも使用することができる。特に、表2の中の遺伝子のリストは、例えば、前立腺癌を含む一連の癌の進行を示すことが分かった。進行しそうかまたは進行する癌は発明者によってD E S N T 癌と呼ばれる。D E S N T 癌は、進行が予測される癌を指す。D E S N T ステータスは、癌が進行するかどうかの指標を指す。侵襲性の癌は進行する癌である。

10

20

30

40

50

【 0 1 0 6 】

本発明の方法により分類することができる癌のタイプとしては以下を包含する；急性リンパ芽球性白血病、急性または慢性のリンパ球あるいは顆粒白血球の腫瘍、急性骨髄性白血病、急性前骨髄球性白血病、腺癌、腺腫、副腎癌、基底細胞癌、骨癌、脳腫瘍、乳癌、気管支癌、子宮頸部形成異常、慢性骨髄性白血病、結腸癌、類表皮癌、ユーイング肉腫、胆嚢癌、胆石腫物、巨細胞腫、膠芽腫多重形、毛様細胞腫瘍、頭部癌、過形成、過形成の角膜神経腫瘍、イン・シトゥー癌腫、腸の神経節腫、島細胞腫、カボジ肉腫、腎臓癌、喉頭癌、平滑筋腫瘍、肝臓癌、肺癌、リンパ腫、悪性カルチノイド、悪性高カルシウム血症、悪性黒色腫、マルファン症候群様の体質腫瘍、髄様癌、転移性の皮膚癌、粘膜神経腫、菌状息肉腫、骨髄異形成症候群、骨髄腫、頸部癌、神経の組織癌、神経芽腫、骨原性肉腫、骨肉腫、卵巣腫瘍、脾癌、副甲状腺癌、クロム親和細胞腫、真性赤血球増加症、原発性の脳腫瘍、前立腺癌、直腸癌、腎の細胞腫瘍、網芽細胞腫、横紋筋肉腫、セミノーマ、皮膚癌、小細胞肺腫瘍、軟部組織肉腫、扁平上皮癌、胃癌、甲状腺癌、局所的の皮膚病変、細網細胞肉腫(vetriculum cell sarcoma)、あるいはウィルムス腫瘍(Wilm's tumor)。

10

【 0 1 0 7 】

本発明が特に関連するものには、前立腺癌、結腸直腸癌および乳癌がある。

【 0 1 0 8 】

「侵襲性の癌」は、「侵襲性の前立腺癌」を含む。侵襲性の前立腺癌は、疾患進行および潜在的なさらなる合併症（たとえば転移あるいは転移性の進行）を防ぐか、停止させるか、低減するために治療を要求する癌として定義することができる。侵襲性の前立腺癌は、治療していないでおかれたならば、前立腺の外部に広がり、患者を殺すことがある前立腺癌である。本発明は、侵襲性の前立腺癌を含むいくつかの侵襲性の癌を検知するのに役立つ。

20

【 0 1 0 9 】

前立腺癌は、アメリカン ジョイント コミッティー オン キャンサー(American Joint Committee on Cancer:AJCC)の腫瘍ノード転移(tumour-nodes-metastasis:TNM)進捗システムにしたがって分類することができる。Tスコアは、主要な(原発性)腫瘍のサイズ、および前立腺の外部で成長しているか、および近くの器官内に存在しているかについて記述する。Nスコアは近くの(局所的な)リンパ節に広がっているか記述する。Mスコアは、癌が身体の他の器官に転移(拡散)したかどうかを示す：

30

【 0 1 1 0 】

T 1 腫瘍は小さすぎるので、スキャン上で見ることができず、前立腺の診察時には自覚症状はない。高いP S Aレベルを検出した後、それらは針生検によって発見されることがある。

【 0 1 1 1 】

T 2 腫瘍は完全に前立腺の内部にあり、3つのより小さなグループに分割される：

T 2 a - 腫瘍は、前立腺の葉のうちの1つの半分だけにある；

T 2 b - 腫瘍は葉のうちの1つの半分以上にある；

T 2 c - 腫瘍は両方の葉にあるが、前立腺の内部にまだある。

【 0 1 1 2 】

T 3 腫瘍は前立腺のカプセル(カバリング)を通り抜けている - それらは2つのより小さなグループに分類される：

T 3 a - 腫瘍は前立腺のカプセル(カバリング)を通り抜けている；

T 3 b - 腫瘍は精嚢へ広がっている。

【 0 1 1 3 】

T 4 腫瘍は、たとえば直腸、膀胱、筋肉あるいは骨盤腔の側のような他の近くの体器官へ広がっている。病期T 3およびT 4腫瘍は局所的に進行した前立腺癌と呼ばれる。

40

【 0 1 1 4 】

リンパ節はそれらが癌細胞を含んでいる場合、「陽性」と評される。リンパ節の内部の癌細胞がある場合、それは通常より大きい。より多くの癌細胞を含むほど、リンパ節は大

50

きくなるだろう：

- N X - リンパ節はチェックすることができない；
- N 0 - 前立腺に近いリンパ節に癌細胞はない；
- N 1 - リンパ節の中に癌細胞がある。

【 0 1 1 5 】

M 進度は転移（癌拡散）を指す：

- M 0 - 癌は骨盤の外部に広がっていない；
- M 1 - 癌は骨盤の外部に広がっている；
- M 1 a - 骨盤の外側のリンパ節に癌細胞がある；
- M 1 b - 硬骨に癌細胞がある；
- M 1 c - 他の場所に癌細胞がある。

10

【 0 1 1 6 】

前立腺癌もグリーソン分類(Gleason grading system)を使用して、スコア付けされることができる。それは、疾病の悪化を類別するために組織学的分析を使用する。1 から 5 のグレードが診察される細胞に割り当てられる。また、2 つの最も一般的な等級が全体的なグリーソンスコアを提供するために加えられる。グレード 1 は、緊密にパックされた形の良い腺を含み健康な組織に非常によく似ている。グレード 5 は認識しうる腺がないか、あるいは非常に少ない。6 未満のスコアでは予後良好である。しかし、6 以上のスコアではより侵襲性のものとして分類される。グリーソンスコアは泌尿器病理学の国際学会によって 2 0 0 5 年に修正された。このスコアリングの基準は以下に記述される(Epstein JI, Allsbrook WC Jr, Amin MB, Egevad LL; ISUP Grading Committee. The 2005 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason grading of prostatic carcinoma. Am J Surg Pathol 2005;29(9):1228-42)。グリーソンスコアは生検で検知される、つまり、サンプリングされた腫瘍の部分で検知される。グリーソン 6 の前立腺は、生検によってサンプリングされない侵襲性の腫瘍の小さな巣(foci)を持っていることがある。したがって、グリーソンはガイドである。より低いグリーソンスコアでは、より小さな割合の患者が侵襲性の癌を持つだろう。前立腺癌の患者の中のグリーソンスコアは 2 から 1 0 までに達することができる。グリーソンの小部分は侵襲性の癌があるので、平均生存率は高い。また、グリーソン増加するとともに平均生存率は減少する。侵襲性の癌を有する患者によって低下される（つまり、各グリーソンスコアでの生存率の混合がある）からである。

20

30

【 0 1 1 7 】

前立腺癌もそれらがどれくらい進んだかによってステージを分けることができる。これは、たとえばグリーソンスコアおよび/または P S A 検査のような他の要因と同様に T M N スコアリングに基づく。進度は以下のように定義することができる：

【 0 1 1 8 】

第 1 期：

- T 1、N 0、M 0、グリーソンスコア 6 以下、1 0 未満の P S A；または
- T 2 a、N 0、M 0、グリーソンスコア 6 以下、1 0 未満の P S A。

【 0 1 1 9 】

病期 I I A：

- T 1、N 0、M 0、7 のグリーソンスコア、2 0 未満の P S A；または
- T 1、N 0、M 0、6 以下のグリーソンスコア、少なくとも 1 0 だが 2 0 未満の P S A；
- または、
- T 2 a あるいは T 2 b、N 0、M 0、7 以下のグリーソンスコア、2 0 未満の P S A。

40

【 0 1 2 0 】

病期 I I B：

- T 2 c、N 0、M 0、任意のグリーソンスコア、任意の P S A；または
- T 1 または T 2、N 0、M 0、任意のグリーソンスコア、2 0 以上の P S A；または
- T 1 あるいは T 2、N 0、M 0、8 以上のグリーソンスコア、任意の P S A。

50

【 0 1 2 1 】

第 3 期：

T 3、N 0、M 0、任意のグリーソンスコア、任意の P S A。

【 0 1 2 2 】

第 4 期：

T 4、N 0、M 0、任意のグリーソンスコア、任意の P S A；または
任意の T、N 1、M 0、任意のグリーソンスコア、任意の P S A；または
任意の T、任意の N、M 1、任意のグリーソンスコア、任意の P S A。

【 0 1 2 3 】

本発明では、侵襲性の癌は機能的にあるいは臨床的に定義される：すなわち進行することができる癌をいう。これは P S A 再発によって測定することができる。患者が手術または放射線療法を受ける場合、前立腺細胞は殺されるか除去される。P S A が前立腺細胞によってのみ作られるので、患者の血液中の P S A レベルは非常に低いか検出できない量になる。癌が再発し始める場合、P S A レベルは増加し、再び検出できるようになる。これは「P S A 再発(PSA failure)」と呼ばれる。代替的測定は転移あるいはエンドポイントとしての死亡の存在である。

10

【 0 1 2 4 】

グリーソンおよび上に定義されるような病期の増加も進行と見なすことができる。しかしながら、D E S N T 特性記述はグリーソン、病期および P S A に依存しない。それは、グリーソン、病期および P S A に加えて、侵襲性の癌の進行に関する追加情報を提供する。したがって、それは結果の有用な独立した予測因子である。しかしながら、D E S N T ステータスは、グリーソン、腫瘍病期および / または P S A と組み合わせることができる。

20

【 0 1 2 5 】

したがって、発明の方法は癌を分類する方法を提供する。いくつかの方法は、バイオマーカー・パネルの 1 または複数のメンバーの発現レベルまたは発現ステータスを決定することを含む。遺伝子のパネルは本発明の方法を使用して決定されることができる。ある実施態様では、遺伝子のパネルは表 3 にリストされた遺伝子の少なくとも 5、少なくとも 10、少なくとも 15 あるいは 20 すべての遺伝子を含むことができる。遺伝子のパネルは表 2 にリストされた遺伝子の少なくとも 5、少なくとも 10、少なくとも 20、少なくとも 30、少なくとも 40 あるいはすべての 45 の遺伝子を含むことができる。本発明の他の

30

【 0 1 2 6 】

それらの遺伝子の 1 つ以上のステータス（例えば遺伝子の少なくとも 50 %、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 95 % あるいは 100 %）が、異常なステータスを持つと考えられる場合、侵襲性であると癌が記述される。異常なステータスは、健康または非侵襲性の癌状態と異なる発現ステータス（例えば D N A メチル化あるいは他の後成の要因による、発現のレベルによって決定される）として定義することができる。これは、分析される同じ遺伝子の以前に決定された参照発現ステータスによって決定されることができるし、あるいは 1 つ以上のコントロールあるいはハウスキーピング遺伝子のステータスの決定により達成されることができる。ハウスキーピング遺伝子は、進行した患者と進行していない患者で同じレベルで発現されると一般に考えられている。したがって、コントロール遺伝子に対するテスト遺伝子の比率を決定することは可能である。比率は正常な細胞と進行した組織において異なるだろう。上に記載されたように、表 6 あるいは表 7 のハウスキーピング遺伝子を使用することができる。

40

【 0 1 2 7 】

例えば、1 つの実施態様では、表 3 の遺伝子の少なくとも 15 遺伝子の、少なくとも 50 %、少なくとも 60 %、少なくとも 70 %、少なくとも 80 %、少なくとも 90 % が異常な発現ステータスを有する場合（たとえば表 3 の 15 遺伝子の少なくとも 80 %）、癌

50

は侵襲性であるか、または潜在的に侵襲性であるとされる。別の実施態様では、表 2 の遺伝子の少なくとも 40 遺伝子の、少なくとも 50 %、少なくとも 60 %、少なくとも 70 %、少なくとも 80 %、少なくとも 90 % が異常な発現ステータスを有する場合（たとえば表 2 の 40 遺伝子の少なくとも 80 %）、癌は侵襲性であるか、または潜在的に侵襲性であるとされる。

【0128】

遺伝子の発現ステータスの決定は遺伝子の発現のレベルを決定することを含むことがある。ここに使用されるような発現ステータスおよび発現のレベルは、当業者に公知の方法で決定することができる。例えば、当業者に公知の方法によって決定されるように、これは特定の遺伝子（1 または複数）のアップまたはダウンレギュレーションを指すことができる。後成の修飾（epigenetic modifications）は、発現の指標、例えば DNA メチル化ステータスの決定、または他の後成の変化、たとえばヒストンマーキング(histone marking)、RNA 変化あるいはコンフォメーション変化の指標として使用されることができる。後成の修飾は、DNA 中の遺伝子の発現を規制し、患者の間の療法の効能に影響を及ぼすことができる。異常の後成の変化は、例えば癌のような多くの疾病に関連している。動物中の DNA メチル化は投与量補正、インプリンティング、およびゲノム安定および発現に影響を及ぼす。DNA メチル化を同定する方法は、当業者に公知であり、例えばメチル化特異性 PCR、マトリックス支援レーザー脱離/イオン化飛行時間型質量分析法、マイクロアレイの使用、リデュースト リプレゼンテーション ビスルフェイト シーケンシング(reduced representation bisulfate sequencing:RRBS)あるいはホール ゲノム ショットガン ビスルフェイト シーケンシング(whole genome shotgun bisulfate sequencing:WGBS)が使用される。さらに、後成の変化は、クロマチンのコンフォメーションの変化を含んでいることがある。

【0129】

遺伝子の発現ステータスも後成の特徴を検討して判断されることがある。たとえばメチル化による DNA 中のシトシンの修飾は、遺伝子発現の変更に関係するかもしれない。評価する後成の変化の他の方法は、ヒストン修飾（マーキング）および関連する遺伝子の検討、非コード RNA の検討およびクロマチンコンフォメーション分析を含む。後成のステータスを検査するために使用することができる技術の例は次の出版物に提供される：

【0130】

- 1.Zhang, G. & Pradhan, S. Mammalian epigenetic mechanisms. IUBMB life (2014).
- 2.Gronbak, K. et al. A critical appraisal of tools available for monitoring epigenetic changes in clinical samples from patients with myeloid malignancies. Haematologica 97, 1380-1388 (2012).
- 3.Ulahannan, N. & Grealley, J. M. Genome-wide assays that identify and quantify modified cytosines in human disease studies. Epigenetics Chromatin 8, 5 (2015).
- 4.Crutchley, J. L., Wang, X., Ferraiuolo, M. A. & Dostie, J. Chromatin conformation signatures: ideal human disease biomarkers? Biomarkers (2010).
- 5.Esteller, M. Cancer epigenomics: DNA methylomes and histone-modification maps. Nat. Rev. Genet. 8, 286-298 (2007).

【0131】

発現ステータスが異常であると見いだされた場合、特定のサンプルのバイオマーカーのステータスと、平均サンプル（つまり健康なサンプルあるいは DE SNT 癌にかかっていない患者から得られたサンプル）に一般に見いだされるステータスとが異なる状況を指す。本発明では、異常な発現ステータスの存在は DE SNT 癌を示すことができる。例えば、異常なステータスは、後成の要因を使用し、または遺伝子発現のレベル（例えば RNA レベル）の決定により決定されることがある。表 2 にリストされた遺伝子に関して、遺伝子発現の減少あるいはその遺伝子の発現の減少に帰着する発現ステータスの変化は、DE

SNT癌を示す。したがって、表2にリストされた遺伝子の少なくとも5、少なくとも10、少なくとも20、少なくとも30、少なくとも40あるいは45のすべての遺伝子の異常な発現ステータスの存在は、DE SNT癌を示す。あるいは、表2にリストされ遺伝子の、少なくとも5、少なくとも10、少なくとも20、少なくとも30、少なくとも40あるいは45のすべての遺伝子の発現ステータスを表示する手段である閾値は当業者によって決定されることできる。与えられた患者のサンプルについては、平均発現ステータスが前述の閾値未満（1つ以上の遺伝子、あるいは好ましくは分析されている遺伝子の過半での発現の減少）である場合、これはDE SNT癌を示す。

【0132】

ある実施態様では、表2にリストされた遺伝子の少なくとも5、少なくとも10、少なくとも20、少なくとも30、少なくとも40あるいは45の遺伝子すべての発現ステータスあるいは発現レベルの減少は、DE SNT癌を示す。ある場合には、新しいバイオマーカー・パネルが、癌の分類に使用された本発明の方法を使用して、生成されることができ。例えば、35遺伝子データセットの異なる組み合わせを使用して行なわれた第2の分析では、遺伝子は、各データセットのLPD分析の100回の少なくとも67回でダウンレギュレートされたことが見いだされた（表5）。最初の分析で、45の共通してダウンレギュレートした遺伝子と27の遺伝子がオーバーラップしていた。したがって、バイオマーカー・パネルは表5にリストされた遺伝子の少なくとも5、少なくとも10、少なくとも20、少なくとも30あるいは35の遺伝子すべてを含むことができる。

【0133】

通常、発現ステータスが異常かどうか決定するために、患者のサンプルの少なくとも1つのコントロール遺伝子の発現ステータスの決定を含むことが必要である。少なくとも1つのコントロール遺伝子の発現ステータスに基づいて、予後の遺伝子のための指標値を決定することできる。指標値が予後の遺伝子の発現の減少のために、ある閾値未満である場合、これは癌進行を示すか、あるいは癌進行を予測する（つまりDE SNT癌）。1つ以上のコントロール遺伝子を使用し、予後の遺伝子の発現レベルを正規化し、予後の遺伝子の少なくとも50%、少なくとも60%、少なくとも70%、少なくとも80%、少なくとも90%あるいは少なくとも95%が減少した発現レベルを持っているかどうか決定することにより、閾値が決定される。ある実施態様では、減少した発現レベルを有する予後の遺伝子の100%は、癌進行を示すかあるいは予測的である（つまりDE SNT癌）。

【0134】

発明の1つの実施態様では、方法は次のものを含む：

- a) 対象の分析物（例えばRNA、DNAあるいは蛋白質）の生体試料を富化すること；および
- b) 富化したサンプル内の対象分析物の後成のステータスを決定すること。

【0135】

蛋白質は発現レベルを決定するためにも使用され、適切な方法は、当業者に知られている。これもさらに以下に議論される。

【0136】

装置とメディア

本発明は、さらに発明の任意の方法を行なうように構成された装置を提供する。図4は本明細書に示された方法の実行のための装置または計算装置100を示す。当業者によって理解されるように、図3に示されるアーキテクチャーと異なるアーキテクチャーが使用されることができ。

【0137】

図を参照する。メーター100は、ビジュアルディスプレイ110およびバーチャルまたは専用のユーザ入力デバイス112を含む多くのユーザー・インタフェースを含んでいる。メーター100はさらにプロセッサ114、メモリ116および電力システム118を含んでいる。メーター100は、プロセッサ114とリモートシステムの間でコミュニケーションを送受信するためにさらにコミュニケーション・モジュール120を含む

10

20

30

40

50

。メーター 100 は、さらに受け取り（例えば、操作された時、プロセッサ 114 にここに記述されるような方法を行なわせる指示を伴うメモリディスクまたは非一時的なコンピュータ読取り可能なメディア）のためのレシービングデバイスまたはポート 122 を含む。

【0138】

プロセッサ 114 はデータを受け取り、メモリ 116 にアクセスし、前記メモリ 116、コミュニケーション・モジュール 120 あるいはユーザ入力デバイス 112 から指示を受け取った際に実行するように構成される。プロセッサはディスプレイ 110 をコントロールし、コミュニケーション・モジュール 120 を介してリモート機器にデータを伝えることができる。

10

【0139】

メモリ 116 は、プロセッサによって読まれた時、プロセッサに本明細書に記述される方法を行なわせるように構成される、コンピュータ読取り可能な指示を含むことができる。

【0140】

本発明は、指示を格納した機械読取り可能な媒体（これは一時的または非一時的であることができる）を提供する。指示は、機械によって読まれた時、指示が実行され本明細書に示される方法を引き起こすように形成される。

【0141】

他の方法および発明の用途

20

さらに診断または癌の進捗の決定を支援するために、発明の方法は、さらなるテスト（例えば P S A 検査、グリーソンスコア分析）と組み合わせることができる。P S A 方法では、血液サンプル中の前立腺特異抗原の量が定量化される。

前立腺特異抗原は前立腺の細胞によって生産された蛋白質である。レベルが血液の中で上がった場合、これは前立腺癌を示す。「上がった」量は、患者の詳細（例えば年齢）に依存するだろうが、一般にレベルが高いほど、前立腺癌が存在するということである。期間（例えば一週、一か月、6 か月あるいは一年）にわたる P S A レベルの連続的上昇も、さらに前立腺癌の表れである。たとえば 4 ナノグラム / m l あるいは 10 ナノグラム / m l 以上の P S A レベルは前立腺癌を示すが、前立腺癌は 4 以下の P S A レベルの患者でも見つかる。

30

【0142】

発明のある実施態様では、方法は、非侵襲性の癌から侵襲性の癌（たとえば侵襲性の前立腺癌）を差別的に診断することができる。これは癌の D E S N T ステータスの決定により達成することができる。別法としてあるいはさらに、定量されている遺伝子のそれぞれについての試験サンプル内の発現レベルと、適切な参照（例えば健康な患者、または非侵襲性の癌に苦しむ患者からのサンプル）の患者、またはコントロールまたはハウスキーピング遺伝子を使用した発現レベルとの比較により達成されることができる。このように、不必要な治療は回避することができる。また、適切な治療が代わりに適用される（例えばフルオキセチン(fluoxetine)、ガバペンチン(gabapentin)あるいはアミトリプチリン(amitriptyline)のような前立腺炎の抗生物質治療あるいはフィナステリド(Finasteride)のようなアルファ・リダクターゼ阻害剤での治療)。

40

【0143】

発明の 1 つの実施態様では、方法は、以下のステップを含む：

- 1) 患者から得られた生体試料の R N A を検知すること；および
- 2) R N A 分子の各々の発現レベルを定量すること。

【0144】

検知された R N A 転写物は、定量されるバイオマーカーに相当する（従ってその発現レベルが測定されている遺伝子）。ある実施態様では、検知されている R N A（例えば m R N A、I n c R N A あるいはスモール R N A）は、表 2 にリストされた少なくとも 40 の遺伝子（任意に表 2 にリストされた遺伝子の少なくともすべて）に対応し、または表 3 に

50

リストされた少なくとも15の遺伝子（任意に表3にリストされた遺伝子のすべて）に対応する。そのような方法は、先に患者から得られたサンプル上で試みられることができる。任意に患者はDREを行い、前立腺をマッサージし、得られるサンプル中のRNAの量を増加させることができる。あるいは、方法はそれ自身、患者から生体試料を得る1ステップを含むことができる。

【0145】

1つの実施態様では、検知されたRNA転写物は、表1にリストされた遺伝子の選択またはすべてに相当する。その後、遺伝子の部分集合は、たとえばLDP分析のようなさらなる分析のために選ぶことができる。

【0146】

発明のある実施態様では、生体試料は検知と定量に先立ってRNA（あるいは蛋白質のような他の分析物）を富化することができる。しかしながら、富化のステップは任意である。また、代わりに、RNAは、全尿のような生の加工されていない生体試料から得ることができる。富化のステップは、サンプル中のRNA（あるいは他の分析物）の濃度を増加させる任意の適切な前処理方法ステップであることができる。例えば、富化のステップは、サンプルから細胞を除去するための遠心分離と濾過を含むことができる。

【0147】

発明の1つの実施態様では、方法は以下を含む：

- a) 増幅、濾過あるいは遠心分離によってRNAについて生体試料を富化すること、任意に生体試料はDREを経験した患者から得られた；
- b) 富化したサンプルのRNA転写物を検知すること；および
- c) 検知されたRNA分子の各々の発現レベルを定量すること。

【0148】

検知のステップは、ハイブリダイゼーション、増幅、シーケンシング、または分子量および/または電荷の検知、または細胞の表現型の変化、あるいは特定の分子の結合の検知、あるいはそれらの組み合わせに基づく検知方法を含むことができる。

【0149】

ハイブリダイゼーションに基づいた方法はノーザンブロット、マイクロアレイ、ナノストリング(NanoString)、RNA-FISH、分岐鎖ハイブリダイゼーション分析、および関連方法を含んでいる。増幅に基づいた方法は、定量逆転写ポリメラーゼ連鎖反応(quantitative reverse transcription polymerase chain reaction:qRT-PCR)、トランスクリプション メディエーティッド 増幅(transcription mediated amplification)、および関連方法を含んでいる。シーケンシングに基づいた方法は、サンガー・シーケンシング(Sanger sequencing)、次世代シーケンシング（合成による高スループットシーケンシング）、およびターゲッティッドRNAseq(targeted RNAseq)、ナノポアメディエーティッドシーケンシング(nanopore mediated sequencing:MinION)、質量分析検知および関連方法を含んでいる。分子量および/または電荷の検知に基づいた方法は、質量分析を含むが、これに制限されない。表現型の変化に基づいた方法は、miRNAsをスクリーニングする方法（例えば、Cullen & Arndt, Immunol. Cell Biol., 2005, 83:217-23を参照）により、テスト・セル、あるいは動物の細胞内の変化を検知することができる。特定の分子の結合に基づいた方法は、例えば抗体あるいはRNAまたはDNA結合性蛋白質のような他の結合分子との結合の検知を含んでいる。

【0150】

ある実施態様では、方法は、RNA転写物をcDNA転写物に変換する1ステップを含むことがある。そのような方法ステップは、方法の任意の時に行うことができ、たとえば富化の前（このステップが行われる場合、富化ステップはcDNA富化ステップである）、検知の前（その場合には検知ステップはcDNA検知のステップである）、あるいは定量の前（その場合には検知された各cDNAシーケンスの転写物の数を数えることによる検知されたRNA分子の各々の発現レベルである）に行うことができる。

【0151】

10

20

30

40

50

発明の方法は、検知され定量されるRNAあるいはcDNAの量を増加させる増幅ステップ含んでいることがある。増幅の方法はPCR増幅を含んでいる。

【0152】

発明のいくつかの方法では、cDNA-結合分子複合体の検知および定量化は遺伝子発現を決定するために使用されることができる。例えば、サンプル中のRNA転写物は、逆転写によってcDNAに変換されることができ、その後サンプルは定量されている遺伝子に特異的な結合分子と接触され、DNA特異的な結合分子複合体の存在を検知し、対応する遺伝子の発現を定量する。

【0153】

したがって、バイオマーカー・パネルの中で同定された1以上の遺伝子に対応するcDNA転写物の、検知方法での使用、前立腺癌の予後の診断または決定での使用を提供する。

【0154】

一旦発現レベルが定量されれば、癌（特に侵襲性の前立腺癌）の診断を決定することができる。発明の方法は患者の予後を決定するため、治療への患者の反応を決定するため、患者の癌の治療適応性を決定するために使用することができる。方法は癌進行を予測するために使用することができるからである。

【0155】

方法はさらに、定量化された発現レベルを参照と比較し、次いで癌、特に侵襲性の前立腺癌の存在または不存在を決定するステップをさらに含むことができる。

【0156】

分析物富化は任意の適切な方法によって達成されることができ、遠心分離および/または濾過によりサンプルからセルデブリスを取り除くことが好ましい。富化したサンプルからRNAを得るステップは、富化したサンプルの中にあるマイクロベシクル(microvesicles)からRNAを収穫することを含んでいることがある。

【0157】

RNAシーケンシングステップは任意の適切な方法によって達成されることができ、直接RNAシーケンシング、RT-PCRあるいは、シーケンシング-バイ-シンセシス(sequencing-by-synthesis)(次世代、あるいはNGS、高スループットシーケンシングが好ましい。定量化は任意の適切な方法によって達成されることができ、たとえば、特定のシーケンスで特定された転写物の数を数えるにによって達成することができる。1つの実施態様では、シーケンス(通常75-100の塩基対)はすべてヒトの参照とアラインされる。適切なデータベース(例えばEnsemblデータベース)に定義された各遺伝子について、シーケンスの数、あるいはその遺伝子とオーバーラップする(そしてどんな他のものともオーバーラップしない)読みを、数える。サンプル間の遺伝子を比較するために、各サンプルを正規化することが通常必要であり、量はシーケンスされたデータと等価な合計となるようにされる。正規化の方法は当業者に明白である。

【0158】

当業者に明白なように、使用されている試験サンプルのタイプを考慮するため、および/または分析に先立って生じた試験サンプルを処理するため、分析物濃度のどんな測定値も正規化されることが必要である。

【0159】

遺伝子の発現レベルはコントロールと比較され、発現のレベルが分析されているサンプルでのものより高いか低いかを決定することができる。分析されたサンプルの発現のレベルが、分析されたサンプルと比較されるサンプル中の発現のレベルより高い場合、遺伝子はアップレギュレートされたと言える。分析されたサンプルの発現のレベルが、分析されたサンプルと比較されるサンプル中の発現のレベルより低い場合、遺伝子はダウンレギュレートされたと言える。

【0160】

発明の実施態様では、遺伝子の発現のレベルは予兆であることがある。そのため、本発明は、介在を要求する前立腺癌(侵襲性の前立腺癌)、および介在を要求しない癌(イン

10

20

30

40

50

ドレントまたは非侵襲性の前立腺癌)を区別するのに有用であり、不必要な治療およびそれらに関連する副作用を回避するのに特に役立つ。本発明の最も有望な使用は、追加の患者のサンプルがL P D分析によりD E S N Tかどうか決定するための5 0 0の遺伝子パネルの使用、患者のサンプル中の遺伝子のダウンレギュレーションを測定することによりD E S N Tかどうか決定するための4 5の遺伝子パネルの使用、およびR F分析による2 0の遺伝子パネルの使用だろう。

【0 1 6 1】

発明のある実施態様では、バイオマーカー・パネルは、P S A検査、P C A 3テスト、プーロラリス(Prolaris)あるいはオンコタイプ(Oncotype) D Xテストのような別のテストと組み合わせることができる。他のテストは、グリーソンスコアを決定する組織学的検査、あるいは癌の進行の病期の事前評価であることがある。

10

【0 1 6 2】

発明の一層の実施態様では、前立腺癌の治療のための患者の適応性を決定する方法であって、本発明の方法によって癌を分類すること、癌進行が診断されるか疑われる場合に、特に侵襲性の前立腺癌が診断されるか疑われる場合に、前立腺癌の治療を続けるかどうか決定することを含む方法が提供される。

【0 1 6 3】

患者の療法に対する反応をモニターする方法であって、本発明の方法によって、先に前立腺癌の治療(例えば化学療法および/または放射線療法)を受けた患者から得られた生体試料を使用して、癌を分類することを含む方法がさらに提供される。ある実施態様では、方法は治療を受ける前後の患者の中で繰り返される。その後、発現のレベルの比較に基づいて、治療を継続するべきか、代替療法を試みるべきであるかどうかの決定をなすことができる。例えば、治療を受けた後にD E S N T癌が検知されるか疑われる場合、代替治療療法が使用されることができる。方法は治療がD E S N Tから非D E S N Tに患者の癌をダウングレードすることに成功するかどうか確かめるために繰り返すことができる。

20

【0 1 6 4】

1つの実施態様では、以下を含む方法が提供される：

- a) 患者から得られたサンプルに本発明の診断法を行い、D E S N T癌(たとえばD E S N T前立腺癌)の存在または非存在を決定すること；
- b) D E S N T癌が見つかるか疑われる癌について治療を提供すること；
- c) 患者から得られたさらなるサンプルに本発明の診断法を続いて行い、D E S N T癌の存在または非存在を決定すること；および
- d) 癌のための療法を維持するか、変更するか、終了すること。

30

【0 1 6 5】

発明のある実施態様では、たとえば治療に引き続く尿中のバイオマーカー発現を測定することにより、容易に異なる治療の影響をモニターすることができるので、本発明の方法とバイオマーカーのパネルは患者の治療を個別的に取り扱うのに有用である。発明の方法およびバイオマーカーは、たとえばホルモン除去療法に対する反応のような治療の有効性を予測するために使用することができる。

【0 1 6 6】

発明の別の実施態様では、患者の癌(たとえば侵襲性の前立腺癌)を治療するか防ぐ方法であって、患者から得られたサンプルに本発明の診断法を適用し、D E S N T癌の存在または非存在を決定し、D E S N T癌が検知されるか疑われれば癌治療を施す方法が提供される。前立腺癌を治療する方法は腫瘍を切除しおよび/または患者へ化学療法および/または放射線療法を施すことを含むことがある。

40

【0 1 6 7】

本発明の癌を治療する方法は、侵襲性の前立腺癌の治療に特に役立つ。ある実施態様では、治療法は、D E S N T癌を持っていると同定された患者に行なわれる。

【0 1 6 8】

可能であれば、前立腺癌の治療は腫瘍の切除あるいは他の外科技術を含んでいる。例え

50

ば、治療は根本的または部分的な前立腺摘除、経尿道的切除(trans-urethral resection)、精巣摘除あるいは両側の精巣摘除を含むことがある。治療は、別法としてあるいはさらに化学療法および/または放射線療法による治療を含むことがある。化学療法の治療はドセタセル(docetaxel)、アビラテロン(abiraterone)あるいはエンザルタミド(enzalutamide)を含んでいる。放射線治療(Radiotherapeutic treatment)は体外照射療法(external beam radiotherapy)、骨盤の放射線療法、手術後の放射線療法、近距離照射療法あるいは場合に依りて予防的な放射線療法を含んでいる。他の治療は、補助ホルモン療法(たとえばアンドロゲン抑制療法、寒冷療法、高密度焦点式超音波療法、免疫療法、近距離照射療法および/またはビスホスホネートおよび/またはステロイド投与)を含んでいる。

【0169】

10

発明の別の実施態様では、次のものを含む癌の治療に役立つ薬を同定する方法が提供される：

- a) 患者から得られたサンプルに発明の診断法を行い、DESENT癌の存在か非存在を決定すること；
- b) 患者への候補薬を投与すること；
- c) 患者から得られたさらなるサンプルに発明の診断法を続けて行い、DESENT癌の存在か非存在を決定すること；および
- d) ステップc)で発見された知見とステップ(a)で発見された知見を比較すること、ここでDESENT癌の可能性の低減は薬候補を癌の可能な治療薬であると同定する。

【0170】

20

生体試料

発明の方法は、生体試料に行なわれるステップを含むことがある。分析される生体試料は、尿サンプル、精液サンプル、前立腺の滲出液サンプル、あるいは前立腺に由来する高分子または細胞を含んでいる任意のサンプル、全血サンプル、漿液サンプル、唾液、あるいは生検試料(たとえば前立腺組織試料あるいは腫瘍サンプル)であることができる。前立腺癌についての最も一般的な生体試料は、前立腺生検、前立腺摘除あるいはTURPからのものである。方法は、生体試料を得るか提供するステップを含んでいることがある。あるいは、サンプルはたとえば生体外の方法で患者から得られたものであることができる。サンプルは、潜在的に癌である前立腺組織、または前立腺内の他の細胞、あるいは前立腺、血液、あるいは免疫系内の細胞によって生産された微小小胞体に関連する遺伝子の発現のレベルを表すと考えられる。従って、本発明の方法は、癌に反応する前立腺、および/または血液システムおよび/または骨髓内の細胞によって生産されたRNAの定量データを使用して前立腺癌の存在か非存在を決定する。

30

【0171】

発明の方法は患者からの1つの試験サンプルに対して実行されることができる。あるいは、複数の試験サンプル(例えば少なくとも2、3、4あるいは5サンプル)が、患者から受け取られることがある。サンプルはそれぞれ発明の方法を使用して個別の分析に供されることができるし、あるいは診断を行う一人の患者からの多数のサンプルに方法が適用されることができる。

【0172】

40

発明の中で使用されるさらなる分析法

発明のバイオマーカー・パネルからの遺伝子または蛋白質の発現のレベルは、多くの方法で決定することができる。バイオマーカーがサンプル内のタンパク質として発現される場合、発現のレベルは、たとえばサンプル中の蛋白質の濃度を決定することによりバイオマーカーを定量することにより決定することができる。あるいは、サンプル(たとえば組織試料)中のRNAあるいは蛋白質の量が決定されることができる。一旦発現のレベルが決定されたならば、レベルは、任意にコントロールと比較することができる。これは先に測定された発現レベル(同じ被験者から異なる時期に得られたサンプル中で、あるいは異なる被験者(例えば健康な被験者、非侵襲性の癌(つまりコントロールまたは参考)の被験者)から得たサンプル中で)、異なる蛋白質またはペプチド、あるいは他のマーカー、

50

または、同じサンプル内において、分析されているサンプル内で発現またはタンパク濃度のレベルがより高いかより低いかがどうか決定するアセスメント手段であることができる。ハウスキーピング遺伝子もコントロールとして使用することができる。理想的には、コントロールはサンプル間で一般に著しく変わらない蛋白質かDNAマーカーである。

【0173】

遺伝子発現を定量する他の方法は、RNAシーケンシングを含んでいる。その1つの態様は全体トランスクリプトームショットガンシーケンシング(whole transcriptome shotgun sequencing: WTSS)として知られている。RNAシーケンシングの使用は、サンプルの中にあるRNAシーケンスの特性(nature)を決定することができ、そして各RNA分子(例えばmRNAあるいはmicroRNAトランスクリプト)の存在量の測定により、更に遺伝子発現を定量する。方法は、サンプルの分析はシーケンシング・バイ・シンセシス(sequencing-by-synthesis)のアプローチを使用する。

10

【0174】

使用することができるいくつかのタイプのRNAシーケンシングがあり、たとえばRNA PolyAテイルシーケンシング(RNA PolyA tail sequencing)(RNAシーケンスのpolyAテイルがpoly Tオリゴヌクレオチドを使用してターゲットとされている)、ランダム・プリムドシーケンシング(random-primed sequencing)(ランダムなオリゴヌクレオチド・プライマーを使用して)、ターゲッティドシーケンス(targeted sequencing)(特定遺伝子トランスクリプトに補足的な特異性オリゴヌクレオチド・プライマーを使用して)、小分子RNA/非コードRNAシーケンシング(それはサイズ分離を使用して、microRNAsのような小さな非コードのRNAを分離することを含むことができる)、ダイレクトRNAシーケンシングおよびリアルタイムPCRを含む。ある実施態様では、RNAシーケンス読み取りは、参照ゲノムにアラインされ、遺伝子発現を決定するために各シーケンスの読み取りの数を定量する。発明のある実施態様では、方法は転写アセンブリ(デノボ(de-novo)あるいはゲノム・ガイドされた)を含む。

20

【0175】

RNA、DNAおよび蛋白質配列(マイクロアレイ)はある実施態様の中で使用されることができる。RNAとDNAのマイクロアレイは、DNAまたはRNAのオリゴヌクレオチドの一連の微視的なスポット(microscopic spots)を含み、各々は補足的な核酸分子に結合することができるヌクレオチドのユニーク配列を含む。このように、オリゴヌクレオチドは、高い厳格な条件の下で正確な標的配列がハイブリダイズするプローブとして使用される。本発明では、標的配列は、その発現が検知されている遺伝子に対応する、転写されたRNA配列またはそれらのユニークな部分であることができる。蛋白質マイクロアレイも直接蛋白質発現を検知するために使用することができる。これらは固体の表面に固定された捕獲分子(capture molecules)を含む、DNAおよびRNAマイクロアレイに似ている。

30

【0176】

捕獲分子は抗体、蛋白質、アプタマー、核酸、レセプタおよび酵素を含んでいる。商用抗体が検知される分析物に利用可能でない場合、それは望ましいことがある。アレイに使用するための捕獲分子は、外部的に合成し、精製し、アレイに付着させることができる。あるいは、それらはその場で合成し、配列に直接付着させることができる。捕獲分子は生合成、無細胞のDNA発現あるいは化学合成によって合成することができる。その場合合成は後の2つで可能である。

40

【0177】

一旦マイクロアレイ上で捕らえられたならば、検査法は公知技術の任意のもののいずれかでありえる。例えば、蛍光検知を使用することができる。それは敏感で、安全で、高解像度でありえる。他の検査法としては、他の光学的手法(例えば比色分析、化学発光、ラベルフリー表面プラスモン共鳴分析(label free Surface Plasmon Resonance analysis)、顕微鏡検査法、反射率など)、質量分析、電気化学的方法(例えばボルタメトリー(voltametry)および電流測定)および無線周波数方法(例えば多極共鳴分光法(multipolar r

50

esonance spectroscopy)) を含んでいる。

【0178】

RNAまたはcDNAの検知のための方法は、ハイブリダイゼーション(例えばノーザンブロット、マイクロアレイ、ナノストリング(NanoString)、RNA-FISH、枝分れ鎖ハイブリダイゼーション分析、TaqManまたはSYBR green生成物検出のような定量的逆転写ポリメラーゼ連鎖反応(qRT-PCR)のための増幅検査法に基づくことができる。次のもののようなプライマー伸張検知方法: 単一のヌクレオチド伸張、サンガー・シーケンシング。あるいは、RNAは、サンガー・シーケンシング、次世代(高生産性)シーケンシング、特にシーケンシング-バイ-シンセシス、プリサイスターゲッティドRNAseq分析(recise targeted RNAseq assays)のようなターゲッティドRNAseq分析、あるいはオックスフォード・ナノポアMinION装置のような分子検出装置によりシーケンスされる。上記の技術の組み合わせは、磁気ビーズを介した分子捕獲、転写増幅およびたとえば化学発光による検知のための第2のプロブとのハイブリダイゼーションを使用するGen-Probe PCA3分析において使用される、Transcription Mediated Amplification (TMA)において使用される。

10

【0179】

RNAは検知に先立ってcDNAに変換されることができる。RNAまたはcDNAは、検知の前にあるいは検知の一部として増幅されることができる。テストは、テスト・セル内の表現型の1または複数の変化によってRNA、蛋白質あるいは他の高分子の存在を検知することができる機能テストを構成することができる。表現型の変化は、運動性または侵襲の変化を含む。

20

【0180】

一般に、電気泳動に供された蛋白質も質量分析によりさらにキャラクタライズされる。そのような質量分析方法はマトリックス-アシスティドレーザ脱離/イオン化飛行時間(matrix-assisted laser desorption/ionisation time-of-flight: MALDI-TOF)を含むことができる。

【0181】

MALDI-TOFは、生体分子(たとえば蛋白質、ペプチドおよび糖)の分析を可能とするイオン化技術である。それは、従来のイオン化方法によってイオン化された時に脆弱でフラグメントになる傾向がある。イオン化はレーザー・ビーム(例えば窒素レーザー)が引き金となって起きる。また、マトリックスが、直接のレーザー・ビームへの暴露によって破壊されることから生体分子を防御し、かつ気化とイオン化を促進するために使用される。サンプルは溶液でマトリックス分子と混合される。また、少量の混合物が表面に残され、乾燥される。溶剤が蒸発するとともに、サンプルとマトリックスは共結晶する。

30

【0182】

タンパク濃度を決定するさらなる方法は、質量分析および/またはたとえばLC-MS、UPLC、タンデムUPLC-MS/MSシステムおよびELISA法のような液体クロマトグラフィーを含んでいる。発明の中で使用されることができる他の方法は、Agilent バイトキャプチャー(Agilent bait capture)およびPCRに基づいた方法(例えば、PCR増幅は分析物の量を増加させるために使用されることができる)を含んでいる。

40

【0183】

発明の方法は、分析物(定量されるRNA分子または蛋白質)に特異的な結合分子あるいは試薬を使用して実行することができる。結合分子と試薬は、検知されるRNA分子または蛋白質への親和性がある分子であり、それらが公知の方法を使用して検知することができる結合分子/試薬-分析物複合体を形成することができるようにされる。発明の結合分子はオリゴヌクレオチド、またはオリゴリボヌクレオチド、ロックド核酸(locked nucleic acid)あるいは他の同様の分子、抗体、抗体フラグメント、蛋白質、アプタマーあるいは分子でインプリントされた重合体構造、あるいはDNAまたはRNAと結合することができる他の分子であることができる。発明の方法は適切な1または複数の結合分子と生体

50

試料と接触することを含むことができる。結合分子が発明のキットの一部を形成することがある。特にそれらは本発明のバイオセンサーの一部を形成することがある。

【0184】

アプタマーは特定の標的分子に結合するオリゴヌクレオチドあるいはペプチド分子である。オリゴヌクレオチド・アプタマーはDNAアプタマーおよびRNAアプタマーを含んでいる。アプタマーはランダム・シーケンス・オリゴヌクレオチドあるいはペプチドのプールから生体外の選別過程によって作成することができる。アプタマーはリボザイムと結合することができ、それらの標的分子が存在する状態で自己開裂する。他のオリゴヌクレオチドは、定量されているRNA分子と相補的なRNA分子を含んでいることがある。例えば、poly T oligosはRNA分子のpoly Aテイルをターゲットとするために使用することができる。

10

【0185】

アプタマーは公知の任意のプロセスによって作ることができる。例えば、アプタマーが同定されることができるプロセスは、systematic evolution of ligands by exponential enrichment(SELEX)である。これは標的分子への選択的な結合に基づいて分割することにより、分子のライブラリーの複雑さを反復して低減すること、およびその後の再増幅を含む。潜在的なアプタマーのライブラリーは目標蛋白質でインキュベートされ、結合していない要素が結合した要素から分離される。結合された要素は回収され増幅され(例えばポリメラーゼ連鎖反応によって)、低減された複雑さ(たとえば富化されたプール)のライブラリーを生産する。富化したプールはSELEXの第2のサイクルを始めるために使用される。

20

目標蛋白質への引き続く富化されたプールの結合はサイクルごとにモニターされる。一旦、結合分子の割合が適切なレベルに上昇したと判断されれば、富化したプールのクローンが作られる。その後、結合分子は個々に分析される。SELEXはFitzwater & Polisky (1996) Methods Enzymol, 267:275-301でレビューされる。

【0186】

抗体はモノクローナルとポリクローナル抗体の両方を含むことができ、公知の任意の手段によって生産することができる。特別の蛋白質に結合するモノクローナルおよびポリクローナル抗体を生産するための技術は、当該技術分野でよく開発されている。それらは、標準的な免疫学教科書、例えば、Roitt et al., Immunology, second edition (1989), Churchill Livingstone, Londonで議論される。抗体は人間からのもの、ヒト化されたもの、あるいは、他の種からのものであることがある。本発明は、抗原に結合することができる抗体誘導体を含んでいる。したがって、本発明は抗体フラグメントおよび合成構成物を含んでいる。抗体フラグメントおよび合成構成物の例は、Dougall et al. (1994) Trends Biotechnol, 12:372-379中で与えられる。たとえばFab, F(ab')₂またはFvのような、抗体フラグメントあるいは誘導体は使用されることができる。Huston et al. (1993) Int Rev Immunol, 10:195-217に記述されたような単鎖抗体(scAb)も使用することができる。ドメイン抗体(dAbs)、例えば単ドメイン抗体、あるいは抗体様の単ドメイン抗原結合性レセプターも使用することができる。さらに、抗体フラグメントおよび免疫グロブリン様分子、ペプチドミメティクス(peptidomimetics)あるいは非ペプチド・ミメティクスは、抗体の結合活性を模倣する様にデザインすることができる。Fvフラグメントは単鎖Fv(scFv)分子として知られている合成構成物を生産するために変成することができる。これは、分子の安定性に寄与するVHおよびVL領域に共有結合で連結するペプチド・リンカーを含んでいる。

30

40

【0187】

他の合成構成物はCDRペプチドを含んでいる。これらは抗原結合決定基(antigen binding determinant)を含む合成ペプチドである。これらの分子は、CDRループの構造を模倣し、抗原相互作用性の側鎖を含んでいるコンフォメーションが制限された有機的な環である。合成構成物はさらにキメラ分子を含んでいる。合成構成物は、さらに抗原結合に加えて分子にある望ましい性質を提供する、共有結合でリンクした部位を含む分子を含ん

50

でいる。例えば、部位は、ラベル（例えば蛍光性か放射性のラベルのような検知できるラベル）、ヌクレオチドあるいは薬学的に活性な作用薬であることができる。

【0188】

結合分子が抗体か抗体フラグメントである発明のそれらの実施態様では、発明の方法は公知の任意の免疫手法を使用して行なうことができる。例えば、ELISA、ラジオイムノアッセイあるいは同様の技術が利用できる。一般に、適切な自己抗体は固体の表面上で不動化される。また、テストされるサンプルは自己抗体と接触させられる。自己抗体によって認識された癌マーカー蛋白質がサンプルの中にある場合、抗体マーカー複合体が形成される。その後、複合体は、例えば、マーカー蛋白質のエピトープを特異的に認識するラベルが付けられた第2の抗体を使用して、定量的に測定することができる。ラベルは、例えばワサビダイコンのペルオキシダーゼ（HRP）あるいはアルカリフォスファターゼ（AP）のような生化学的標識で第2の抗体に付けられることがある。また、複合体の検知は、比色定量、化学発光、または蛍光性の生成物を生成する酵素用の基質の追加によって達成することができる。あるいは、複合体の存在は、検知できるラベル（例えば適切な酵素）でラベルが付けられたマーカー蛋白質の追加によって決定されることができる。この場合、酵素活性の測定量は、形成された複合体の量に反比例する。また、負のコントロールが、サンプル中の抗原の存在の決定するために参照として必要である。複合体を検知する別の方法は、放射性同位元素でラベルが付けられた抗体あるいは抗原を利用し、放射能の測定を行うものである。抗原のための放射性のラベルの例は ^3H 、 ^{14}C および ^{125}I を含んでいる。

【0189】

発明の方法は定性的フォーマット（それはサンプル中の癌マーカー分析物の存在か非存在を決定する）、あるいは定量的フォーマット（それはさらに、サンプルの中にある癌マーカー分析物の量の測定を提供する）で行なうことができる。一般に、発明の方法は定量的である。サンプルの中にあるバイオマーカーの量は上記の任意の技術を使用して計算されることができる。この場合、分析を行なう前に、既知量の癌マーカー分析物を含んでいるか既知の濃度の癌マーカー分析物の一連の標準試料からの分析に使用するために同じ検知反応を使用して得られた信号の測定により、標準曲線を描くことが必要なことがある。スクリーンされるサンプルの中に存在する癌マーカーの量は、その後標準曲線から推定された。

したがって、本発明の中で使用されるような遺伝子発現を決定する方法は、ポリヌクレオチドのハイブリダイゼーション分析に基づいた方法、ポリヌクレオチドのシーケンシングに基づいた方法、プロテオミクス・ベースの方法、逆転写PCR、マイクロアレイ・ベースの方法および免疫組織化学ベースの方法を含んでいる。計量遺伝子発現に関係のある参照も上に提供される。

【0190】

キットの部分とバイオセンサー

発明の一層の実施態様では、発明のバイオマーカーの発現または濃度を定量するための手段、あるいは発明のバイオマーカーの発現ステータスを決定する手段を含む、癌進行予測（DESENT癌の検知）のための部分のキットが提供される。手段は任意の適切な検出手段であることがある。例えば、ここに議論されるように、手段はバイオセンサーであることがある。キットは、さらに1または複数のサンプルの包装容器および/または生体試料からバイオマーカーを抽出するための溶剤を含むことができる。キットは、さらに使用のための指示を含むことがある。

【0191】

発明のある実施態様では、発明のバイオマーカーの発現ステータス（例えば発現のレベル）を検知するための手段を含む、癌の分類（例えば、癌進行の可能性を決定する）のための、部分のキットが提供される。バイオマーカーを検知するための手段は、定量されるバイオマーカーに特異的に結合するか、反応する試薬であることがある。したがって、発明の1つの実施態様では、定量されるバイオマーカー分析物に特異的に結合する試薬また

は分子と、患者からの生体試料とを接触させること、分析物 - 試薬あるいは分析物 - 結合分子の複合体の量を測定すること、生体試料中の関連する蛋白質あるいは遺伝子の発現のレベルと、分析物 - 試薬あるいは分析物 - 結合分子の複合体の量を相関させることを含む前立腺癌を診断する方法を提供する。

【0192】

例えば、発明の1つの実施態様では、方法は、以下のステップを含む：

- a) 発明のバイオマーカーの1つ以上に特異性の試薬または結合分子と生体試料を接触させること；
- b) バイオマーカーについて分析物 - 試薬あるいは分析物 - 結合分子複合体の量を定量すること；および
- c) 分析物 - 試薬あるいは分析物 - 結合分子複合体の量を生体試料中のバイオマーカーの発現レベルと関連させること。

【0193】

方法は、ステップc)中のバイオマーカーの発現レベルを参照と比較し癌のステータスを分類すること、特には癌進行の可能性、そしてその結果治療の必要性（侵襲性の前立腺癌）を決定するステップd)をさらに含むことができる。もちろん、ある実施態様では、方法は本発明に述べられていたような統計分析を行なうことをさらに含むことができる。その後、患者は治療されることができる。適切な試薬あるいは結合分子としては、抗体が抗体フラグメント、オリゴヌクレオチド、アプタマー、酵素、核酸、細胞小器官、細胞、生物学的組織、インプリントされた分子あるいは小分子を含む。そのような方法は発明のキットを使用して実行されることができる。

【0194】

部分のキットは、メモリとプロセッサを有する器具または装置を含むことがある。メモリは格納された指示を含むことができ、プロセッサによって読まれた時、プロセッサに上に記述された方法の1つ以上を行なわせる。メモリは、さらにランダムフォレスト分析で使用される複数のデシジョンツリーを含むことができる。

【0195】

発明の部分のキットはバイオセンサーであることができる。バイオセンサーは生物学的感知要素を組み込み、生体試料についての情報（例えば分析物の存在（あるいは非存在）または濃度）を提供する。具体的には、それらは生理化学的な検出器とバイオ認識コンポーネント（バイオレセプタ）を組み合わせ、分析物（たとえばRNAまたは蛋白質）の検知および/または定量を行う。例えば、バイオレセプタは対象の分析物と特異的に相互作用するかまたは結合し、たとえば抗体が抗体フラグメント、酵素、核酸（たとえばアプタマー）、細胞小器官、細胞、生物学的組織、インプリントされた分子あるいは小分子であることができる。

バイオレセプタは、支持体（例えば金属、ガラス、高分子支持材、ヒドロゲル支持体のような3次元の格子支持体）上に不動化されることができる。

【0196】

バイオセンサーは、存在するバイオ変換器(biotransducer)のタイプによってしばしば分類される。例えば、バイオセンサーは、電気化学（たとえば電位差計）、電子、圧電気、重量分析、ピロ電気のバイオセンサーあるいはイオンチャネル・スイッチ・バイオセンサーであることができる。分析物の存在量を正確に決定することができるように、変換器は、対象の分析物とバイオレセプタの間の相互作用を定量的な信号に翻訳する。光学的バイオセンサーは、バイオレセプタと対象の分析物の間の相互作用に起因する表面プラズモン共鳴に依存することがある。SPRは試験サンプル中の分析物の量を定量化するために使用することができる。他のタイプのバイオセンサーは、エバネセント波バイオセンサー(evanescent wave biosensor)、ナノバイオセンサーおよび生物学的バイオセンサー（例えば酵素、核酸（たとえばRNAまたはアプタマー）、抗体、エピジェネティック(epigenetic)、細胞小器官、細胞、組織あるいは微生物バイオセンサー）を含んでいる。

【0197】

本発明は、定量されるバイオマーカーの各々に特異性の捕獲分子（たとえばRNAまたはDNAオリゴヌクレオチド）を含むマイクロアレイ（RNA、DNAあるいは蛋白質）を提供する。そこでは捕獲分子は固体担体上に不動化される。マイクロアレイは発明の方法に有用である。

【0198】

発明の1つの実施態様では、発明のバイオマーカーの1つ以上の発現レベルを決定することを含む前立腺癌を分類する方法が提供され、方法は決定された値を参照と比較することを任意に含む。

【0199】

分析されるバイオマーカーは、発明の方法によって決定することができる。あるいは、ここに提供されるバイオマーカー・パネルを使用することができる。表3にリストされた遺伝子の少なくとも15（好ましくは20すべて）は、前立腺癌を分類するのに役立つ。

【0200】

表2にリストされた遺伝子の少なくとも40（好ましくは45すべて）は、いくつかの異なるタイプの癌を分類し、前立腺癌の分類を含む癌の進行の可能性を決定するのに役立つ。第2および引き続く発明の態様の特徴は、発明の最初の態様の特征に準ずる。

【0201】

本発明は、以下の例を参照してさらに記述される。それは例示のためにのみ存在し、本発明を制限するものではない。

【0202】

実施例では、以下の多くの図が言及される：

図1：ラテントプロセスデコンポジション(Latent Process Decomposition:LPD)、遺伝子相関および臨床結果

a. MSKCCデータセットからのアフィメトリックス発現データ(Affymetrix expression data)のLPD分析は、サンプルを8つのプロセス（各々棒グラフによって表わされた）に分割した。サンプルは8つのプロセスすべてで表わされる。また、各棒の高さは、各LPDプロセスに割り当てることができるシグネチャーの割合(p_i)に相当する。サンプルは、それらが p_i の最高値を示すLPDグループに割り当てられる。LPDは、500の遺伝子プローブを使用して、MSKCCデータセットの中でサンプル間の発現における最も大きな変異について行なわれた。DESENT癌を含むプロセスが示される。

b. LPD分析において使用されたデータセットのリスト。LPDの中で使用される原発性癌および正常な標本のユニークな数が示される。FF：新鮮な凍結された標本；FFPE、ホルマリン-固定化パラフィン埋め込み標本。Cancer MapとCamCapは独立しておらず、40の癌を共通して有していた。Cancer Mapデータセットのための臨床および分子の詳細は、補助情報表2および補足的データ1に与えられる。

c. DESENTとして指定された癌の間の遺伝子発現の平均値の相関。MSKCC、Cancer Map、StephensonおよびKleinデータセットのための6つの比較がすべて示される。各遺伝子の発現レベルはすべてのサンプルにわたり平均0および標準偏差1で正規化した。

d. MSKCC、Cancer MapおよびStephensonデータセットのためのKaplan-Meier PSA再発プロット。各グループ中の癌の数は各Kaplan-Meierプロットの右下隅に示される。PSA再発の患者の数は括弧の中で示される。

【0203】

図2：DESENT予後不良前立腺癌内で共通してダウンレギュレートした遺伝子

a. 非DESENT癌と比較してDESENT癌中で著しく変化した発現を示した遺伝子の数($P < 0.01$ 、再発発見比率(False Discovery Rate)について補正後)。100の独立したLPDの実験で少なくとも80でダウンレギュレートしたとの厳格な要求に基づいて、4つの発現マイクロアレイ・データセットすべてにおいて45の遺伝子がDESENT癌中でより低い発現をした。

10

20

30

40

50

b. 生物学的グループによる45の遺伝子のリスト。コード化された蛋白質の機能は、補助情報表3に示される。45の遺伝子のうちのいくつかは間質組織で優先的に発現されるが、我々は、細胞の組成物のデータが利用可能であるCancer MapおよびCamCapの患者のシリーズの両方において間質組織の含有量と臨床の結果の間には相関が無いことを見いだした。患者が2つのグループ（中央の間質含有量の上および下）へ階層化された時、Kaplan-Meierプロットは、Cancer Map（ログ-ランク検定、 $p = 0.159$ ）およびCamCap（ $p = 0.261$ ）の患者のシリーズの両方の結果の差を示さなかった。

c. circosプロットとして表わされる、前立腺癌のための公表された予後不良シグネチャーでの遺伝子と、人間の前立腺癌のためのDESENT分類の関係。45の共通してダウンレギュレートされる遺伝子へのリンクは茶色で示される。

【0204】

図3：The Cancer Genome Atlasデータセット中のRF-DESENTおよび非-RF-DESENT癌の比較

20の遺伝子のランダムフォレスト（RF）クラシファイヤーはDESENT癌（RF-DESENT癌と同定される）を同定するために使用された。遺伝子変化のタイプは各遺伝子（突然変異、融合、欠失および過剰発現）について示される。生化学的再発（BCR）を含む臨床のパラメーターは、iCluster、メチル化、ソマティックコピー数変化(somatic copy number alteration:SVNA)およびmRNA^{7,20}のグループと一緒に底部に表わされる。各遺伝子の突然変異およびホモ接合体欠失は、BRCA2（ $P = 0.021$ 、²テスト）およびTP53（ $P = 0.0038$ ）の中の遺伝子変化の過剰を含むRF-DESENT癌と組み合わせられた時、複数の試験について修正した後に、これらの相違は有意でなかった（ $P > 0.05$ ）。

【0205】

図5：ログ尤度プロット(Log-likelihood plots)

各データセットについてのログ尤度（縦軸）対プロセスの数（横軸）。MAP解決策（上のカーブ）および最大尤度解決策（下のカーブ）を使用した。最大尤度モデルについては、ログ尤度のピークは、使用するプロセスの数を示す。MAPモデルについては、ベイズ事前確率(Bayesian prior)が、過剰に複雑なモデルの罰則を構築するために使用される。ログ尤度は後にプラトーへ上昇し、その後それ以上のゲインはえられず、使用されるべきプロセスの最大数を示す。

【0206】

図6：トランスクリプトーム・データセットのラテントプロセスデコンポジション（LPD）分析

MSKCC、Stephenson、Cancer Map、CamCapおよびKleinデータセットは各々、それらのログ尤度プロット（図5）から示されたプロセスの最適な数へデコンポジションされた。単一のサンプルはすべてのプロセスにわたり表わされる。また、各棒の高さは、各LPDプロセスに帰することができるシグネチャーの割合（ p_i ）に相当する。サンプルは、それらが p_i の最高値を示すLPDグループに割り当てられる。MSKCC、Cancer MapおよびCamCapデータセットについて、赤、青、緑は、方法の中で定義されるような臨床のパラメーターに基づいた進行の異なる危険を有する癌を表示する。Stephensonデータセットについては、3つの危険グループへの指示に必要なパラメーターのうちのいくつかが見当たらないので、病理学的病期だけが示される。Kleinデータセットからの臨床データは公的に利用可能ではない。各データセットについては、DESENT癌を含むプロセスが示される。ログ尤度プロットとLPDは、その発現がMSKCCデータセットの中で最大に変化する500の遺伝子座を使用して行なわれた。

【0207】

図7：LPDによって同定されたDESENT癌についての結果の分析

(a-d) MSKCC (a)、Cancer Map (b)、Stephenson (c)

）およびCamCap(d)データセットについてのKaplan-Meier PSA再発プロット。各データセットについては、LPDによってDESENTプロセスに割り当てられた癌は、残りの癌と比較される。各グループ中の癌の数は各プロットの右下隅に示される。PSA再発の癌の数は括弧の中に示される。示されたKaplan-Meierプロットは、無作為に選ばれたシードパラメーターを使用して、各行なわれた、100のLPD実験からの最も頻繁な値(モード)p値を表わす(図12)。

(e-i) MSKCC(e)、CancerMap(f)およびStephenson(g)データセットのための方法に述べられているように、多変量分析が行なわれた。(h)については、多変量分析が、結合したMSKCC、CancerMapおよびStephensonデータセットについて行なわれた。

(i) 多変量分析をCamCapデータセット上で行なった。CamCapは、CancerMapデータセットの40の癌がオーバーラップしたために、別途分析された。方法に述べられているように、MSKCCとStephensonのデータセット用の病理学的病期共変量はコックス・モデルのその比例ハザード仮定を満足せず、時間従属変数としてモデル化された。

【0208】

図8：LPD分類によって同定されたDESENT癌の遺伝子発現の相関

癌の間の遺伝子発現の平均値の相関(corr.)を、MSKCC、CancerMap、Stephenson、KleinおよびCamCapデータセットの各々からのLPDを使用して、DESENTプロセスとの間で示した。その発現レベルがMSKCCデータセットの中で最大に変化し、LPDのために使用された、500の遺伝座からのデータが示される。平均0および標準偏差1で、各遺伝子の発現レベルがすべてのサンプルにわたり正規化された。10の可能な比較がすべて示される。

【0209】

図9：20の遺伝子シグネチャーを使用するRF分類によるDESENT癌の検知

ランダムフォレスト分類が、1669の遺伝子のlasso回帰分析(lasso regression analysis)の中の5つのデータセット；MSKCC、CancerMap、Stephenson、KleinおよびCamCapの少なくとも2つでDESENT癌の有意な変化された発現と同定された20の遺伝子のシグネチャーを使用して行なわれた。各データセットについて使用された参照は、図12に示されるモードp値に対応するDESENTグループの癌であった。図は各予測についてのAUC、正確さ、感度および特異性を示す。偽陽性(右上)・偽陰性(左下)割付の数を示すグリッドが、各データセットのために示される。

【0210】

図10：RF分類によって同定されたDESENT癌についての結果の分析

(a-e) MSKCC(a)、CancerMap(b)、Stephenson(c)、CamCap(d)およびTCGA(e)データセットのためのKaplan-Meier PSA再発プロット。各データセットについては、20の遺伝子RF分類を使用してDESENTに割り当てられた癌は、残りの癌と比較された。各グループ中の癌の数は各プロットの右下隅に示される。PSA再発癌の数は括弧の中に示される。MSKCC(f)、CancerMap(g)、Stephenson(h)、CamCap(i)およびTCGA(j)データセットのための方法に述べられているように、多変量分析が行なわれた。方法に述べられているように、MSKCCとStephensonのデータセットについての病理学的病期共変量はコックス・モデルの比例ハザード性(proportional hazards assumption)を満足せず、時間従属変数としてモデル化された。

【0211】

図11：RF分類によって同定されたDESENT癌の遺伝子発現の相関

癌の間の遺伝子発現の平均値の相関は、MSKCC、CancerMap、Stephenson、Klein、CamCapおよびTCGAデータセットの各々からのRF分類を使用して、DESENTプロセスに割り当てた。その発現レベルがMSKCCデータセ

10

20

30

40

50

ットの中で最大に変化し、LPDのために使用された500の座からのデータが示される。平均0および標準偏差1として、各遺伝子の発現レベルはすべてのサンプルにわたり正規化された。15の可能な比較がすべて示される。各データセットについては、LPDとRFによって同定されたDESENTプロセスの間に同様の相関が観察された（データは示されない）。

【0212】

図12：LPD検定の分布

100のLPDのPSA再発ログ・ランクp値の分布を、データセット(a)MSKCC、(b)CancerMap、(c)CamCapおよび(d)Stephensonについてランダムシードで再開した。モードログ・ランクp値に対応するKaplan-Meierプロットの例は、図1dおよび図7a-dに示される。

10

【0213】

図13：MSKCCデータセットのLPDデコンポジション。

(a) サンプルは8つのプロセスすべてで表わされる。また、各棒の高さは、各LPDプロセスに割り当てることができるシグネチャーの割合（ガンマ、垂直軸）に相当する。7番目の列は、各サンプルの中で同定されたDESENT発現シグネチャーのパーセンテージを示す。

(b) 各サンプルの中にあるDESENT癌の割合を示す棒グラフ。

(c, d) 個々の癌の組成を示す円グラフ。DESENTは赤である。他のLPDグループはキーの中で示されるような異なる色によって表わされる。円グラフの隣の数値は、それが上記の棒グラフからどの癌を表わすかを示す。DESENTシグネチャーが最も豊富なものだった時、個々の癌は「DESENT癌」として割り当てられた；例は左の箱（DESENT）の中で示される。他の多くの癌は少ない割合のDESENT癌(d)を含んでおり、さらに不良な結果を持つと予測される：例はより大きなボックス(c、いくつかのDESENT)中に示される。

20

【0214】

図14：存在するDESENT癌のパーセンテージに基づいた前立腺癌の層化

これらの分析については、MSKCC、CancerMap、CamCapおよびStephensonデータセットからのデータは組み合わせられた(n=517)。

(a) プロットは各癌へのDESENT癌の寄与を示し、4つのグループへ分けられた。グループ1のサンプルは0.1%未満のDESENT癌がある。

30

(b) LPDによって決定されるDESENT癌存在の割合に基づいて、Kaplan-Meierプロットは生化学的再発(BCR)の無い生存を示す。各グループ中の癌の数は右下に示される。各グループ中のPCR再発の数は括弧内に示される。グループ1-4の定義は図2aの中に示される。DESENT30%まで(グループ2)のガンマ値の癌は、DESENTを欠く癌(<0.1%)と比較して、より不良な臨床の結果(X2テスト、 $p=0.015$)を示した。中間物(0.3から0.6)および高い(>0.6)ガンマ値を有する癌は、さらにDESENTを欠く癌と比較して有意により悪い結果(それぞれ、 $P=2.69 \times 10^{-6}$ および $P=2.22 \times 10^{-14}$)を示した。結合したログ・ランクp値= 1.28×10^{-14} 。

40

【0215】

図15：LPDについて、ノモグラムモデルは1、3、5および7年でPSAの無い生存を予測するために開発された。

一人の患者のそれぞれの臨床の変数の評価は、対応するポイントスコア(トップスケール)を有する。各変数のポイント・スコアは加えられ、各患者について合計のポイント・スコアを生成する。1、3、5および7年でのPSAの無い生存の予測された可能性は、トップのポイント・スコアから確率目盛に鉛直線を下に引き下げることにより決定することができる。

【0216】

図16：LPDのためのコックス・モデル

50

- (a) 各共変量のHRのグラフ表示およびHRの95%の信頼区間。
 (b) HR、95%のCIおよびコックス・モデルのワルド検定統計量。
 (c) 1000のブートストラップリサンプルを使用するモノグラムの内部バリデーション(internal validation)のためのキャリブレーション・プロット。
 黒実線は、モノグラムの明白なパフォーマンスを表わす、青色線は、バイアス補正されたパフォーマンスを表わす、点線は、理想的なパフォーマンスを表わす。
 (d) CamCapデータセットを使用するモノグラムの外部バリデーション(external validation)のためのキャリブレーション・プロット。実線は観測されたパフォーマンスに相当し、点線は、理想的なパフォーマンスを表わす。

【0217】

10

実施例

実施例1

前立腺癌の臨床の管理での重大問題は、それが高度に異質であるということである^{1, 2}。したがって、個々の癌行動の正確な予測は診断の時に達成可能ではなく、本質的な過剰治療に結びつく^{3, 4}。それが乳癌⁵とは対照的に、グローバルな発現プロファイルのスーパーバイズしていない分析は、別個の臨床結果のロバストカテゴリー(robust categories)の前立腺癌を定義していないことは謎のままである^{6, 7}。現在の研究では、ラテントプロセスデコンポジション(Latent Process Decomposition⁸ (LPD))とも呼ばれるスーパーバイズしていないベイズ手順の適用が、4つの独立した前立腺癌トランスクリプトーム・データセット中の共通のプロセスを同定する。このプロセスに割り当てられた癌(DESNT癌と同定)は、サイトスケルトンマシーナリー(cytoskeleton machinery)、イオン輸送および細胞接着に含まれる多くのコード化蛋白質の45遺伝子のコアセットの低い発現により特徴づけられる。前立腺摘除に続くリンクしたPSA再発データを備えた3つのデータセットについては、DESNT癌の患者は、他の患者に比較して非常に不良な結果を示した($P = 2.65 \times 10^{-5}$ 、 $P = 4.28 \times 10^{-5}$ および $P = 2.98 \times 10^{-8}$)。ランダムフォレスト分類を使用したThe Cancer Genome Atlas内で注釈された前立腺癌の分析は、DESNT癌と、ETS遺伝子ステータスを含む任意の特別のクラスの遺伝子突然変異の存在の間のリンクを明らかにできなかった。我々の結果は、人間の前立腺癌の予後不良カテゴリーの存在を実証し、療法のターゲッティングを助け、インドレント疾病を持った人の治療に関連する罹患を回避するのを支援するだろう。

20

30

【0218】

前立腺癌のためのほとんどの発現に基づいた予後のシグネチャーは、侵襲性および非侵襲性の疾病の比較^{9, 10}、または特定の生物学的機能を表わす遺伝子¹¹⁻¹⁴の選択を含む、スーパーバイズされたステップを使用して導かれたものを共通に持っている。あるいは、発現バイオマーカーは体細胞のコピー数多型(somatic copy number variations: SCN)の存在にリンクされることがある⁷。latent Dirchelet allocation method¹⁵に基づいたLPDは、臨床結果か生物学的役割⁸についての知識がない状態でデータセットの構造を検討する。標準のスーパーバイズされていないクラスターリングモデル(例えばk-手段および階層的クラスタリング)とは対照的に、個々の癌は単一のクラスターに割り当てられない：代わりに、各癌の中の遺伝子発現レベルはラテントプロセスの組み合わせによってモデル化される。個々の癌の組成が非常に異種混合の場合^{16, 17}、および単一の標本が1を超える寄与系統(contributing lineage)¹⁸⁻²⁰を含んでいる場合、この種の分析は前立腺癌に特に適している。LPDは、乳癌データセットにベサル(basal)とERBB2の過剰発現サブグループの存在を確認し⁸、かつ重度の前立腺癌の患者が2つの臨床的に別個のカテゴリーへ階層化することができること²¹を示すために以前から使用されていた。

40

【0219】

前立腺摘除標本から得られた4つの独立したトランスクリプトーム・データセット(MSKCC⁶、CancerMap、Klein²²およびStephenson²³、図1b)が分析された。LPDは、ログ尤度プロット(図1b、図5)から示されるように、

50

全体的な発現プロフィールに寄与する、3から8の間の根本的なプロセスを使用して行なわれた。各データセットのデコンポジションに続いて、癌は、図1aおよび図6の中で示される結果を生成するそれらの最も高い p_i 値に基づいて個々の潜在プロセスに割り当てられた。 p_i は、個々の癌の発現プロフィールへの各プロセス i の寄与である：すべてのプロセスでの p_i の合計は1である。デコンポジション間での関係を求めて、遺伝子発現レベルの相関に基づき、4つのデータセットすべてにわたり共通するように見える1つのプロセスが同定された(図1c)。さらに各データセットについて、この集合体を調査するために、我々は、プロセスに割り当てられた癌と同じデータセットからの他のすべての癌と比較して、著しくより低くまたは高く発現した遺伝子を同定した(False Discovery Rateについて補正の後、 $p < 0.01$)。これは、より低い発現の45の遺伝子すべてで共有されるセットを明らかにした(図2a、拡張データ表1)。これらの45のコア遺伝子によってコード化された蛋白質の多くは細胞骨格のコンポーネントであるかまたはその力学を規制する。その一方で他のものは細胞接着およびイオン輸送に関係している(図2b)。45の遺伝子のうちの11は、前立腺癌について公表された予後のシグネチャーの要素だった(図2cおよび補足データ1)。例えば、MYLK、ACTG2およびCNN1は癌転移²⁴のためのシグネチャーでダウンレギュレートされる。一方、TMP2のより低い発現がOncotype DXシグネチャーの一部としてより不良な結果に関連している²⁵。この一般的なプロセスに割り当てられた癌は、「DESENT」(ラテン語でDESCENDUNT、下がる)と呼ばれる。

【0220】

MSKCC発現データセットに利用可能にリンクされた臨床データを使用して、発明者は、他のプロセスに割り当てられた患者と比較された時、DESENT癌の患者が非常に不良な結果を示すことを見いだした($P = 2.65 \times 10^{-5}$ 、ログランク検定、図1d)。前立腺摘除に続くPSA再発データが利用可能であることがさらなる2つのデータセットで確認された(図1d)：DESENT癌を有するStephensonおよびCancerMapデータセット患者の両方について、非常に不良な結果が示された(それぞれ、 $P = 4.28 \times 10^{-5}$ および $P = 2.98 \times 10^{-8}$)。Gleason sumを含む多変量解析において、DESENT癌とされたStageおよびPSAはStephensonおよびCancerMapデータセットにおける不良な結果の独立予測因子だった($P = 1.83 \times 10^{-4}$ および $P = 3.66 \times 10^{-3}$, Cox regression model)が、MSKCCデータセットではそうではなかった($P = 0.327$) (表8、図7)。3つのデータセットが組み合わされた時、DESENTメンバーシップの独立した予測値は、 $P = 1.61 \times 10^{-7}$ (図7)であり、Gleason Sumでは $P = 1.00 \times 10^{-5}$ だった。予後不良DESENTプロセスは、CamCap dataset⁷ (表8、図7および8)の中で同定された。それは独立していなかったため、上記の分析から除外された：CancerMapに含まれた癌との本質的なオーバーラップがあった(図1b)。

【0221】

発明者は、LPDと異なり、コンピューター処理集約的でなく、データセットのより広い範囲および個々の癌の両方に適用することができたクラシファイヤーの開発を望んでいた。少なくとも2つのデータセット中でDESENTおよび非DESENT癌の間で著しく変化した発現を有する1669の遺伝子が、分析に選ばれた。LASSOロジスティック回帰モデル(LASSO logistic regression model)は20遺伝子のセットの選択に導くMSKCCデータセット内のDESENTメンバーシップの最良の予測である遺伝子を同定するために使用された(拡張データテーブル2)。それはDESENT癌の中で有意により低い発現を備えた45の遺伝子に対して1つの遺伝子オーバーラップ(ACTG2)を持っていた。ランダムフォレスト(RF)分類を使用して、これらの20の遺伝子が、個々の癌がDESENTであるとの予測について高い特異性および感度を、MSKCCトレーニング・データセット、および3つの確認データセット中で有するとされた(図9)。PSA再発データにリンクされた2つの確認データセット(StephensonとCancerMap)については、予測された癌サブグループは、一変量および多変量分析の両方でより不良な臨床の結果を

示し、これはLPDを使用して観察された結果と合致した(表8、図10)。RF分類がThe Cancer Genome Atlas (TCGA)²⁰によって注釈された(annotated)333の前立腺癌からのRNAseqデータに適用された時、以下のものに基づいてDESENTとして確認された患者のサブグループが同定された:

- (i) 他のデータセット中のDESENT癌グループとの遺伝子発現レベルの相関(図11);
- (ii) コアダウンレギュレートした遺伝子セット(45/45の遺伝子)でのDESENTおよび非DESENT癌の間での差別的に発現した遺伝子のオーバーラップのデモンストラーション;および
- (iii) 非DESENT患者と比較してより不良な臨床の結果($P = 5.4 \times 10^{-4}$)(表8、図10e)。

10

【0222】

TCGAデータセットについての、我々は任意の特定の遺伝子変化の存在と、DESENT癌としての割付との間の相関を見つけられなかった(False Discovery Rate、²テストについての補正後 $P > 0.05$ 、図3)。特別の留意点として、ETS遺伝子ステータスに相関はなかった($P = 0.136$ 、²テスト、図3)。蛍光インシティハイブリダイゼーション ブレーク-アパート分析(the fluorescence in situ hybridization break-apart assay)²⁶を使用して決定されたDESENT癌とERG遺伝子転位の間の相関の欠如は、Cancer Mapサンプル(LP D - DESENT、 $P = 0.549$; RF - DESENT、 $P = 0.2623$ 、²テスト:それぞれLP D - DESENTおよびRF - DESENTと呼ばれる、LP DおよびRFアプローチによって同定されたDESENT癌)を使用して確認された。これらの観察は、ERGステータスと臨床結果の間の相関の不足と一致している²⁷。前立腺癌のほぼ半分で見つかったETS遺伝子変化²⁰、²⁶は、前立腺癌の進行の初期のステップであると考えられる¹⁷、²⁸ので、DESENT癌の生成に関係する変化は、ETS陽性と、ETS陰性癌に共通する後の出来事を表わす。

20

【0223】

TGCAシリーズ中のRF - DESENT癌については、45のコア遺伝子のうちのいくつかは、制御遺伝子発現における可能な役割を示唆する、非-RF - DESENT癌(補助情報表1)と比較して、CpG遺伝子メチル化の変更されたレベルを示した。このアイデアを支援するものとして、45のコア遺伝子のうちの16については、ヒト癌中の後成のダウンレギュレーションは、前立腺癌中の6つの遺伝子を含めて、以前に報告されている(CLU、DPYSL3、GSTP1、KCNMA1、SNAI2およびSVIL)(図2b、拡張データ表1)。5つの遺伝子(FBLN1、GPX3、GSTP1、KCNMA1、TIMP3)のCpGメチル化は、以前に癌攻撃性にリンクされた。細胞骨格構造(cytoskeleton structure)を決定し、DESENT癌の中の細胞接着に含まれる遺伝子のダウンレギュレーションは、癌攻撃性を決定する際にアメーバ状のタイプの移動および間充織の移動の寄与に反対するだろうが、集合的な移動または膨張性増殖型(expansive growth phenotypes)を反映することができるだろう²⁹。

30

【0224】

前立腺癌についてのThe European Randomized研究からの証拠は、PSAスクリーニングが前立腺癌からの死亡率を21%引き下げることができることを実証する³⁰。しかしながら、PSAスクリーニングについての重大問題は、それが臨床的に無関係である癌の50%以内の検知に結びつくということである³、⁴:それは人のスクリーニングがない状態での一生の症状を引き起こさなかったであろう癌である。前立腺癌トランスクリプトーム・データセットへのLPDの適用に関する我々の研究は、すべての試験された前立腺摘除シリーズにわたる共通の新しい予後不良カテゴリーの前立腺癌の存在を明らかにした。いくつかの異なるプラットフォーム(Illumina HT12 v4 BeadChip array、RNAseq、Affymetrix array)によって生成されたデータを使用して、冷凍された組織とホルマリ

40

50

れはより攻撃的な疾患を有する人への放射線療法、手術および化学療法のターゲッティングを可能とし、無関係の癌の人の中で、インポテンスを含む治療の副作用を回避するのを支援するだろう。

【0225】

方法

Cancer Map データセット

新鮮な前立腺癌標本は、Royal Marsden NHS Foundation Trust and Addenbrooke's Hospital, Cambridgeで前立腺摘除を経験した系統的な一連の患者から得られた。適切な地方の研究倫理委員会(Research Ethics Committee)の許諾が、この研究について得られた。冷凍された前立腺スライスを集め³¹、RNAが以前に記述されるように調製された^{7, 32}。

10

【0226】

発現プロファイルは、メーカーの指示によって1.0 Human Exon ST (Affymetrix, Santa Clara, CA, USA)を使用して、以前に記述された³²ように決定された。Affymetrix GeneChip (登録商標)のWhole Transcript Sense Target Labelling Assay が使用され、バイアスなしで発現されたゲノム全体(RNAの合計の1.5 μg)から増幅されビオチン化されたセンス-ストランドDNA目標(biotinylated sense-strand DNA targets)を生成した。メーカーの指示に、ハイブリダイゼーション、洗浄および走査のステップについて従った。アレイは、16時間、45でAffymetrix Gene Chipハイブリダイゼーション・オープンの中で60 rpmでそれらを回転させることによりハイブリダイズされた。ハイブリダイゼーションの後、アレイはAffymetrix GeneChip Fluidics station FS 450の中で洗われた。アレイはAffymetrix Gene Chip scanner 3000 7Gシステムを使用して走査された。データはGene Expression Omnibus GSE (出版でリリースされるデータ)から利用可能である。

20

【0227】

進行カテゴリーのリスク

進行カテゴリーの前立腺摘除リスクが、前立腺癌(Chris Foster、パーソナルコミュニケーション)についての英国の国際的な癌ゲノム・コンソーシアム層化(the UK International Cancer Genome Consortium stratification)に基づいて定義された。

【0228】

30

【表1】

低いリスク	PSA ≤ 10 ナノグラム/ml AND (グリーソン=3+3 OR (グリーソン=3+4 AND 余分な皮膜の伸展はない(no extra capsular extension)))
中程度のリスク	10 ナノグラム/ml < PSA ≤ 20 ナノグラム/ml OR (グリーソン=4+3 AND 余分な皮膜の伸展はない) OR (グリーソン=3+4 AND 余分な皮膜の伸展)
高いリスク	PSA > 20 ナノグラム/ml OR グリーソン合計 > 7 OR (グリーソン=4+3 AND 余分な皮膜の伸展) OR 精囊腺浸潤

【0229】

40

追加トランスクリプトーム・データセット

MSKCC、Cancer Map、CamCap、StephensonおよびKleinと呼ばれる5つの前立腺癌マイクロアレイ・データセットが分析された。分析されたすべてのデータは根治的前立腺摘除標本からだった。MSKCCデータセットは370のAffymetrix Human Exon 1.0 ST Array 実験(GEO: GSE21034)を含む⁶。50のマイクロアレイは細胞系統、異種移植片および転移性の組織に対応して除去された。残りの320のマイクロアレイは、原発性の腫瘍および正常組織サンプルからの160の複製を表わす：各サンプルからの1つのデータセットだけがLPD分析で使用された。Stephensonデータセットは、Affymetrix U133A ヒトの遺伝子アレイ²³を使用して得られた78の癌および11の正常な前立腺サンプルからのデータを含んで

50

いる。Kleinは、Affymetrix Human Exon 1.0 ST Arrays (GEO: GSE62667)²²で分析された、182のホルマリン固定され、パラフィン埋込(FFPF)された原発性の腫瘍サンプルから成る。我々の研究で使用されるCamCapデータセットは、2つの前立腺摘除シリーズ(CambridgeとStockholm)から得られたIllumina HumanHT-12 V4.0発現ビーズチップ(ビーズ・マイクロアレイ)データセット(GEO: GSE70768とGSE70769)を組み合わせで生成され、147の癌および73の正常なサンプルから成った⁷。CamCapとCancerMapのデータセットは40人の患者に共通し、したがって独立していない。The Cancer Genome Atlasからの333の前立腺癌から成る1つのRNAseqデータセットが分析され、TCGA²⁰と呼ばれた。TCGAによって供給された遺伝子あたりのカウントが使用された。

10

【0230】

データ処理

遺伝子レベルおよびエキソン・レベル(exon-level)発現信号評価は、Affymetrix発現コンソール・ソフトウェア・パッケージ(Affymetrix Expression Console software package)の中でインプリメントされたロバスト多重配列分析アルゴリズム(robust multiarray analysis algorithm)³³を使用して、Affymetrix GeneChip Exon 1.0 STアレイから生成されたCELファイルから導かれた。ビーズマイクロアレイデータセットについては、あらかじめ正規化されたデータが使用されて、illuminaHumanv4.db Rアノテーションパッケージを使用して、UCSC hg¹⁹に注釈された(annotated)不良な品質プローブ(「悪い」および「マッチなし」プローブ)が除去された。あらかじめ正規化されたStephensonデータセットはhgu133a.db Rパッケージを使用して注釈された。必要な時、データセット/コアバッチ効果は、sva Rパッケージの中でインプリメントされたComBat アルゴリズム³⁴の使用のために調節された。

20

【0231】

Latent Process Decomposition

Latent Process Decomposition(LPD)^{8, 35}、スーパーバイズされていないベイズアプローチが、サンプルをプロセスと呼ばれるサブグループに分類するために使用された。Rogersら³⁵の中でのように、LPDでの使用のためにMSKCCデータセットにわたる最も大きな変異を備えた500 プローブセットが選ばれた。これらのプローブセットは492の遺伝子をマップする。各データセットのために、これらの遺伝子をマップするすべてのプローブがLPD分析(CancerMap: 507、CamCap: 483、Stephenson: 609)の中で使用された。

30

【0232】

LPDは客観的にプロセスの中で最も有望な数进行评估することができる。発明者は、プロセスの様々な数で見積もられたデータのホールドアウトバリデーションログ尤度(hold-out validation log-likelihood)を評価し、またプロセスの数を選ぶために均一(最大尤度アプローチと等価)および不均一(MAPアプローチ)の組み合わせを使用した。MAPアプローチのために、モデルの平均パラメーターが先に観察されたように0.1にセットされた。この使用された値は結果にほとんど影響を与えず、最大のログ尤度に相当する先の値に変化パラメーターがセットされた、つまりMSKCCには-0.5、CancerMapには-0.5、CamCapには-0.05、Stephensonには-0.75、Kleinには-0.3。強健さ(robustness)のために、発明者は、各データセットのために、異なるシードでLPDを100回行った。100回の実施から、発明者は、後の分析に使用される代表的な検定を選択した。代表的な検定は、モードに近い生存ログ-ランクp値での検定だった。Kleinデータセットについては、利用可能な臨床データを持っていなかったため、LPDからのホールドアウトログ尤度が代わりに使用された。

40

【0233】

統計的検定

統計的検定はすべてRバージョン3.2.2(<https://www.r-project.org/>)により行

50

なわれた。特定の遺伝子セットとサンプル・サブグループの2つのデータセット間の発現プロファイル間の相関は、以下のように計算された：

- 1．各遺伝子については、1つのプローブセットは任意に選択されている；
- 2．各プローブセットについては、その分布は標準正規分布にすべてのサンプルにわたり変形される；
- 3．サブグループ中のサンプルにわたる各プローブセットの平均の発現が決定され、サブグループについての発現プロファイルを得た；および
- 4．2つのデータセット中のサブグループの発現プロファイル間のピアソンの相関が決定される。

【0234】

差別的に発現されたプローブセットは、limma R パッケージ³⁶に実装された(implemented)緩和されたt検定(moderated t-test)を使用して同定された。遺伝子は、調整されたp値が0.01未満(フォルスディスカバリーレート(False Discovery Rate))を使用して調整されたp値)だった場合、著しく差別的に発現すると考えられる。

【0235】

生存率分析は前立腺摘除の後の生化学的再発をエンドポイントとして、コックス比例ハザード・モデル(Cox proportional hazards models)およびKaplan-Meier 評価(Kaplan-Meier estimator)を使用して行なわれた。1人の患者当たりいくつかのサンプルが利用可能だった時、腫瘍組織の最も高い割合を備えたサンプルだけが使用された。正常組織からの発現プロファイルは含まれなかった。多変数の生存率分析は、臨床の共変量グリーソン・グレード(7および>7)、病理学的病期(T1/T2およびT3/T4)およびPSAレベル(10および>10)で行なわれた。発明者は、ヘヴィサイド関数(heavyside function)での変数の生成物として、比例ハザード仮定(MSKCC 中のT-病期)を満たさなかった変数をモデル化した：

【0236】

【数1】

$$g(t) = \begin{cases} 1, & t \geq t_0 \text{ の場合} \\ 0, & \text{それ以外} \end{cases}$$

【0237】

ここで t_0 は時間閾値である。ヘヴィサイド関数での予測の増加は、予測を時間区間に分割し、拡張コックス・モデルがそれについて異なる危険比率を計算した。

【0238】

DESENTメンバーシップの最適な予測の算出

DESENTメンバーシップの最適な予測を引き出すために、それらが比較可能なように、データセットが調製された：関連する遺伝子がすべてのマイクロアレイ・プラットフォームで見つけられた場合のみプローブは保持された。1つの無作為に選ばれたプローブだけが、1つの遺伝子当たり保持された。また、バッチ効果はComBat アルゴリズム³⁴を使用して調節した。MSKCCデータセットは、トレーニング・セットとして使用され、他のデータセットはテスト・セットとして使用された。遺伝子選択は、glmnet R パッケージ³⁷の中で実装された、正則化された一般的な線形モデル・アプローチ(LASSO)を使用して、5つのマイクロアレイ・データセット(1669の遺伝子)の合計の少なくとも2つのDESENTの中で、著しくアップまたはダウンレギュレートされたすべての遺伝子から始めて行なわれた。LASSOは、100回実行された。また、実行の少なくとも25%で選択された遺伝子だけが保持された。その後、最適な予測は、ランダムフォレスト R パッケージ³⁹の中で実装されたランダムフォレストモデル³⁸を使用して導き出された。デフォルトパラメーターが使用され、ツリーの数10001にセットされた。クラスサイズの不均衡は、大多数のクラスのサンプリングを少数のクラスの

10

20

30

40

50

度数に減らすことにより調節された。

【0239】

実施例2

DESENTシグネチャーの存在は不良な臨床結果を予測する。従来の研究では、発現マイクロアレイ・データセットの最適なデコンポジションは3から8の基本的なプロセスを使用して行なわれた。8つのプロセスへのMSKCCデータセットのデコンポジションの例は、図13aに示され、プロセスがそれぞれ棒グラフによって表わされる。サンプルは8つのプロセスすべてで表わされる。また、各棒の高さは、各LPDプロセスに割り当てることができるシグネチャーの割合（または p_i ）に相当する。LPDプロセス7は、各サンプルの中で同定されたDESENT発現シグネチャーのパーセンテージを示し、図13bおよび13dの中で示されるように、DESENTシグネチャーが最も豊富だった時、個々の癌が「DESENT癌」として割り当てられる。PSA再発について、DESENT癌の患者は、同じデータセット中の他の癌に比べて、常により不良な結果を示した。含意は、それが不良な結果を与えるDESENTシグネチャーを含んでいる癌の部位の存在であるということである。このモデルは、MSKCCデータセットのために図13cで示されるもののような、DESENTシグネチャーのより小さな寄与を含んでいる癌がさらにより不良な結果を示すに違いないと予測する。

【0240】

この予測をテストする力を高めるため、MSKCC、CancerMap、StephensonおよびCamCapからの癌からのこのデータが組み合わせられた（ $n = 515$ ）。連続変数としてDESENTプロセスに割り当てられた発現の割合（）を処理し、発明者は、それがPSA再発（ $P = 2.66 \times 10^{-15}$ 、 $HR = 1.5$ 、95% CI = [1.35, 1.66]、コックス比例ハザード回帰モデル）と重要な関連を持つことを見いだした。が増加するとともに、結果はより悪くなった。これは、癌を存在するDESENTプロセスの割合に基づいた4つのグループに分割することにより説明され（図14a）、PSA再発のない生存は、以下のとおり（図14b）である；

(i) DESENT癌無し、70か月で74.4%。

(ii) 0.3未満の、70か月で63.1%；

(iii) 0.3から0.6の、70か月で45.5%、および(iv) > 0.60 の、70か月で20.4%（図14b）。

全体的に癌の47%は少なくとも幾分のDESENT癌を含んでいた（図14a）。

【0241】

PSA再発を予測するDESENTノモグラム

DESENT癌の割合は、Cox比例ハザードモデルで他の臨床の変数（グリーソン・グレード、PSAレベル、病理学的病期および外科的マージンステータス）と組み合わせられ、330の癌のコンバイン・データセットとフィッティングされた。DESENTは、グリーソン・グレード = 4 + 3（ $P = 1.8 \times 10^{-3}$ 、 $HR = 3.26$ 、95% CI = [1.55, 6.86]）、およびグリーソン・グレード > 7 （ $P < 1 \times 10^{-4}$ 、 $HR = 5.41$ 、95% CI = [2.46, 11.92]）、病理学的病期（ $P = 2.45 \times 10^{-2}$ 、 $HR = 1.62$ 、95% CI = [1.06, 2.48]）、陽性の外科的マージン（ $P = 1.74 \times 10^{-2}$ 、 $HR = 1.69$ 、95% CI = [1.10, 2.60]）と共に、より悪い臨床の結果（ $P = 3 \times 10^{-4}$ 、 $HR = 1.30$ 、95% CI = [1.13, 1.50]、図16a、b）の独立予測因子だった。予測としてのPSAレベルは、統計的有意差（ $P = 0.1145$ 、 $HR = 1.13$ 、95% CI = [0.97, 1.32]）の我々の閾値未満だった。この生存モデルを使用して、他の変数とともにDESENT癌への使用のための、前立腺摘除後の、1、3、5および7年での生化学的再発のリスクを予測するノモグラムが考案された（図15、図16）。内的妥当性(internal validation)では、ノモグラムはブートストラップ-補正された0.761のC指標、外的妥当性(external validation)ではCamCapデータセットで、0.799のC指標を得た。

【0242】

10

20

30

40

50

【表 2 - 1】

表 1:MSKCCデータセットのほとんどにわたり発現が異なる500の遺伝子プローブ

HGNC シンボル	アセッションナンバー	MME	NM_007288	C7	NM_000587
TGM4	NM_003241	RBPM5	L17325	NTN4	NM_021229
RLN1	NM_006911	HLA-DRB1	NM_002124	FAM36A	NM_198076
ORM1	NM_000607	FOLH1	NM_001193471	CNTNAP2	NM_014141
OLFM4	NM_006418	LUZP2	NM_001009909	SC4MOL	NM_006745
OR51E2	NM_030774	MSMB	NM_002443	CH17- 189H20.1	AK000992
SERPINB11	NM_080475	GSTT1	NM_000853		ENST000004270 89
CRISP3	NM_006061	MMP7	NM_002423	TRGC2	NM_015646
TDRD1	NM_198795	ODZ1	NM_001163278	RAP1B	NM_01098484
SLC14A1	NM_001128588	ACTB	NM_001101	SLC4A4	NM_178430
IGJ	NM_144646	SPON2	NM_012445	LCE2D	NM_001964
ERG	NM_001136154	SLC38A11	NM_173512	EGR1	NR_001447
GDEP	NR_026555	FOS	NM_005252	MT1L	SCUBE2
TMEFF2	NM_016192	OR51T1	NM_001004759	FAM55D	NM_001077639
CST1	NM_001898	HLA-DMB	NM_002118	PDK4	NM_002612
LTF	NM_002343	KRT15	NM_002275	CXCL13	NM_006419
AMACR	NM_014324	ITGA8	NM_003638	CACNA1D	NM_000720
SERPINA3	NM_001085	CXADR	NM_001338	GPR160	NM_014373
NEFH	NM_021076	LYZ	NM_000239	CPM	NM_001874
ACSM1	NM_052956	CEACAM20	NM_001102597	PTGS2	NM_000963
OR51E1	NM_152430	C8orf4	NM_020130	TSPAN8	NM_004616
MT1G	NM_005950	DPP4	NM_001935	BMP5	NM_021073
ANKRD36B	NM_025190	PGC	NM_002630	GOLGA8A	NR_027409
LOC100510 059	XM_003120411	C15orf21	NR_022014	OR4N2	NM_001004723
PLA2G2A	NM_000300	CHORDC1	NM_012124	FAM135A	NM_001105531
TARP	NM_001003799	LRRN1	NM_020873	DYNLL1	NM_001037494
REXO1L1	NM_172239	MT1M	NM_176870	DSC3	NM_024423
ANPEP	NM_001150	EPHA6	NM_001080448	C4orf3	NM_001001701
HLA-DRB5	NM_002125	PDE11A	NM_001077197	HIST1H2BK	NM_080593
PLA2G7	NM_001168357	TMSB15A	NM_021992	LCN2	NM_005564
NCAPD3	NM_015261	LYPLA1	NM_006330	STEAP4	NM_024636
OR51F2	NM_001004753	FOSB	NM_006732	RPS27L	NM_015920
SPINK1	NM_003122	F5	NM_000130	TRPM8	NM_024080
RCN1	NM_002901	C15orf48	NM_032413	ID2	NM_002166
CP	NM_000096	MIPEP	NM_005932	LUM	NM_002345
SMU1	NM_018225	HSD17B6	NM_003725	EDNRB	NM_001122659
ACTC1	NM_005159	SLP1	NM_003064	PGM5	NM_021965
AGR2	NM_006408	CD38	NM_001775	SFRP4	NM_003014
SLC26A4	NM_000441	MMP23B	NM_006983	STEAP1	NM_012449
IGKC	BC032451	OR51A7	NM_001004749	FADS2	NM_004265
MYBPC1	NM_002465	CFB	NM_001710	CXCL11	NM_005409
NPY	NM_000905	CCL2	NM_002982	CWH43	NM_025087
PI15	NM_015886	POTEM	NM_001145442	SNRPN	BC043194
SLC22A3	NM_021977	TPMT	NM_000367	GPR110	NM_153840
PIGR	NM_002644	FAM3B	NM_058186	THBS1	NM_003246
		FLRT3	NM_198391		

【 0 2 4 3 】

10

20

30

40

50

【表 2 - 2】

表 1 の続き

APOD	NM_001647	ATP8A2	NM_016529	SPOCK1	NM_004598
HPGD	NM_000860	PRIM2	NM_000947	GSTP1	NM_000852
LEPREL1	NM_018192	ADAMTSL1	NM_001040272	OAT	NM_000274
LCE1D	NM_178352	NELL2	NM_001145108	HIST2H2BF	NM_001024599
GSTM5	NM_000851	RPS4Y1	NM_001008	ACSM3	NM_005622
SLC30A4	NM_013309	CD24	NM_013230	GLB1L3	NM_001080407
SEMA3D	NM_152754	GOLGA6L9	NM_198181	SLC5A1	NM_000343
CACNA2D1	NM_000722	ZFP36	NM_003407	OR4N4	NM_001005241
GPR116	NM_015234	TRIB1	NM_025195	MAOB	NM_000898
C7orf63	NM_001039706	BNIP3	NM_004052	BZW1	NM_014670
FAM198B	NM_001128424	KL	NM_004795	GENSCAN0000007309	GENSCAN0000007309
SCD	NM_005063	PDE5A	NM_001083	IFI44L	NM_006820
NR4A2	NM_006186	DCN	NM_001920	KRT5	NM_000424
ARG2	NM_001172	LDHB	NM_001174097	SCN7A	NM_002976
ZNF385B	NM_152520	PCDHB5	NM_015669	GOLM1	NM_016548
RGS1	NM_002922	ACADL	NM_001608	HIST4H4	NM_175054
DNAH5	NM_001369	ZNF99	NM_001080409	IL7R	NM_002185
NPR3	NM_000908	CPNE4	NM_130808	CSGALNACT1	NM_018371
RAB3B	NM_002867	CCDC144B	NR_036647	A2M	NM_000014
CHRD1	NM_145234	SLC26A2	NM_000112	LRRC9	AK128037
ZNF208	NM_007153	CYP1B1	NM_000104	ARHGEF38	NM_017700
MBOAT2	NM_138799	SELE	NM_000450	ACSL5	NM_016234
ATF3	NM_001040619	CLDN1	NM_021101	SGK1	NM_001143676
ST6GAL1	NM_173216	KRT13	NM_153490	TMEM45B	NM_138788
GDF15	NM_004864	SFRP2	NM_003013	AHNAK2	NM_138420
ANXA1	NM_000700	SLC25A33	NM_032315	NEDD8	NM_006156
FOLH1	NM_004476	HSD17B11	NM_016245	GREB1	NM_014668
C4B	NM_001002029	HSD17B13	NM_178135	UBQLN4	NM_020131
ELOVL2	NM_017770	UGT2B4	NM_021139	SDHC	NM_003001
GSTM1	NM_000561	CTGF	NM_001901	TCEAL2	NM_080390
GLIPIR1	NM_006851	SCIN	NM_001112706	SLC18A2	NM_003054
C3	NM_000064	C10orf81	NM_001193434	HIST1H2BE	NM_003523
MYO6	NM_004999	CYR61	NM_001554	RARRES1	NM_206963
ORM2	NM_000608	PRUNE2	NM_015225	PLN	NM_002667
RAET1L	NM_130900	IFI6	NM_002038	OGN	NM_033014
PCDHB3	NM_018937	MYH11	NM_022844	GPR110	NM_025048
C1orf150	ENST000000366488	PPP1R3C	NM_005398	CLGN	NM_001130675
ALOX15B	NM_001141	KCNH8	NM_144633	NIPAL3	NM_020448
LSAMP	NM_002338	ZNF615	NM_198480	ACTG2	NM_001615
SLC15A2	NM_021082	ERV3	NM_001007253	RCAN3	NM_013441
PCP4	NM_006198	F3	NM_001993	KLK11	NM_001167605
MCCC2	NM_022132	TTN	NM_133378	HMGCS2	NM_005518
GCNT1	NM_001097634	LYRM5	NM_001001660	EML5	NM_183387
C5orf23	BC022250	FMOD	NM_002023	EDIL3	NM_005711
SCGB1D2	NM_006551	NEXN	NM_144573	PIGH	NM_004569
CXCL2	NM_002089	IL28A	NM_172138	GLYATL1	NM_080661
AFF3	NM_001025108	FHL1	NM_001159702		
		CXCL10	NM_001565		

10

20

30

40

【 0 2 4 4 】

50

【表 2 - 3】

表 1 の続き

ATP1B1	NM_001677	CXCR4	NM_001008540	FGFR2	NM_000141
GJA1	NM_000165	OR51L1	NM_001004755	SNAI2	NM_003068
PLA1A	NM_015900	SLC12A2	NM_001046	CALCRL	NM_005795
MPPED2	NM_001584	AGAP11	NM_133447	MON1B	NM_014940
AMD1	NM_001634	SLC27A2	NM_003645	PVRL3	NM_015480
EMP1	NM_001423	AZGP1	NM_001185	VGLL3	NM_016206
PRR16	NM_016644	VCAN	NM_004385	SULF1	NM_001128205
CNN1	NM_001299	ERAP2	NM_022350	LIFR	NM_002310
GHR	NM_000163	KRT17	NM_000422	SH3RF1	AB062480
ALDH1A1	NM_000689	SLC2A12	NM_145176	C12orf75	NM_001145199
TRIM29	NM_012101	CCL4	NM_002984	GNPTAB	NM_024312
IFNA17	NM_021268	RPF2	NM_032194	CALM2	NM_001743
TAS2R4	NM_016944	SLC45A3	NM_033102	KLF6	NM_001300
SEPP1	NM_001093726	SEC11C	NM_033280	C7orf58	NM_024913
GREM1	NM_013372	IFIT1	NM_001548	RDH11	NM_016026
RASD1	NM_016084	PAK1IP1	NM_017906	NR4A1	NM_002135
C1S	NM_201442	HIST1H3C	NM_003531	RWDD4	NM_152682
CLSTN2	NM_022131	ERRFI1	NM_018948	ABCC4	NM_005845
DMXL1	NM_005509	ADAMTS1	NM_006988	ZNF91	NM_003430
HIST1H2BC	NM_003526	TRIM36	NM_018700	GABRE	NM_004961
NRG4	NM_138573	FLNA	NM_001456	SLC16A1	NM_001166496
ARL17A	NM_001113738	CCND2	NM_001759	DEGS1	NM_003676
GRPR	NM_005314	IFIT3	NM_001031683	CLDN8	NM_199328
PART1	NR_024617	FN1	NM_212482	HAS2	NM_005328
CYP3A5	NR_033807	PRY	NM_004676	ODC1	NM_002539
KCNC2	NM_139136	HSPB8	NM_014365	REEP3	NM_001001330
SERPINE1	NM_000602	CD177	NM_020406	LYRM4	AF258559
SLC6A14	NM_007231	TP63	NM_003722	PPFIA2	NM_003625
EIF4A1	NM_001416	IFI44	NM_006417	PGM3	NM_015599
MYOF	NM_013451	COL12A1	NM_004370	ZDHHC8P1	NR_003950
PHOSPHO2	NM_001008489	EDNRA	NM_001957	C6orf72	AY358952
GCNT2	NM_145649	PCDHB2	NM_018936	HIST1H2BD	NM_138720
AOX1	NM_001159	HLA-DRA	NM_019111	TES	NM_015641
CCDC80	NM_199511	TUBA3E	NM_207312	PDE8B	NM_003719
ATP2B4	NM_001001396	ASPN	NM_017680	DNAJB4	NM_007034
UGDH	NM_003359	FAM127A	NM_001078171	RGS5	NM_003617
GSTM2	NM_000848	DMD	NM_000109	EPHA3	NM_005233
MEIS2	NM_172316	DHRS7	NM_016029	COX7A2	NR_029466
RGS2	NM_002923	ANO7	NM_001001891	MT1H	NM_005951
PRKG2	NM_006259	MEIS1	NM_002398	HIST2H2BE	NM_003528
FIBIN	NM_203371	TSPAN1	NM_005727	TGFB3	NM_003239
FDXACB1	NM_138378	CNTN1	NM_001843	VEGFA	NM_001025366
SOD2	NM_001024465	TRIM22	NM_006074	CRISPLD2	NM_031476
SEPT7	NM_001788	GSTA2	NM_000846	TFF1	NM_003225
PTPRC	NM_002838	SORBS1	NM_001034954	LOC100128 816	AY358109
GABRP	NM_014211	GPR81	NM_032554	SYT1	NM_001135805
CBWD3	NM_201453	CSRP1	NM_004078	CPE	NM_001873
TOR1AIP2	NM_022347	C3orf14	AF236158		

【 0 2 4 5 】

10

20

30

40

50

【表 2 - 4】

表 1 の続き

TRPC4	NM_016179	TPM2	NM_003289	LOC286161	AK091672
RAB27A	NM_004580	REPS2	NM_004726	NAALADL2	NM_207015
CD69	NM_001781	EA2F	NM_018456	TMPRSS2	NM_001135099
RPL17	NM_000985	CAV1	NM_001172895	SERPINF1	NM_002615
PSCA	NM_005672	PRUNE2	NM_015225	EPHA7	NM_004440
ATRNL1	NM_207303	TMEM178	NM_152390	SDAD1	NM_018115
MYOCD	NM_001146312	MFAP4	NM_001198695	SOX14	NM_004189
MS4A8B	NM_031457	SYNM	NM_145728	RPL35	NM_007209
TNS1	NM_022648	EFEMP1	NM_004105	HSPA1B	NM_005346
BAMBI	NM_012342	RND3	NM_005168	MSN	NM_002444
IGF1	NM_001111283	SCNN1A	NM_001038	MTRF1L	NM_019041
RALGAPA1	NM_014990	B3GNT5	NM_032047	PTN	NM_002825
S100A10	NM_002966	LMOD1	NM_012134	CAMKK2	NM_006549
PMS2CL	NR_002217	UBC	NM_021009	RBM7	NM_016090
MMP2	NM_004530	LMO3	NM_018640	OR52H1	NM_001005289
SLC8A1	NM_021097	LOX	NM_002317	C1R	NM_001733
OAS2	NM_002535	NFIL3	NM_005384	CHRNA2	NM_000742
ARRDC3	NM_020801	C11orf92	NR_034154	MRPL41	NM_032477
AMY2B	NM_020978	C11orf48	NM_024099	PROM1	NM_001145847
SPARCL1	NM_001128310	BCAP29	NM_018844	LPAR6	NM_005767
IQGAP2	NM_006633	EPCAM	NM_002354	SAMHD1	NM_015474
ACAD8	NM_014384	PTGDS	NM_000954	SCNN1G	NM_001039
LPAR3	NM_012152	ASB5	NM_080874	DNAJC10	NM_018981
HIGD2A	NM_138820	TUBA1B	NM_006082	MOXD1	NM_015529
NUCB2	NM_005013	SERHL	NR_027786	HIST1H2BG	NM_003518
HLA-DPA1	NM_033554	ITGA5	NM_002205	ID1	NM_181353
SLITRK6	NM_032229	SPARC	NM_003118	SEMA3C	NM_006379

10

20

【 0 2 4 6】

30

【表 3】

表 2: MSKCC、KLEIN、CANCERMAP および STEPHENSON データセット (少なくとも 80/100 の LPD 実施) の中で共通にダウンレギュレートされる 45 遺伝子

C7	CSR1	GPX3	EPAS1	CRISPLD2	PCP4
JAM3	FBLN1	LMOD1	CNN1	ETS2	ACTN1
MYLK	ATP2B4	SPG20	CLU	ILK	CDC42EP3
ACTG2	PPAP2B	STOM	GSTP1	MYL9	SORBS1
STAT5B	PLP2	ITGA5	TIMP3	PALLD	PDK4
TPM2	RBPM5	TNS1	SVIL	FERMT2	
FLNA	CALD1	SNAI2	TPM1	TGFBR3	
KCNMA1	ACTA2	PDLIM1	DPYSL3	VCL	

40

【 0 2 4 7】

50

【表 4】

表 3：表 4 で識別された 1 6 6 9 の遺伝子からの L A S S O 分析によって識別された 2 0 遺伝子

DST	CYP27A1	SP100	ALDH2	MME
CHRD1	RND3	PARM1	WDR59	S100A13
THSD4	ACTG2	ZNF532	LDHB	MSRA
GSTM4	PLEKHA6	DLG5	CDK6	EPHX2

10

【 0 2 4 8 】

20

30

40

50

【表 5 - 1】

表 4:少なくとも2つのデータセットでDESENTおよび非DESENT癌の間で著しく異なる発現を示す
1669の遺伝子

LPP	CX3CL1	NSFL1C	PFKFB3	USP11	CCND2	CLIC4
UGP2	RGL1	CCDC69	PER3	DLD	FBX07	DKC1
MFN2	ATAD1	TRIM29	RFWD2	C11orf54	S100A13	WLS
UTY	CHD1	EIF5	AOC3	ATP2B4	SQRDL	EMP2
SPRY1	ZNF589	STMN1	ATF3	FBX018	COPZ2	SLC2A5
CTNNB1	SETD5	MITF	GON4L	WSB1	ALDH3A2	FBXW4
CAT	ABR	TNKS	TMF1	ST8SIA1	TPP2	GALM
MBTPS1	WDR19	MSRB2	NR4A1	ID1	FAM129A	ECHDC2
SLC38A2	ZCCHC11	MBNL2	SPTBN1	CASC3	PCDH9	ACTA2
CCT3	STK24	TRIP11	FLOT1	RBMS3	FHL2	MADD
ITSN1	PI4KA	PIAS2	DGKA	SPG11	WIPF1	EYA4
SCYL3	NFAT5	RYK	VCL	CCDC121	RBPM5	DLX1
TPST1	CAPNS1	GPR161	TDFP1	SERINC3	SIK2	FAM198B
MGP	METTL3	ACTC1	PREX2	RBBP6	ACOX1	TAB2
SMC1A	RFC2	BRE	PRRG4	CRTAP	LYST	SMARCA2
KCNMB1	SNAI2	ZC3H18	ANKRD12	NUB1	PPIC	TCF7L2
LMBRD1	ANKRD34B	SLC1A1	APEX1	FOXN3	NCOA1	FBLN1
TJP2	SF3A1	GABBR1	MEF2A	AMT	CNOT1	SET
DVL2	ATP2A2	PPP1R10	PI15	EPS15	LONRF3	PPAP2B
IL4R	CDK5RAP2	ROCK2	LARGE	MATR3	UBE2E3	CDH11
FBXO32	DHX9	RARA	PARP6	SKP2	ILF3	SP110
RAB2A	STAU1	SVIL	ANXA7	TUBB	WAC	CAST
ZMYND8	MAPKAPK2	PMP22	PDSS2	PEX10	LRP1	EP400
MTMR9	ATP10D	KIF1B	LRPPRC	RAB27A	TCF12	AFF1
GLIPR2	USP9X	PBRM1	COX7A1	LASP1	TSNAP13	PCBP1
GLT8D1	SLC41A1	ANAPC1	GPBP1L1	NUP214	NRBP1	AIMP2
CLK1	METTL7A	LGALS3BP	WDR11	YTHDC1	CDC45	GNG12
CDC5L	LTBP1	PRMT1	SFXN3	HEG1	KLF3	FAM13B
POGZ	BNC2	BAG3	SON	MORF4L2	LRP10	ADAMTS1
PRDM2	EPC2	ACSS3	TRAF3IP2	AMFR	SERPINC1	EFEMP1
PER1	RUVEL1	MSN	PKP1	COX11	GPM6B	GTF2I
GCNT2	PARM1	PPP1R15A	UBR2	STARD13	PCDH7	MANBAL
SLC22A17	VEZF1	FGA	MXI1	RBMS1	FYTTD1	SSTR1
APP	MYLK	MYH11	STAT3	ROBO1	AMMECR1	TEAD1
DMD	ZYG11B	CDH7	MTUS1	VSIG2	WDFY3	RBAK
KCTD9	SH3RF1	TCF20	HEPH	TRERF1	NDRG2	SORBS2
CUL3	VPS13D	C2orf43	DNAJB5	FAF1	ATOX1	BIN1
ADRA1A	MDH1	POPDC2	TGFB1	WHSC1L1	PITPNC1	HSPB1
SCMH1	APCDD1	LRPAP1	PDLIM4	C9orf72	PPP1R15B	PPARD
ZNF483	AHNAK2	ARID1B	AGL	SYNE1	USP25	C9orf3
NAMPT	ACTR3	ERC1	ELF1	GAB1	EXOSC10	NID1
ITGB4	CBX7	LIMK2	CELF2	PINK1	ZNF207	SF3B1
SMC6	LEPREL1	DYRK1A	MEIS2	PLD3	PDSSA	FAM124A
NBEAL1	MT1M	HIPK1	TP53BP1	TRPM7	IRF2BP2	RNF213
EPB41L5	TSPYL2	TTC17	PTGDS	NF1	MED13	LPAR1

【 0 2 4 9 】

10

20

30

40

50

【表 5 - 2】

表 4 の続き

TMEM51	RHOT1	JAZF1	NBAS	ASAP1	DDX42	PDE8A
IGF1R	DYNLT1	SMAD3	TACC2	CLSPN	KPNA6	TNP01
SYNM	HERC4	PRKCD	CELF1	CAP2	MPHOSPH8	TSPAN18
MYL9	SERPINB1	SMG6	SLC37A3	RNF185	PYGL	UST
UBA6	HSPA9	PDZRN4	DICER1	SEC31A	KCNAB1	SAP130
HSD17B11	DPYSL3	VWA5A	TP53INP2	CLU	CTSB	ALAS1
DDX17	PEL1	PDGFC	SS18	MAPKAP1	STOM	FST
MYADM	ARSJ	UNG	ST5	SNX2	EGFR	CLASP1
SMURF2	PSIP1	CCNL1	FLNA	PARP14	RB1	ELOVL6
ZFP36L1	PPFIBP1	PRICKLE2	DHX8	KHDRBS3	TLN1	DDX24
YY1AP1	AGPAT1	JAK2	CAV1	RAPH1	NEO1	CD99L2
FN1	SETD3	DCN	CPT1A	SMNDC1	TLE4	PRUNE2
PPFIBP2	BRIX1	VPS45	TGFB3	CCN1	LMNA	SLK
UBE4B	GSTP1	IP6K2	BTBD7	USO1	TTC14	ENSA
APOBEC3C	WBP5	HFE	ATP12A	DNAH10	THOC2	GBP1
PDS5B	SLC25A23	CSDE1	NCOA7	CTDSP2	LATS1	PTEN
DDR2	FAM65A	TMLHE	C16orf45	CEBPB	ANG	HP1BP3
WDR1	DYM	SPATS2L	C7	VEGFA	PRPF4B	TBX3
COL4A6	MAP1B	MED13L	PSMC4	AKAP11	MON2	TIMP3
SH3BP5	PAXIP1	STAT5B	PIK3C3	ZC3H7A	PLCL1	CDC73
NEBL	MYOCD	CLIP1	RCAN3	KIAA0513	ACSS2	ZYX
ARHGEF7	CDKL5	NUP98	BPTF	PDCD6IP	LPHN2	DLG1
DST	NSF	NIPAL3	ZMYM4	RTN4	KIAA1109	EFS
KPNA1	ITM2C	CYB5B	UBQLN1	ASPH	TRIM38	TTLL7
DIP2C	CREG1	EPB41L1	APOL1	MLXIP	NCK1	SH3BGRL
MSRB3	NUCKS1	TOPBP1	ZEB2	FAM114A1	PITRM1	PSMD1
NHS	BAZ1B	PKD4	PJA2	PLEKHO1	RBM3	ADHFE1
ZNF460	LMOD1	TNFRSF1A	UTP14A	ARIH1	NFX1	ZRANB2
JPH2	CACNA1D	CAPN7	OGN	NFIX	ORMDL1	WRNIP1
MTMR3	HIF1A	ANXA11	CDK4	YWHAB	TMEM43	AKAP7
SLIT2	PAN3	CACHD1	PIK3R1	ROR2	NID2	CSRNP1
CCDC91	UACA	MCAM	CXCL12	TCIRG1	NHLRC2	FREM2
YEATS2	BACH1	TPM1	MAEA	SCP2	PALLD	MAP11C3B
SYNE2	MAP4K4	HBP1	ZBTB20	MATN2	ASXL2	ATF6
GEM	MXK	DZIP1	NOL8	LMO3	DES	LMO4
SLC10A7	PAPD4	CBLB	CD81	SLC7A8	MAPKAPK5	SLC16A2
PTGS2	PDGFRA	PCNA	CDC42BPB	ZFYVE9	UBR4	KPNB1
USP24	RSPO3	ACOX2	NIPBL	PDE5A	MSMB	TNFRSF19
C1S	AP2B1	EIF4A2	ANO5	FERMT2	TNFRSF10B	UBE3C
ANGPT1	SRI	NFATC3	DUOX1	PDLIM3	IK	LIMA1
AFF3	PSMA4	B4GALT5	BCLAF1	TSC22D3	TUBB6	EPAS1
MAP3K4	NT5C2	POLDIP3	SMG7	PTPRA	PHF21A	MARK3
MME	MIER1	TOP2A	QKI	MRPL10	SLC8A1	CYP27A1
RHOJ	DCAF7	THBS1	PCM1	SEMA3C	ACIN1	NDEL1
CHRD1	FGFR2	NSD1	MKL2	DCBLD1	APBB1	NUMB
AASS	TRIM33	GGA2	VAPA	MAX	ZNF516	TMBIM1
INO80	PCP4	CWC27	CMIP	KIF20A	CLIC6	RELA
FBXO11	IER3	FAM127B	STAT2	CLK4	DEDD	PIK3CA

【 0 2 5 0 】

10

20

30

40

50

【表 5 - 3】

表 4 の続き

PKD2	ABCC13	LITAF	RCC2	FLOT2	AFAP1L2	MACF1
DMXL2	AKAP13	TRIP6	ETS2	TGFBR2	TPR	PRNP
MEIS1	F5	RDH10	TRIP12	RALGAPA2	USH2A	CTNS
RPRD2	EPHX2	PTK2	LRRN1	THSD4	TEX2	PER2
NXF1	CHMP1A	ITSN2	SETBP1	SNX9	CPE	TLL13
RICTOR	CPM	FBXO17	LRCH2	IREB2	ATP1A1	HS1BP3
TTBK2	ALMS1	VAMP3	MAPK14	GPX3	ITIH5	DHX36
DDHD2	YAP1	OGDHL	CSNK1D	PSME1	DDX3Y	TMEM185A
NUP153	SRPX	TNRC6A	ZFP36	PPFIA1	ARHGAP1	USP48
SNRNP200	PGM5	HOXD10	SSX2IP	MYO6	COL6A1	ADH5
LONRF1	IGF1	UBE2C	PYGM	GJA1	PTK2B	PRKAR1A
KANK1	CMBL	ITGB1	BAZ2B	REST	ILK	PRPF8
HECTD1	B3GALT2	UBR3	ABI2	CALU	LRP6	PIGT
ABHD6	ATP8B1	MAGI2	TOMM34	OLFML3	ITGB8	PLP2
DSTN	PARD3	PRPF3	HSPB6	XRN2	BCAS1	ATG9A
KDM3B	MTMR8	ATP6V0E1	ID4	S100A16	RALGAPB	ABI1
COL6A3	ZNF451	CCDC80	GDAP1	EIF4EBP2	ITGB3	LSAMP
KLHL5	RC3H2	ITPK1	RYBP	LDHB	AKT3	DOPEY1
TAGLN	IFI16	MAPK11P1L	TIMP2	SLMO2	TRMU	ETV5
PKN2	MMP19	FLNC	RNF217	SPAG9	KAT2B	NKAIN1
TCERG1	YPEL5	DOCK1	CCT6A	PUM2	OGDH	NFE2L1
ADAR	VPS37A	KANK2	CCNT2	YWHAH	IARS	USP34
REV1	NUDT5	STARD4	KDM3A	ZNF655	YME1L1	CASP14
SORT1	STK4	CKAP5	CDKN1B	TCF21	KIF2A	CYB5R2
DDX19B	ANXA4	ATF7IP	HLF	IL17RA	ZMYND11	ROCK1
CFL2	ARHGAP26	RAB7A	MYH9	STXBP1	ATG2B	PPP1R12A
CDK12	TIMP1	ENAH	SCARA3	SDCCAG8	TIMELESS	DKK3
RGN	SKP1	NPC2	LRP2	DDX3X	SEC24B	SBN01
MPDZ	GPBP1	BOD1	CHST3	SCAMP1	CAV2	PCNX
SLC1A5	ANXA2	GSTM4	OTUB1	C11orf57	DCBLD2	SPEG
MAP4	ANKRD17	AQR	LGALS1	EFTUD1	CDC42SE2	ZNF234
LCLAT1	FOXO4	IVNS1ABP	NR2C2	TOR1AIP1	KCNJ8	CYB5R3
LIX1L	BCOR	SORBS3	AXIN2	C16orf62	NISCH	KCNMA1
DPT	PPP1R3B	SPTA1	SESTD1	GMPR	CNOT4	RAB11FIP2
FAM127A	TIA1	CALD1	CIZ1	GDPD1	SNX33	CHMP2B
OTUD4	NVL	EML4	NCK2	OPA1	ITPR2	KLHDC2
EPS15L1	HADHA	ARHGAP17	NIN	VDAC3	ARHGAP10	USP30
ARL6IP1	LRRC41	GADD45B	CD59	RNF216	CDC42EP3	HOOK3
BIN3	AES	KCTD10	PARN	MPZL2	CD74	SMAD4
CNN2	GSTM2	EDARADD	TSPAN31	ZSCAN18	TMED10	HPS1
AFF4	SMARCA5	CTSA	FOSL2	CASP7	DIXDC1	CLCN6
ADCY5	CYP20A1	WDR26	GSK1	FMNL2	LRRC16A	SERINC1
RDX	VAMP2	CTTNBP2NL	RASA1	NPHP3	SKIL	SSFA2
RABGAP1L	LDB3	MAF	TNRC6B	GNAO1	GGT7	RNF121
RAD50	PRKCB	SYTL4	YTHDC2	GCLC	FLII	CEP350
EAF2	ATM	TMEM63A	PTPLA	ARRB1	MAT2A	TAPBP
FYCO1	S100A6	NFKBIZ	PAK1IP1	LGALS3	BCL6	MEF2C
RBM4	CYTH3	TNC	CDC27	RUFY3	N4BP2L2	MTPAP

【 0 2 5 1 】

10

20

30

40

【表 5 - 4】

表 4 の続き

MKLN1	DEK	CAPRN1	COMMD6	NPAS2	CD47	CD44
TRA2B	ATF2	BCL7B	MID2	MAML2	PEA15	VILL
EXOC4	MAPK10	ADCY8	SRGN	NUFIP2	RRM1	NFIB
DIRAS2	MBNL1	R3HDM1	LIMS2	REL	GLI3	CD40
TUBA1A	ALDH1A2	FNBP1L	NETO2	MRVI1	GLG1	PUS7
EEA1	MRAS	TTLL5	GIT2	SUPT4H1	SUN1	UTP18
CA11	REXO2	ZCCHC24	GNL2	GATAD2B	PDHA1	PTPRG
AKT2	RIC3	FAT1	COMMD1	MSL2	KIF16B	KLF4
ACOX3	AUTS2	DHX15	ARID4B	MFAP4	ARPP19	TBC1D14
MLLT10	PSAP	TBC1D1	EFHD2	AOX1	GAS1	PSMC5
HNRNPU	CUL1	MAN1A2	EIF4G2	SOS1	STRN3	DYNC1H1
ATXN2	SORBS1	TTC28	CSTB	ZNF280D	GPR124	RBM23
TSHZ3	EXOC7	CALCOCO2	MMP2	MAPK1	OAZ1	RRAS
ELP3	PPARGC1A	HK1	ZEB1	TBC1D5	NFKBIA	CEP120
GNS	DMTF1	DIP2B	ARNT	SCPEP1	SCN7A	STAM2
EP300	PTPN14	STK38L	HELZ	BBS2	DOCK9	DUSP1
FGF2	ATP2B1	CPEB3	EGR1	AFTPH	USP4	RHOA
DLG5	GIGYF2	PARVA	CHD9	GAS6	SMARCC2	CDC42BPA
TBL1X	GSTM5	SCRN1	NEU1	PRPSAP1	PAICS	SUPT16H
PTPRM	ACO1	SMURF1	STAT6	IL13RA1	TGFB11	TRAK2
RHOBTB3	STXBP6	EIF5B	MEF2D	BHLHE40	MED21	PRRG1
GGCT	SERPINH1	MCL1	CHMP1B	UCK2	STX12	ASH1L
CLINT1	SMAD2	RBL2	TNKS2	FXYD6	TMEM165	ATP8B2
LAPTM4A	ATL3	SMC5	TOP1	AP3B1	NT5DC3	KIAA1033
ANO4	CREB3L2	ASAP2	SETX	LBR	CALCOCO1	LAMP1
ZNFX1	ABCC9	LRCH3	PSME4	MTOR	NR4A3	TRPC4
CDC42EP4	FOSB	PTRF	ZC3H13	GLIPR1	CDC42EP5	NR4A2
PLSCR1	COQ10B	TPM2	ANPEP	FRMD6	NCAPD2	POLR2A
IFI35	CHRM1	NEIL3	ACACB	SETD2	DNAJB1	CNN3
HNRNPM	ITM2B	ZNF611	SEC63	PRKDC	EIF4G3	VIM
PCDH15	ALOX15B	INO80D	C1R	RIN2	GNAI2	JMMT
BBX	TMEM55A	NFIA	STXBP3	SLFN5	SPATA6	PAGE4
EXOC1	ERAP1	PRPS2	JARID2	JAM3	EPHA3	ARHGAP20
A2M	DNAJC13	PIBF1	CDC37L1	TBCK	ZNF396	GALNT8
ASCC3	ITCH	RARS2	DAB2	ARL6IP5	TBCEL	SLMAP
TGFBR1	DAAM2	HMGXB4	SOS2	IDE	FUBP1	FBN1
CORO1C	LARP6	TSC1	TECPR2	RBBP7	PHF11	NEXN
GNAL	IFNAR1	NEK7	GPATCH8	TACC1	ATP1A2	PUM1
GTF3C2	FAM160B1	IDS	SLC39A14	BTG2	APOOL	EPRS
IL6ST	LAMB2	FAM107B	SH3PXD2B	VPS39	NCBP1	MORC3
TTLL4	KIF15	SUPT6H	ZNF384	AHCYL1	NOTCH2	TGFBR3
TNRC6C	IPO8	EARS2	AP3D1	KRT15	STX6	SECISBP2L
SAT2	WNK1	ANKRD40	JMJD1C	TEP1	CALM1	UBP1
HIPK3	PLXDC2	IRS1	COL6A2	WDR12	SEN7	KBTBD2
PHF1	CD63	ADD1	TCEAL2	COPS3	PYGB	SBF2
TSG101	STAT5A	DENND4A	STAM	BNIP2	TRAPPC10	PBX1
EYA1	HDDC2	NNT	EMP3	PNMA1	KLF9	HERPUD2
C15orf41	NPTN	RND3	SHKBP1	FBXO31	ZNF3	SWAP70

【 0 2 5 2 】

10

20

30

40

50

【表 5 - 5】

表 4 の続き

DENND5A	SMOC1	FNBP1	TRIO	ROS1	SLC18A2	AHR
PPIP5K2	HSPB8	DUSP3	FHL1	LDB1	HIST1H4C	RASD2
TTL3	ITGA7	PLEKHA6	SIN3A	FAM20B	MRGPRF	RAB8B
SMTN	EZH1	CAP1	MYO1D	PLEK2	KHDRBS1	MYO9A
PRKD1	PDE4D	PHF3	JUN	DERA	LSM14A	XPO7
GPRC5B	KRT23	CHURC1	ENTPD4	COPA	SLC12A4	KCNS3
PRKACA	SPON1	LNPEP	ACADVL	CSRP1	M6PR	DDX1
HERC1	C10orf76	CAPZB	VPS53	MYCBP2	POLR2B	ANXA1
ZZEF1	ZNF318	PCDH18	HEXB	C11orf30	OTUD5	CYR61
SNTB2	PHC3	KIF4A	UBE2E1	PRPSAP2	SPRY2	RGS2
RBM5	AMOT	SNRNP40	USP14	TGFB2	TMEM109	ARRDC3
WWTR1	STAG1	CST3	TINAGL1	MYO1C	SPTAN1	REV3L
DAAM1	ARPC2	CSNK2B	PRPF18	ANTXR2	PLEKHA5	OSMR
GSN	AGFG1	LDB2	PKD2	ITGA9	SAMD8	SLC15A2
C2orf88	TMEM59	RLF	UBAP1	PDE11A	TMEM220	REPS1
GPRASP1	STX7	SMG1	TNS1	RAF1	XRCC5	PPWD1
CDKAL1	VPS4B	DCUN1D4	GNG2	PTN	FNBP4	TMEM35
SLC25A12	ITGA5	BIRC6	KIF14	DARS	UFC1	TBC1D23
PCGF5	DAPK3	EMP1	RBPMS2	TEAD3	CTGF	MSRA
KIF5B	ZHX2	KRT5	PPP1R7	ZFR	NPAT	ABCB11
ARMCX1	KIAA0430	PRDM8	SLC4A7	PSMB7	CISD1	ACTN1
SNX19	JAK1	RHOB	DRAM2	SMARCA4	CNPY2	CD38
WBP2	MED12	PTTG1IP	EHD2	TCF4	SEC24A	QRICH1
PHIP	RNF38	ITGA1	STRBP	TRPS1	FOXJ3	SP100
KLF8	ALDH2	SPEN	NPR2	DEPDC1B	TMEM47	CYLD
TET2	XYLB	CDK6	MYL6	UBAP2L	EXT1	TRO
MIB1	SIDT1	EPHB6	XRN1	TLE2	PAK3	CD46
SRD5A2	ZFAND5	PPP3CB	RAP1A	TCF25	IGFBP5	OSBPL9
PDLIM1	SPARCL1	MTMR12	PITPNB	CYC1	CNOT6	NCKAP1
GDAP2	USP53	ZNF185	DCP1A	PLAGL1	FABP3	SOD2
DCTN1	ACTG2	FAM160B2	VAMP5	MTR	TP63	PTP4A2
BMPR2	SPOP	SF3B3	VPS13C	SMAD9	SHISA5	CHD2
CCDC25	WDR59	BIRC5	CREBBP	LZTFL1	SERPINF1	SPOCK3
ITPR1	LAMA4	MXRA5	CAMK2G	FCHSD2	ZNF148	G3BP1
GTF3C3	MCC	EHBP1	CNN1	SOX4	CRIM1	PREPL
ETV6	DPYD	AEBP2	MAP3K7	CREB1	MAN2A1	FUBP3
TBC1D9B	ASB2	ZFC3H1	MYOF	HNRNPA2B1	QSER1	RSRC2
ARFGEF2	ZBTB4	IQGAP1	SGCB	PIP4K2A	MPPED2	SMARCA1
SEC23A	CHMP7	BOC	NFYC	UBC	RCBTB2	AP1G1
PHACTR2	VPS41	SPRED1	IL1R1	RQCD1	AKIRIN2	PPP1CB
CRISPLD2	CRY2	FZD7	ARHGEF12	SLC22A3	GABARAPL1	MAP4K5
ADSL	FGFR1	GNG4	DCUN1D1	FASTKD2	STK38	GALC
XPC	ASNS	CTNNA1	RNF11	SENK6	KDSR	FNDC3B
NFE2L2	GABARAPL2	ERBB2IP	RARRES2	ESYT2	GBF1	PPIL4
CDS2	TRIP13	SYNRG	CYP3A5	RABGAP1	SHOC2	ZNF532
HUWE1	EDNRA	DDX5	PTPRK	STIM1	EPCAM	MARVELD1
AH11	ABCA8	EPB41L2	CCDC88A	GRAMD3	TRIP10	SLAIN2
YPEL3	AZGP1	SLC14A1	SCAPER	NCAPG2	NEK1	RAB3GAP2

【 0 2 5 3】

【表 5 - 6】

表 4 の続き

KDM2A	DCAF8	MYO15B	ZNF638	FAM69A
RAD54L2	RIMKLB	CRTC3	WFDC2	L3MBTL4

【 0 2 5 4】

10

20

30

40

50

【表 6】

表5:MSKCC、KLEIN、CAMCAPおよびSTEPHENSONデータセット（少なくとも67／100のLPD実施）の中で共通してダウンレギュレートされる35の遺伝子
100の67の35の遺伝子

ACTN1	ANXA2	HSPB8	ILK	CSRP1	FERMT2
ATP2B4	ACTG2	PCP4	MYLK	CNN1	JAM3
LMOD1	TPM2	SORBS1	MYH11	DPYSL3	VCL
LPAR1	MYL9	STOM	FBLN1	KCNMA1	PALLD
GSTP1	C7	TGFB3	RND3	CXCL12	ITGA5
PTRF	ACTA2	TGFBR3	FZD7	FLNA	

10

【0255】

【表 7】

表6:コントロール遺伝子の例：ハウスキーピング・コントロール遺伝子

HPRT	18S rRNA	RPL9	PFKP	H2A.X	RPL23a
B2M	28s rRNA	SRP14	EF-1d	IMP	RPL37
TBP	PBGD	RPL24	IMPDH1	accession number X56932	RPS11
GAPDH	ACTB	RPL22	IDH2	ODC-AZ	RPS3
ALAS1	UBC	RPS29	KGDHC	PDHA1	SDHB
RPLP2	rb 23kDa	RPS16	SRF7	PLA2	SNRPB
KLK3_ex2-3	TUBA1	RPL4	RPLP0	PMI1	SDH
KLK3_ex1-2	RPS9	RPL6	ALDOA	SRP75	TCP20
SDH1	TFR	OAZ1	COX IV	RPL3	CLTC
GPI	RPS13	RPS12	AST	RPL32	
PSMB2	RPL27	LDHA	MDH	RPL7a	
PSMB4	RPS20	PGAM1	EIF4A1	RNAP II	
RAB7A	RPL30	PGK1	FH	RPL10	
REEP5	RPL13A	VIM	ATP5F1		

20

30

【0256】

【表 8】

表7:コントロール遺伝子の例：前立腺特異性のコントロール・トランスクリプト

KLK2	PCGEM1	TGM4	PSCA	HOXB13	SPINK1
KLK3	PCA3	RLN1	NKX3.1	PMEPA1	
KLK4	TMPRSS2	ACPP	SPDEF	PAP	
FOLH1(PSMA)	TMPRSS2/ERG	PTI-1	PMA	STEAP1	

40

【0257】

表 8 :

DESENT癌の患者の不良な臨床結果

各データセットについて、DESENTおよび非DESENT癌について報告されたPSA

50

再発間で比較された。L P D、ラテントプロセスデコンポジション；R F - ランダムフォレスト。L P Dについては、ログランク p 値は、方法に述べられているように100の独立したL P D検定から選ばれたモーダルL P D検定を表す。多変量分析については、グリーンソンおよび臨床病期データだけが利用可能なT C G Aデータセットを例外として、診断時のグリーンソン、P S Aおよび病理学的病期がすべてがデータセットに含まれている。完全な分析は図7に示される。

【0258】

【表9】

データセット	ラテントプロセスデコンポジション	
	一変量のp値	多変量のp値
MSKCC	2.65×10^{-5}	3.27×10^{-1}
CancerMap	2.98×10^{-8}	3.66×10^{-3}
Stephenson	4.28×10^{-5}	1.21×10^{-4}
CamCap	1.22×10^{-3}	2.90×10^{-2}

データセット	ランダムフォレスト	
	一変量のp値	多変量のp値
MSKCC	1.85×10^{-3}	6.05×10^{-1}
CancerMap	4.80×10^{-4}	1.45×10^{-2}
Stephenson	1.75×10^{-4}	4.56×10^{-4}
CamCap	1.61×10^{-5}	1.31×10^{-4}
TCGA	5.41×10^{-4}	2.59×10^{-2}

【0259】

拡張データテーブル

拡張データテーブル1：

D E S N T 癌グループ中の変更された発現を示す遺伝子。
各データセットについては、非D E S N Tグループと比較されたD E S N T 癌グループ中の著しく変化した発現 ($p < 0.05$) を示す遺伝子が計算された：p 値は複数の試験のために補正された。L P Dは、異なる無作為に選択されたシード値を使用して、各データセットのために再度100回実行された。4つのデータセットすべてについて、少なくとも80/100実施での発現が変化した45の遺伝子について結果がリストされる。遺伝子がそれぞれ著しく発現が変化した実施の正確な数が示される。D E S N T 癌グループの中のすべての遺伝子はダウンレギュレートされた。強調は、その生成物が以下のコンポーネントであるかリンクされた遺伝子を表わす：細胞骨格（太字）；付着、インテグリンおよび細胞間マトリックス（下線）、転写因子および翻訳レギュレーター（二重下線）、およびイオンチャネル（点線の下線）。シンボル：* 癌の中のC p Gメチル化によるダウンレギュレーション；** 前立腺癌中のC p Gメチル化によるダウンレギュレーション；† 不良な結果に関係するC p Gメチル化；‡ の前立腺癌の機能的な接続性ハブ；|| 前立腺癌にフォーカスされた遺伝子遺伝子相互作用。

【0260】

10

20

30

40

50

【表 1 0】

遺伝子	MSKCC	CancerMap	Stephenson	Klein
<u>ACTA2</u>	100	92	100	98
<u>ACTG2</u>	100	98	100	98
<u>ACTN1</u>	100	92	100	100
<u>ATP2B4</u>	100	92	100	100
<u>C7</u>	100	89	100	100
<u>CALD1</u>	100	92	92	100
<u>CDC42EP3</u>	100	92	100	95
<u>CLU**</u>	100	92	100	100
<u>CNN1</u>	100	92	100	98
<u>CRISPLD2</u>	100	92	100	98
<u>CSRP1*†</u>	100	93	100	100
<u>DPYSL3**</u>	100	92	100	86
<u>EPAS1* </u>	100	92	100	100
<u>ETS2</u>	100	92	100	100
<u>FBLN1*†</u>	100	92	100	100
<u>FERMT2</u>	100	92	100	100
<u>FLNA</u>	100	92	100	98
<u>GPX3*†</u>	100	92	100	100
<u>GSTP1***†</u>	100	92	100	81
<u>ILK</u>	100	92	100	100
<u>ITGA5</u>	100	92	100	100
<u>JAM3*</u>	92	85	100	100
<u>KCNMA1***†</u>	100	92	100	99
<u>LMOD1</u>	100	92	100	91
<u>MYL9</u>	100	92	100	98
<u>MYLK*†</u>	100	92	100	98
<u>PALLD</u>	100	92	100	100
<u>PCP4</u>	100	92	100	100
<u>PDK4</u>	100	83	100	96
<u>PDLIM1</u>	100	91	100	81
<u>PLP2</u>	100	92	100	100
<u>PPAP2B</u>	100	92	100	100
<u>RBPM5</u>	100	92	100	100
<u>SNAI2**</u>	100	93	100	91
<u>SORBS1*</u>	100	92	100	98
<u>SPG20*</u>	100	92	100	100
<u>STAT5B</u>	100	92	100	100
<u>STOM</u>	100	92	100	100
<u>SVIL**</u>	100	83	100	100
<u>TGFBR3</u>	100	92	93	87
<u>TIMP3*†</u>	100	92	100	97
<u>TNS1</u>	100	92	100	100
<u>TPM1*</u>	100	92	100	100
<u>TPM2</u>	100	92	100	80
<u>VCL</u>	100	92	100	100

10

20

30

【0 2 6 1】

拡張データテーブル 2 :

20の遺伝子ランダムフォレストクラシファイヤー。5つのデータセット(MSKCC、CancerMap、Stephenson、KleinおよびCamCap)の内少なくとも2つのデータセットでDESN T癌中で著しく変化した発現を示す1669の遺伝子のリストが、出発点として使用された。lassoロジスティック回帰モデルを適用し、MSKCCデータセット内のDESN Tメンバーシップを予測し、この表に示される20の遺伝子のセットを選択した。各遺伝子については、ランダムフォレスト分類を行なう場合の変数としての重要性も記録される。

【0 2 6 2】

40

50

【表 1 1】

遺伝子	変数重要性	
<i>DST</i>	2.146140965	
<i>CHRD1</i>	1.758974273	
<i>THSD4</i>	1.561264948	
<i>GSTM4</i>	1.550345548	
<i>CYP27A1</i>	1.408713974	
<i>RND3</i>	1.339094656	
<i>ACTG2</i>	1.304989674	
<i>PLEKHA6</i>	0.735553263	10
<i>SP100</i>	0.680938431	
<i>PARM1</i>	0.671688267	
<i>ZNF532</i>	0.630661162	
<i>DLG5</i>	0.492853186	
<i>ALDH2</i>	0.481637788	
<i>WDR59</i>	0.467824475	
<i>LDHB</i>	0.449345969	
<i>CDK6</i>	0.351043941	
<i>MME</i>	0.275274353	
<i>S100A13</i>	0.250416073	
<i>MSRA</i>	0.229702526	20
<i>EPHX2</i>	0.213536527	

【0 2 6 3】

補助情報表

補助情報表 1 :

差別的メチル化

T C G A データセットの中で同定された D E N S T および非 D E S N T 癌の間の差別的メチル化が示される。D E S N T 癌は、ランダムフォレスト分類を使用して、拡張データ表 2 に示される 2 0 の遺伝子シグネチャーを使用して同定された。その後、我々は、R パッケージに実装された「methyAnalysis」(<http://www.bioconductor.org/packages/release/bioc/html/methyAnalysis.html>)で、差別的にメチル化された部位 (D M R) を検知する方法を適用した。有意な結果がリストされる。

【0 2 6 4】

30

40

50

【表 1 2 - 1】

C h r	スタート	エンド	プローブ ナンバー	遺伝子シンボル	距離 T S S *	プロモータ	最小 p 値	調整最小 p 値
1	56992372	56992372	1	PPAP2B	52885	FALSE	1.71E-28	4.92E-27
1	92197531	92197531	1	TGFBR3	130072	FALSE	1.56E-12	7.59E-12
1	92295946	92295946	1	TGFBR3	31657	FALSE	3.72E-16	2.56E-15
1	203598330	203599089	7	ATP2B4	2415	FALSE	7.48E-25	1.23E-23
1	203605590	203605590	1	ATP2B4	9675	FALSE	1.34E-26	2.70E-25
1	203670963	203671140	2	ATP2B4	19093	FALSE	1.44E-39	4.05E-37
10	29923736	29924258	3	SVIL	0	TRUE	1.11E-29	4.22E-28
10	29936149	29948428	3	SVIL	76302	FALSE	2.48E-35	3.09E-33
10	29981216	29981216	1	SVIL	43514	FALSE	4.12E-21	4.57E-20
10	79150517	79150517	1	KCNMA1	247060	FALSE	3.38E-22	4.17E-21
10	79396584	79396793	3	KCNMA1	784	FALSE	1.08E-12	5.32E-12

10

【 0 2 6 5 】

20

30

40

50

【表 1 2 - 2】

10	97049610	97049610	1	<i>PDLIM1</i>	1295	FALSE	2.64E-29	8.99E-28	10
10	97169147	97175479	4	<i>SORBS1</i>	6351	FALSE	1.75E-33	1.31E-31	
11	67350976	67350976	1	<i>GSTP1</i>	-90	TRUE	1.65E-14	9.79E-14	
11	67351271	67352041	6	<i>GSTP1</i>	205	FALSE	1.03E-36	1.92E-34	
11	134020750	134020750	1	<i>JAM3</i>	81930	FALSE	4.58E-28	1.29E-26	
12	54811762	54812085	3	<i>ITGA5</i>	965	FALSE	4.27E-27	9.99E-26	
13	36919344	36919960	6	<i>SPG20</i>	686	FALSE	8.29E-18	6.41E-17	
14	69443362	69443362	1	<i>ACTN1</i>	921	FALSE	5.45E-35	6.12E-33	
15	63345124	63345124	1	<i>TPM1</i>	4488	FALSE	1.25E-12	6.11E-12	
16	84870066	84870203	2	<i>CRISPLD2</i>	16479	FALSE	1.14E-25	2.00E-24	
16	84918794	84918851	2	<i>CRISPLD2</i>	65207	FALSE	7.28E-18	5.79E-17	20
2	46526843	46527098	2	<i>EPAS1</i>	2302	FALSE	7.50E-10	3.02E-09	
2	218767655	218767655	1	<i>TNS1</i>	881	FALSE	9.13E-16	6.24E-15	
20	35169380	35169594	3	<i>MYL9</i>	-293	TRUE	6.90E-31	3.09E-29	
22	45899736	45899736	1	<i>FBLN1</i>	1017	FALSE	6.75E-35	6.89E-33	
3	123339417	123339568	2	<i>MYLK</i>	0	TRUE	9.65E-23	1.27E-21	
3	123414733	123414733	1	<i>MYLK</i>	5623	FALSE	1.68E-32	1.18E-30	
3	123535716	123535716	1	<i>MYLK</i>	14614	FALSE	1.31E-33	1.23E-31	
3	123602485	123602485	1	<i>MYLK</i>	664	FALSE	3.14E-32	2.07E-30	
4	169664785	169664785	1	<i>PALLD</i>	112017	FALSE	2.54E-26	4.99E-25	30
4	169737224	169737224	1	<i>PALLD</i>	184456	FALSE	1.02E-26	2.12E-25	
4	169754328	169754534	2	<i>PALLD</i>	1172	FALSE	9.98E-11	4.27E-10	
4	169770092	169770092	1	<i>PALLD</i>	16936	FALSE	2.81E-24	4.37E-23	
5	40933444	40982092	2	<i>C7</i>	23845	FALSE	3.70E-10	1.51E-09	
7	134575145	134575524	5	<i>CALD1</i>	110981	FALSE	1.24E-22	1.62E-21	
7	134626083	134626083	1	<i>CALD1</i>	8344	FALSE	1.31E-15	8.70E-15	
8	27468981	27469186	3	<i>CLU</i>	82	FALSE	7.22E-28	1.84E-26	
8	30243241	30243260	2	<i>RBPM5</i>	1297	FALSE	2.72E-15	1.74E-14	
8	30254923	30254923	1	<i>RBPM5</i>	12979	FALSE	8.21E-29	2.56E-27	
8	30290489	30290489	1	<i>RBPM5</i>	48545	FALSE	2.39E-11	1.06E-10	40
8	30419935	30419935	1	<i>RBPM5</i>	84620	FALSE	6.82E-32	4.25E-30	
X	153598077	153598077	1	<i>FLNA</i>	4929	FALSE	1.28E-10	5.39E-10	

【 0 2 6 6 】

補助情報表 2 :

C a n c e r M a p データセットの臨床的特徴

【 0 2 6 7 】

10

20

30

40

【表 1 3】

	カテゴリー	カウント/メジアン (レンジ)
患者		154
前立腺切除時年齢		62 (21-74)
前立腺切除時P S A		7.9 (2.4-40)
フォローアップ期間 (月)		56 (1-129)
再発	有	35
	無	102
	不明	17
グリーソン	6	40
	7 (3+4)	83
	7 (4+3)	20
	8	4
	9	7
病期	T1c	1
	T2a	6
	T2b	3
	T2c	49
	T2x	26
	T3a	50
	T3b	17
	T4x	2

10

20

【0 2 6 8】

補助情報表 3 : 差別的に発現した遺伝子の機能

4 5 の遺伝子のリストは、M S K C C、S t e p h e n s o n、C a n c e r M a p および K l e i n データセットの中で同定された D E S N T 癌で共通してダウンレギュレートした。

30

【0 2 6 9】

40

50

【表 1 4 - 1】

遺伝子	同一性	注記
ACTA2	平滑筋アクチン・アルファ 2	アクチン細胞骨格および細胞運動；筋線維芽細胞のためのマーカー
ACTG2	平滑筋アクチン γ 2	細胞骨格のコンポーネント、細胞運動に関係する一発現は実際に広範囲である。
ACTN1	アクチニン・アルファ 1	α -アクチニンは、異なる細胞型中で多数の役割を備えたアクチン結合タンパク質である。 非筋肉細胞では、細胞骨格のアイソフォームが微小繊維束、およびアクチンの細胞膜への結合に関係するアドヘレンスタイプジャンクション (adherens-type junction) に沿って見つかる。 対照的に、骨格のアイソフォーム、心臓のアイソフォームおよび平滑筋アイソフォームは、Z-ディスクとデンスボディ (dense bodies) 相似体に局地化し、それらは、筋原繊維アクチンフィラメントのアンカーを支援する。 この遺伝子は非筋肉、細胞骨格、 α -アクチニン・アイソフォームをコード化し、構造上類似する赤血球のベータ・スペクトリン遺伝子と同じ部位へマップする。[2]
ATP2B4	ATPアーゼ原形質膜 Ca^{2+} 輸送 4	細胞内 Ca ホメオスタシス
C7	補体成分 7	補体系賦活；ピトロネクチン受容体による付着への潜在的なリンク
CALD1	カルデスモン	平滑筋および非筋肉収縮の調整に本質的な役割を果たすカルモデュリンおよびアクチン結合タンパク質
CDC42EP3		CDC42 に結合し、負の調節をする一糸状仮足内のアクチンフィラメントアセンブリに含まれる低分子量 GTP アーゼ。
CLU	クラステリン	CLU は、分泌された蛋白質の蛋白質の折量の援助、細胞デブリのクリアランスとアポトーシスに責任を負う分子シャペロンである；
CNN1	calponin 1	calponin 1 はアクトミオシン相互作用を阻害することによって平滑筋収縮性を抑制する調整物質として機能する。[2] [23] [24] この調整では、 Ca^{2+} とカルモデュリンの結合および PKC リン酸化はアクチンフィラメントから calponin 1 を分離し、平滑筋収縮を容易にする。[25]
CRISPLD2	システイン富化分泌蛋白質 LCCL ドメイン含有 2	分泌蛋白質；aka 妊娠後期肺-1。細胞移動の調整に関係する

【0 2 7 0】

10

20

30

40

50

【表 1 4 - 2】

CSRP1	システインおよびグリシン富化蛋白質 1	CSRP1は、LIMドメイン蛋白質のグループをコード化する遺伝子のCSRPファミリーのメンバーである。それは細胞の成長および分化にとって重要な調節プロセスに関係する。CRP1に見られるLIM/ニ重ジクフィンガー・モチーフは、遺伝子調節、細胞増殖および分化で重要な機能を備えた一群の蛋白質に見いだされる。ファミリーの他の遺伝子はCSRP2とCSRP3を含んでいる。[3]
DPYSL3	ジヒドロピリミジナーゼ様 3	推定上の腫瘍抑制遺伝子；フォーカル付着複合体(focal adhesion complexes)のスタビライザー細胞移動へのリンク
EPAS1	内皮のPASドメイン含有蛋白質 1 低酸素症誘引可能な要因-2アルファ(HIF-2アルファ)として知られているEPAS1	HIF2アルファ低酸素症に対する細胞の反応を規制する重要な転写因子
ETS2	Ets-2	転写因子のc-Etsファミリーの要素
FBLN1	fibulin 1	Fibulin-1は、フィブロネクチン含有ファイバー、エラスチン含有ファイバーおよび基底膜を含む細胞外マトリックス構造で見られる、分泌された糖タンパク質である。Fibulin-1は、フィブロネクチン、[3] ニドゲン-1およびプロテオグリカン、パーシカン[3][4]を含む、多くの細胞外マトリックス要素に結合する。Fibulin-1はさらにフィブリノゲンに結合することができる血液蛋白である。[5]
FERMT2	ファーマチンファミリーメンバー 2	FERMT2は哺乳動物細胞中の細胞外マトリックス構造のコンポーネントで、細胞形変化の適正管理に必要である。[4]
FLNA	フィラミンA	アクチン結合タンパク質またはフィラミンは、皮質細胞質中の直交ネットワークへアクチンフィラメントと交差結合し、アクチン細胞骨格用蛋白質のアンカリングに参加する、280-kDの蛋白質である。細胞骨格のリモデリングは細胞形および移動の調整の中心となる。FLNA遺伝子によってコード化されたフィラミンAは、インテグリン、膜内外の受容体複合体および二次伝達物質[OMIMによって供給された][3]との相互作用により、アクチン細胞骨格の再編成を規制する、広く表現された蛋白質である。
GPX3	グルタチオン・ペルオキシダーゼ-3	Aka血漿、あるいは細胞外のグルタチオン・ペルオキシダーゼ；過酸化水素の解毒に関係する。
GSTP1	グルタチオンS-トランスフェラーゼ・P11	グルタチオンS-トランスフェラーゼ(GST)は、還元グルタチオンとの多くの疎水性親電子性の化合物との結合を触媒することにより、解毒に重要な役割を果たす酵素のファミリーである。
ILK	インテグリン-リンクドキナーゼ	ベータ1インテグリン、付着中の役割、移動生存などを伴う。
ITGA5	インテグリン・アルファ5	インテグリン・ベータ1は、フィブロネクチン受容体を構成する；付着、移動、シグナリング(signal)ing)に関係していた。
JAM3	ジャンクション接着分子-3	AkaJAM-C；密着結合による細胞間相互作用；Mac-1によるプレートリット白血球相互作用において重要。
KCNMA1	カルシウム-活性化されたカリウムチャンネル・サブユニットアルファ-1	電位依存性カリウムチャンネル
LMOD1	レイオモジン(Leiomod)in-1	平滑筋アクチンおよびトロポミオシン結合
MYL9	ミオシン軽鎖9	筋収縮および結合CaおよびMLCKによる作用
MYLK	ミオシン軽鎖キナーゼ	MLCK；Ca/カルモジュリン依存性
PALLD	パラジン(Palladin)	パラジンは、細胞形、付着および収縮をコントロールする、アクチン含有微小繊維のコンポーネントである。[4]
PCP4	ブルキン細胞蛋白質-4	PCP4は、カルモジュリン(CaM)とカルシウム(Ca ²⁺)の会合および解離の両方を加速する。それはCaM依存性酵素、特にCaMKII(CaMK-III)の活性に影響を及ぼすと仮定される。
PKD4	ピルビン酸デヒドロゲナーゼ・リボアミド・キナーゼ・イソチーム4、ミトコンドリア	クレブス回路の調整；ミトコンドリアのマトリックス内に配置され、アセチルCoAのピルベートへの転換を低減し、そのサブユニットのうちの1つの磷酸化によりピルビン酸デヒドロゲナーゼ複合体を阻害する。

【0 2 7 1】

10

20

30

40

50

【表 1 4 - 3】

<u>PDLIM1</u>	P D Z および L I M ドメイン 蛋白質 1	α - アクチニン-1 およびアクチンフィラメントに結合し、細胞移動を規制する。
<u>PLP2</u>	プロテオリピッドタンパク質 2	ケモカインレセプター C C R 1 と相互作用し、細胞移動を調整する。
<u>PPAP2B</u>	脂質リン酸塩フォスホヒドロ ラーゼ 3	フォスファチジン酸ホスファターゼ (P A P) ファミリーのメンバー。 P A P は、 ジアシルグリセロールへフォスファチジン酸を変換し、およびホスホリパーゼ D に よって仲介された、レセプター活性化信号伝達、およびグリセロ脂質の新規合成に おいて機能する。
<u>RBPM5</u>	複数のスプライシングを備えた RNA-結合蛋白質	RNA-結合蛋白質の R R M ファミリーのメンバー：胃腸の平滑筋の成長を規制する。
<u>SNAI2</u>	ジンクフィンガー蛋白質 S N A I 2	C 2 H 2 タイプ・ジンクフィンガー転写因子の S n a i l ファミリーのメンバー。 コード化された蛋白質は、 α -ボックスモチーフ (E-box motifs) に結合して転写 リプレッサーとして作用し、さらに、乳癌中の α -カドヘリン転写 (E-cadherin transcription) を抑制する。
<u>SORBS1</u>	ソルビンとして知られている C A P / P o n s i n 蛋白質、および S H 3 ドメイン含む蛋白質 1	C A P / P o n s i n は、細胞接着、発育因子信号および細胞骨格の形成を規制 するアダプタータンパク質の小さなファミリーの一部である。
<u>SPG20</u>	Spartin	蛋白質はエンドソーム間トラフィック、微小管力学あるいは両方の機能に関係するこ とがある。
<u>STAT5B</u>	転写 5 B の信号の変換および 活性化体	I L 2 、 I L 4 、 C S F 1 および異なる成長ホルモンのような様々な細胞配位子が 引き金となって起きる信号伝達を仲介する転写因子
<u>STOM</u>	ヒト赤血球インテグラル膜蛋白質 バンド 7 として知られる stomatin (Stomatin)。	インテグラル膜蛋白質、イオンチャネルのレギュレーター
<u>SVIL</u>	スーパービリン (Supervillin)	さらに核移行シグナルを有するアクチン結合タンパク質；その機能は、原形質膜で、 および細胞を育てる核の中で専門の構造の中へのアクチンおよび他の細胞骨格蛋白質 の漸増を含んでいることがある。
<u>TGFBR3</u>	ベータグリカン T G F b e t a レセプター I I I	形質転換増殖因子ベータ・レセプター I I I (T G F B R 3) としても知られるベータグ リカン (Betaglycan) は、分子量 > 3 0 0 k D a の細胞表面コンドロイチン硫酸塩／ヘ パラン硫酸プロテオグリカンである。 ベータグリカンは、そのコア蛋白質を介して配位子の T G F ベータ・スーパーファミ リーの様々なメンバーと、およびそのヘパラン硫酸鎖を介して b F G F に結合する。 それは、 T G F ベータ信号伝達に直接関与するのではなく、 T G F ベータ・レセプターに ついての配位子のリザーバーとしてそれが作用する細胞表面で T G F ベータ・スーパー ファミリーの様々なメンバーに結合することによって関与する。 [1] [2]
<u>TIMP3</u>	コラーゲン分解酵素阻害-3	MMP およびさらに他の A D A M 、 A D A M T S メタロプロテアーゼ (metalloproteinases) の陰性のレギュレーター； E C M リモデリングおよび細胞信号 の調整に関係する。
<u>TNS1</u>	Tensin-1	接着点 (focal adhesion) に見いだされる細胞骨格レギュレーター、アクチンフィラメ ントを架橋し、 S H 2 ドメインを有し、またたぶん細胞信号に関与する；最近の論文 では R h o A を陽性に規制する。
<u>TPM1</u>	トロポミオシン・アルファ-1 鎖	横紋筋および平滑筋の収縮システムに包含されるアクチン結合タンパク質で、非筋細 胞の細胞骨格に含まれる。
<u>TPM2</u>	β - トロポミオシン β 鎖として 知られているトロポミオシン	β - トロポミオシンは、アクチンフィラメントを安定させて、かつ筋収縮を規制 するように機能する横紋筋に特有の多段コイル二量体である。
<u>VCL</u>	ビンキュリン	ビンキュリンは、アクチン細胞骨格へのインテグリン接着分子の結合に関係する、接 着点プラーク中の薄膜細胞骨格蛋白質である。

【 0 2 7 2 】

補足データ

補足データ：

C a n c e r M a p データセット中のサンプルの臨床および分子の特性。この出願で提
供される情報を使用することを可能にするために、この表が分割されたことに注意。それ
ぞれのカラムは、列、サンプル I D 、ドナー I D 、バッチ、試料タイプ、抽出方法、セン
ター、E R G F I S H ステータス、腫瘍パーセンテージ、人種・民族、診断時の年齢、
病理学病期、病理学的サブステージ、P S A プレ-前立腺摘除、グリーソンスコア、マー
ジン、嚢外伸展 (Extra Capsular Extension)、B C R フリータイム月 (BCR FreeTime

months)、B C R イベントおよびI C G C カテゴリーを含む。

【 0 2 7 3 】

【表 1 5 - 1】

列	サンプルID	ドナーID	バッチ	試料タイプ	抽出方法	セクター	ERG FISH ヌーターヌ	腫瘍パーセント	人種・民族
1	TB08.0234_v1	TB08.0234	CamFinal	Normal	Qiagen miRNA kit	Cambridge		0	White - British
2	TB08.0234_v3	TB08.0234	CamFinal	Normal	Qiagen miRNA kit	Cambridge		0	White - British
3	TB08.0262_v3	TB08.0262	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	2N	75	White - British
4	TB08.0268_v3	TB08.0268	CamFinal	Low Tumour	Qiagen miRNA kit	Cambridge	2N	5	White - British
5	TB08.0271_v1	TB08.0271	CamFinal	Low Tumour	Qiagen miRNA kit	Cambridge	2N	10	White - British
6	TB08.0311_v2	TB08.0311	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	MixedEdel	33	White - British
7	TB08.0311_v3	TB08.0311	CamFinal	Low Tumour	Qiagen miRNA kit	Cambridge	Edel	10	White - British
8	TB08.0327_v1	TB08.0327	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	Edel	30	White - British
9	TB08.0341_v1	TB08.0341	CamFinal	Normal	Qiagen miRNA kit	Cambridge		0	White - British
10	TB08.0341_v5	TB08.0341	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	2N	25	White - British
11	TB08.0359_v16	TB08.0359	CamFinal	Normal	Qiagen miRNA kit	Cambridge		0	White - British
12	TB08.0359_v2	TB08.0359	CamFinal	Stroma	Qiagen miRNA kit	Cambridge	2N	0	White - British

【 0 2 7 4 】

10

20

30

40

50

【表 1 5 - 2】

13	TB08.0368_v14	TB08.0368	CamFinal	Normal	Qiagen miRNA kit	Cambridge	Esplitt	0	White - British
14	TB08.0429_v7	TB08.0429	CamFinal	Low Tumour	Qiagen miRNA kit	Cambridge		3	White - British
15	TB08.0489_v5	TB08.0489	CamFinal	Normal	Qiagen miRNA kit	Cambridge		0	White - British
16	TB08.0489_v13	TB08.0489	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	Esplitt	30	White - British
17	TB08.0501_v8	TB08.0501	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	2N	33	White - British
18	TB08.0519_v14	TB08.0519	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	Edel	75	Turkish
19	TB08.0533_v6	TB08.0533	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	2N	50	White - British
20	TB08.0588_v1	TB08.0588	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	MixedEsplitt	40	White - British
21	TB08.0589_v1	TB08.0589	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	2N	36	White - British
22	TB08.0589_v2	TB08.0589	CamFinal	Low Tumour	Qiagen miRNA kit	Cambridge	MixedPloidy	10	White - British
23	TB08.0589_v4	TB08.0589	CamFinal	Stroma	Qiagen miRNA kit	Cambridge	2N	0	White - British
24	TB08.0589_v5	TB08.0589	CamFinal	Low Tumour	Qiagen miRNA kit	Cambridge	MixedPloidy	8	White - British
25	TB08.0598_v12	TB08.0598	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	2N	45	White - British
26	TB08.0609_v11	TB08.0609	CamFinal	Low Tumour	Qiagen miRNA kit	Cambridge	MixedRearrangement	15	White - British
27	TB08.0667_v9	TB08.0667	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	2N	40	White - British
28	TB08.0667_v6	TB08.0667	CamFinal	Stroma	Qiagen miRNA kit	Cambridge		0	White - British
29	TB08.0689_v14	TB08.0689	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	MixedRearrangement	40	White - British
30	TB08.0689_v15	TB08.0689	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	MixedEdel	70	White - British
31	TB08.0689_v2	TB08.0689	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	Esplitt	21	White - British
32	TB08.0689_v8	TB08.0689	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	2N	33	White - British
33	TB08.0691_v13	TB08.0691	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	MixedEsplitt	50	White - British
34	TB08.0716_v18	TB08.0716	CamFinal	Stroma	Qiagen miRNA kit	Cambridge	2N	0	White - British
35	TB08.0719_v11	TB08.0719	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	2N	50	White - British
36	TB08.0731_v13	TB08.0731	CamFinal	Low Tumour	Qiagen miRNA kit	Cambridge	Esplitt	3	White - British
37	TB08.0816_v2	TB08.0816	CamFinal	Low Tumour	Qiagen miRNA kit	Cambridge	Edel	18	White - British

【 0 2 7 5 】

10

20

30

40

50

【表 1 5 - 3】

38	TB08.0817_v14	TB08.0817	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	MixedPloidy	34	White - British
39	TB08.0848_v10	TB08.0848	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	Esplitt	35	White - Other
40	TB08.0869_v4	TB08.0869	CamFinal	Low Tumour	Qiagen miRNA kit	Cambridge	MixedRearrangement	5	White - British
41	TB08.0869_v6	TB08.0869	CamFinal	Low Tumour	Qiagen miRNA kit	Cambridge	MixedEsplitt	15	White - British
42	TB08.0869_v7	TB08.0869	CamFinal	Low Tumour	Qiagen miRNA kit	Cambridge	MixedEsplitt	15	White - British
43	TB08.0870_v18	TB08.0870	CamFinal	Low Tumour	Qiagen miRNA kit	Cambridge	MixedPloidy	8	Black or Black British - Caribbean
44	TB08.0872_v2	TB08.0872	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	MixedRearrangement	20	White - Other
45	TB08.0877_v19	TB08.0877	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	Edel	40	White - British
46	TB08.0879_v11	TB08.0879	CamFinal	Low Tumour	Qiagen miRNA kit	Cambridge	Edel	5	White - British
47	TB08.0884_v2	TB08.0884	CamFinal	Normal	Qiagen miRNA kit	Cambridge	2N	0	White - British
48	TB08.0927_v5	TB08.0927	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	2N	20	White - British
49	TB08.0943_v7	TB08.0943	CamFinal	Stroma	Qiagen miRNA kit	Cambridge	2N	0	White - British
50	TB08.0958_v12	TB08.0958	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	2Edel	55	White - British
51	TB08.0958_v13	TB08.0958	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	MixedRearrangement	45	White - British
52	TB08.0973_v9	TB08.0973	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	2N	23	White - British
53	TB08.0978_v7	TB08.0978	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	MixedPloidy	20	White - British
54	TB08.0978_v8	TB08.0978	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	2N	45	White - British
55	TB08.0978_v9	TB08.0978	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	2N	29	White - British
56	TB08.0986_v2	TB08.0986	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	MixedEsplitt	38	White - British
57	TB08.0987_v6	TB08.0987	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	2N	49	White - British
58	TB08.0993_v12	TB08.0993	CamFinal	Low Tumour	Qiagen miRNA kit	Cambridge	MixedRearrangement	4	White - British
59	TB08.0997_v6	TB08.0997	CamFinal	Stroma	Qiagen miRNA kit	Cambridge		0	White - British
60	TB08.0999_v11	TB08.0999	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	2N	30	White - British
61	TB08.0999_v2	TB08.0999	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	MixedRearrangement	48	White - British
62	TB08.1015_v10	TB08.1015	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	MixedEdel	78	White - British

【表 1 5 - 4】

63	TB08.1015_v11	TB08.1015	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	MixedEdel	78	White - British
64	TB08.1015_v9	TB08.1015	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	MixedEdel	50	White - British
65	TB08.1019_v1	TB08.1019	CamFinal	Low Tumour	Qiagen miRNA kit	Cambridge	MixedRearrangement	10	White - British
66	TB08.1019_v14	TB08.1019	CamFinal	Low Tumour	Qiagen miRNA kit	Cambridge	2ESplit	10	White - British
67	TB08.1019_v15	TB08.1019	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	MixedRearrangement	20	White - British
68	TB08.1019_v2	TB08.1019	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	MixedRearrangement	30	White - British
69	TB08.1026_v17	TB08.1026	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	2N	78	White - British
70	TB08.1044_v7	TB08.1044	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	2N	40	White - British
71	TB08.1053_v5	TB08.1053	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	MixedRearrangement	48	White - British
72	TB08.1063_v16	TB08.1063	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	MixedRearrangement	50	White - British
73	TB08.1063_v8	TB08.1063	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	2N	31	White - British
74	TB08.1083_v3	TB08.1083	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	2ESplit	33	White - British
75	TB08.1116_v2	TB08.1116	CamFinal	Low Tumour	Qiagen miRNA kit	Cambridge	MixedRearrangement	15	White - British
76	TB08.1116_v3	TB08.1116	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	MixedESplit	56	White - British
77	TB08.1116_v9	TB08.1116	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	MixedRearrangement	30	White - British
78	TB08.1159_v2	TB08.1159	CamFinal	Normal	Qiagen miRNA kit	Cambridge	Edel	0	White - British
79	TB08.0601_v16	TB08.0601	CamFinal	Normal	Qiagen miRNA kit	Cambridge			White - British
80	TB09.0217_v16	TB09.0217	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	Edel	63	White - British
81	TB09.0217_v7	TB09.0217	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	MixedPloidy	28	White - British
82	TB09.0219_v13	TB09.0219	CamFinal	Low Tumour	Qiagen miRNA kit	Cambridge	2N	10	White - British
83	TB09.0219_v2	TB09.0219	CamFinal	Low Tumour	Qiagen miRNA kit	Cambridge	MixedRearrangement	11	White - British
84	TB09.0219_v21	TB09.0219	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	ESplit	57	White - British
85	TB09.0219_v8	TB09.0219	CamFinal	Low Tumour	Qiagen miRNA kit	Cambridge	2N	4	White - British
86	TB09.0238_v12	TB09.0238	CamFinal	Stroma	Qiagen miRNA kit	Cambridge	2N	0	White - British
87	TB09.0238_v18	TB09.0238	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	MixedRearrangement	50	White - British

【 0 2 7 7 】

10

20

30

40

50

【表 1 5 - 5】

88	TB09.0238_v5	TB09.0238	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	2N	25	White - British
89	TB09.0272_v6	TB09.0272	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	Esplit	65	White - British
90	TB09.0272_v7	TB09.0272	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	2N	35	White - British
91	TB09.0295_v2	TB09.0295	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	2N	70	White - British
92	TB09.0413_v11	TB09.0413	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	2N	68	Black or Black British - Caribbean
93	TB09.0413_v8	TB09.0413	CamFinal	Low Tumour	Qiagen miRNA kit	Cambridge	MixedPloidy	5	Black or Black British - Caribbean
94	TB09.0443_v3	TB09.0443	CamFinal	Low Tumour	Qiagen miRNA kit	Cambridge	Edel	2	White - British
95	TB09.0443_v8	TB09.0443	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	2N	65	White - British
96	TB09.0448_v8	TB09.0448	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	MixedPloidy	33	White - British
97	TB09.0462_v7	TB09.0462	CamFinal	Low Tumour	Qiagen miRNA kit	Cambridge	MixedEsplit	8	White - British
98	TB09.0471_v11	TB09.0471	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	Edel	20	White - British
99	TB09.0504_v4	TB09.0504	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	2N	50	White - British
100	TB09.0550_v15	TB09.0550	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	MixedEsplit	55	White - British
101	TB09.0606_v3	TB09.0606	CamFinal	Low Tumour	Qiagen miRNA kit	Cambridge	MixedPloidy	18	White - British
102	TB09.0706_v5	TB09.0706	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	Esplit	54	White - British
103	TB09.0720_v19	TB09.0720	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	Edel	23	White - British
104	TB09.0721_v14	TB09.0721	CamFinal	Low Tumour	Qiagen miRNA kit	Cambridge	MixedPloidy	10	White - British
105	TB09.0721_v15	TB09.0721	CamFinal	Low Tumour	Qiagen miRNA kit	Cambridge	RG	3	White - British
106	TB09.0725_v9	TB09.0725	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	2N	68	White - British
107	TB09.0774_v1	TB09.0774	CamFinal	Stroma	Qiagen miRNA kit	Cambridge	Esplit	0	White - British
108	TB09.0774_v15	TB09.0774	CamFinal	Low Tumour	Qiagen miRNA kit	Cambridge	2N	10	White - British
109	TB09.0850_v2	TB09.0850	CamFinal	Low Tumour	Qiagen miRNA kit	Cambridge	MixedEsplit	5	White - British
110	TB09.0962_v13	TB09.0962	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	MixedPloidy	23	White - British
111	TB09.0962_v16	TB09.0962	CamFinal	Tumour	Qiagen miRNA kit	Cambridge	Esplit	75	White - British
112	NP1	ICR_38	1208	Normal	Trizol	ICR	2N	0	White - British

【表 1 5 - 6】

113	NP10	ICR_47	309	Normal	Trizol	ICR	2N	0	White - British
114	NP11	ICR_50	309	Normal	Trizol	ICR	2N	0	White - British
115	NP12	ICR_58	309	Normal	Trizol	ICR	2N	0	White - British
116	NP14	ICR_35	309	Normal	Trizol	ICR	2N	0	White - British
117	NP15	ICR_65	309	Normal	Trizol	ICR	2N	0	White - British
118	NP16	ICR_69	309	Normal	Trizol	ICR	2N	0	Black or Black British - African
119	NP17	ICR_51	509	Normal	Trizol	ICR	2N	0	
120	NP18	ICR_66	509	Stroma	Trizol	ICR	2N	0	White - British
121	NP19	ICR_73	509	Stroma	Trizol	ICR	2N	0	White - British
122	NP2	ICR_37	1208	Normal	Trizol	ICR	2N	0	White - British
123	NP20	ICR_57	509	Normal	Trizol	ICR	2N	0	White - British
124	NP21	ICR_56	509	Stroma	Trizol	ICR	2N	0	White - British
125	NP4	ICR_47	1208	Normal	Trizol	ICR	3N	0	White - British
126	NP5	ICR_59	1208	Normal	Trizol	ICR	2N	0	White - British
127	NP8	ICR_34	309	Normal	Trizol	ICR	2N	0	White - British
128	NP9	ICR_54	309	Normal	Trizol	ICR	2N	0	White - Other
129	PRC140	ICR_20	509	Low Tumour	Trizol	ICR	Esplit	10	White - British
130	PRC101	ICR_28	908	Tumour	RNAeasyPlus	ICR	Edel	40	White - British
131	PRC102	ICR_44	908	Tumour	Trizol	ICR	2N	60	White - British
132	PRC103	ICR_34	908	Tumour	RNAeasyPlus	ICR	2N	20	White - British
133	PRC105	ICR_43	908	Tumour	RNAeasyPlus	ICR	2N	45	White - Other
134	PRC106	ICR_54	908	Low Tumour	RNAeasyPlus	ICR	2N	15	White - Other
135	PRC109	ICR_54	1008	Tumour	Trizol	ICR	Edel	60	White - British
136	PRC10	ICR_49	507	Tumour	Trizol	ICR	Edel		White - British
137	PRC110	ICR_22	1008	Tumour	Trizol	ICR	2Edel	55	White - British

【 0 2 7 9】

10

20

30

40

50

【表 1 5 - 7】

138	PRC111	ICR_49	1008	Tumour	Trizol	ICR	2N	20	White - British
139	PRC112	ICR_49	1008	Normal	Trizol	ICR	2N	0	White - Other
140	PRC113	ICR_60	1008	Tumour	Trizol	ICR	2N	70	White - British
141	PRC114	ICR_63	1008	Tumour	Trizol	ICR	2ESplit	40	White - British
142	PRC115	ICR_41	1008	Tumour	Trizol	ICR	2ESplit	30	White - British
143	PRC116	ICR_41	1008	Tumour	Trizol	ICR	MixedRearrangement	50	White - British
144	PRC117	ICR_17	1008	Tumour	Trizol	ICR	ESplit	20	White - British
145	PRC118	ICR_17	1008	Tumour	Trizol	ICR	2N	90	White - British
146	PRC119	ICR_50	1008	Tumour	Trizol	ICR	Edel	30	White - British
147	PRC11	ICR_59	507	Tumour	Trizol	ICR	Edel	60	White - British
148	PRC122	ICR_4	1008	Low Tumour	Trizol	ICR	ESplit	3	White - British
149	PRC123	ICR_17	1008	Low Tumour	Trizol	ICR	2N	5	
150	PRC124	ICR_40	1008	Tumour	Trizol	ICR	2N	20	White - British
151	PRC125	ICR_61	1208	Tumour	Trizol	ICR	2N	45	
152	PRC126	ICR_40	1208	Tumour	Trizol	ICR	2Edel	70	White - British
153	PRC127	ICR_48	1208	Tumour	Trizol	ICR	Edel	50	White - British
154	PRC128	ICR_48	1208	Low Tumour	Trizol	ICR	2ESplit	15	White - British
155	PRC129	ICR_55	1208	Tumour	Trizol	ICR	2ESplit	70	White - British
156	PRC12	ICR_55	507	Tumour	Trizol	ICR		85	White - British
157	PRC130	ICR_25	1208	Tumour	Trizol	ICR	2N	70	White - British
158	PRC133	ICR_58	309	Tumour	Trizol	ICR	MixedPloidy	90	White - British
159	PRC134	ICR_35	309	Normal	Trizol	ICR	MixedPloidy	0	White - British
160	PRC135	ICR_35	309	Tumour	Trizol	ICR	2ESplit	60	White - British
161	PRC136	ICR_68	309	Tumour	Trizol	ICR	MixedPloidy	70	White - British
162	PRC137	ICR_71	309	Tumour	Trizol	ICR	2N	30	White - British

【表 1 5 - 8】

163	PRC138	ICR_65	309	Tumour	Trizol	ICR	2N	60	Black or Black British - African
164	PRC139	ICR_69	309	Tumour	Trizol	ICR	2N	70	Black or Black British - African
165	PRC13	ICR_69	507	Tumour	Trizol	ICR	2Edel	25	White - British
166	PRC141	ICR_2	509	Tumour	Trizol	ICR	Edel	60	White - Other
167	PRC142	ICR_68	509	Normal	Trizol	ICR	2N	0	White - British
168	PRC143	ICR_67	509	Low Tumour	Trizol	ICR	Edel	5	White - British
169	PRC144	ICR_73	509	Tumour	Trizol	ICR	Edel	70	White - British
170	PRC145	ICR_57	509	Low Tumour	Trizol	ICR	NG	5	White - British
171	PRC146	ICR_45	ICRFinal	Low Tumour	Trizol	ICR	2N	2	White - British
172	PRC147	ICR_56	ICRFinal	Low Tumour	Trizol	ICR	MixedEdel	5	White - British
173	PRC148	ICR_70	ICRFinal	Tumour	Trizol	ICR	2N	35	White - British
174	PRC149	ICR_70	ICRFinal	Low Tumour	Trizol	ICR	MixedPloidy	5	White - British
175	PRC14	ICR_39	507	Normal	Trizol	ICR	2N	0	White - Other
176	PRC150	ICR_72	ICRFinal	Tumour	Trizol	ICR	Esplrit	30	White - British
177	PRC151	ICR_7	ICRFinal	Tumour	Trizol	ICR	2N	50	White - British
178	PRC152	ICR_53	ICRFinal	Low Tumour	Trizol	ICR	2N	15	White - British
179	PRC153	ICR_64	ICRFinal	Tumour	Trizol	ICR	2N	20	White - British
180	PRC154	ICR_33	ICRFinal	Tumour	Trizol	ICR	MixedPloidy	65	White - British
181	PRC155	ICR_33	ICRFinal	Tumour	Trizol	ICR	2N	65	White - British
182	PRC156	ICR_1	ICRFinal	Tumour	Trizol	ICR	Edel	50	White - Other
183	PRC157	ICR_62	ICRFinal	Tumour	Trizol	ICR	2N	85	White - British
184	PRC158	ICR_74	ICRFinal	Tumour	Trizol	ICR	MixedPloidy	70	White - British
185	PRC159	ICR_8	ICRFinal	Tumour	Trizol	ICR	4N	40	White - Other
186	PRC15	ICR_80	507	Normal	Trizol	ICR	2N	0	White - British
187	PRC160	ICR_79	ICRFinal	Tumour	Trizol	ICR	2N	75	White - Other

【表 1 5 - 9】

188	PRC161	ICR_23	ICRFinal	Tumour	Trizol	ICR	2N	60	White - British
189	PRC162	ICR_76	ICRFinal	Tumour	Trizol	ICR	Espllt	50	White - British
190	PRC163	ICR_80	ICRFinal	Tumour	Trizol	ICR	2N	50	White - British
191	PRC164	ICR_81	ICRFinal	Tumour	Trizol	ICR	Espllt	40	White - Irish
192	PRC165	ICR_73	ICRFinal	Tumour	Trizol	ICR	Edel	30	White - Other
193	PRC166	ICR_3	ICRFinal	Tumour	Trizol	ICR	Edel	65	White - British
194	PRC167	ICR_36	ICRFinal	Tumour	Trizol	ICR	Espllt	70	White - British
195	PRC168	ICR_19	ICRFinal	Tumour	Trizol	ICR	Edel	70	White - British
196	PRC169	ICR_78	ICRFinal	Low Tumour	Trizol	ICR	Espllt	10	White - British
197	PRC16	ICR_77	507	Normal	Trizol	ICR		0	White - British
198	PRC17	ICR_75	507	Low Tumour	Trizol	ICR	Espllt	10	White - British
199	PRC18	ICR_6	507	Tumour	Trizol	ICR			White - British
200	PRC19	ICR_25	507	Low Tumour	Trizol	ICR		5	White - British
201	PRC1	ICR_27	507	Tumour	Trizol	ICR	Edel	45	White - British
202	PRC20	ICR_2	507	Low Tumour	Trizol	ICR	Espllt	15	
203	PRC21	ICR_82	507	Low Tumour	Trizol	ICR	2Espllt	15	
204	PRC22	ICR_82	507	Normal	Trizol	ICR		0	White - British
205	PRC23	ICR_24	507	Normal	Trizol	ICR		0	White - British
206	PRC24	ICR_26	507	Tumour	Trizol	ICR	2Edel	30	White - British
207	PRC25	ICR_12	507	Tumour	Trizol	ICR	Edel	35	White - British
208	PRC26	ICR_29	507	Low Tumour	Trizol	ICR	2N	15	White - British
209	PRC27	ICR_30	407	Tumour	Trizol	ICR		50	Black or Black British - Caribbean
210	PRC28	ICR_13	407	Low Tumour	Trizol	ICR		5	White - British
211	PRC29	ICR_15	407	Low Tumour	Trizol	ICR	MixedPloidy	15	White - British
212	PRC2	ICR_18	507	Low Tumour	Trizol	ICR	Edel	10	White - Other

【表 1 5 - 1 0】

213	PRC30	ICR_7	407	Tumour	Trizol	ICR	Edel		White - British
214	PRC31	ICR_22	507	Low Tumour	Trizol	ICR		5	White - British
215	PRC32	ICR_14	507	Low Tumour	Trizol	ICR		5	White - Other
216	PRC34	ICR_21	407	Normal	Trizol	ICR	2N	0	White - Irish
217	PRC35	ICR_5	407	Normal	Trizol	ICR	Edel	0	White - Irish
218	PRC36	ICR_5	407	Low Tumour	Trizol	ICR	Edel	5	White - British
219	PRC38	ICR_12	407	Low Tumour	Trizol	ICR	Edel	15	White - British
220	PRC39	ICR_11	407	Low Tumour	Trizol	ICR		10	White - British
221	PRC3	ICR_32	507	Tumour	Trizol	ICR	Edel	50	White - British
222	PRC40	ICR_9	407	Tumour	Trizol	ICR	Edel	70	White - British
223	PRC42	ICR_20	407	Low Tumour	Trizol	ICR	Edel	5	White - British
224	PRC45	ICR_10	407	Normal	Trizol	ICR		0	White - British
225	PRC4	ICR_14	507	Tumour	Trizol	ICR	2Esplitt	25	White - British
226	PRC5	ICR_16	507	Low Tumour	Trizol	ICR	Esplitt	3	White - British
227	PRC6	ICR_23	507	Tumour	Trizol	ICR		80	White - British
228	PRC7	ICR_10	507	Tumour	Trizol	ICR		50	White - British
229	PRC8	ICR_23	507	Tumour	Trizol	ICR		80	White - British
230	PRC9	ICR_31	507	Tumour	Trizol	ICR		30	White - British
231	ST1	ICR_48	1208	Stroma	Trizol	ICR	Edel	0	White - British
232	ST2	ICR_46	ICRFinal	Stroma	Trizol	ICR	2N	0	White - British
233	ST3	ICR_52	ICRFinal	Stroma	Trizol	ICR	2N	0	White - British
234	ST4	ICR_66	ICRFinal	Stroma	Trizol	ICR	2N	0	White - British
235	ST5	ICR_76	ICRFinal	Stroma	Trizol	ICR	2N	0	White - Other

【 0 2 8 3】

上記の表中、マージンの欄における「Normal」、「Tumour」、「Low Tumour」、「Stroma」は、それぞれ「正常」、「腫瘍有」、「少量の腫瘍」、「間質」であり、センターの欄における「Cambridge」は、「ケンブリッジ」であり、抽出方法の欄における「Trizol」は「トリゾール」であり、E R G F I S Hステータスの欄における「MixedEdel」、「MixedEsplitt」、「MixedPloidy」、「MixedRearrangement」は、それぞれ「混合Edel」、「混合Esplitt」、「混合Ploidy」、「混合リアレンジメント」であり、人種・民族の欄の「White - British」、「Turkish」、「Black or Black British - Caribbean」、「White - Other」、「Black or Black British - African」、「White - Ir

ish」はそれぞれ、「白人 - 英国人」、「トルコ人」、「黒人または黒人英国人 - カリブ人」、「白人 - その他」、「黒人または黒人英国人 - アフリカ人」、「白人 - アイルランド人」である。

【 0 2 8 4 】

【表 1 6 - 1】

列	サンプルID	診断時の年齢	病理学病期	病理学的サブステージ	PSAレベル - 前立腺摘除	グリーensonスコア	マージン	嚢外伸展	BCR フリータイム月	BCRイベント	ICG Cカテゴリー
1	TB08.0234_v1	64	T2	b	5.80	3+5	negative margins	N	66.00	FALSE	normal
2	TB08.0234_v3	64	T2	b	5.80	3+5	negative margins	N	66.00	FALSE	normal
3	TB08.0262_v3	69	T3	a	8.30	3+4	Positive circumferential	Y	65.00	FALSE	cat_1
4	TB08.0268_v3	56	T3	a	8.70	3+4	Positive circumferential	Y	59.00	FALSE	cat_1
5	TB08.0271_v1	74	T2	x	15.40	3+4	Negative	N	73.00	FALSE	cat_2
6	TB08.0311_v2	69	T3	a	15.30	3+4	Positive base	Y	64.00	FALSE	cat_2
7	TB08.0311_v3	69	T3	a	15.30	3+4	Positive base	Y	64.00	FALSE	cat_2
8	TB08.0327_v1	57	T2	x	4.80	3+4	Negative	N	64.00	FALSE	cat_1
9	TB08.0341_v1	57	T2	x	5.10	3+4	negative margins	N	6.00	TRUE	normal
10	TB08.0341_v5	57	T2	x	5.10	3+4	negative margins	N	6.00	TRUE	cat_1
11	TB08.0359_v16	63	T2	a	9.90	3+4	positive Apex margin	N	2.00	TRUE	normal
12	TB08.0359_v2	63	T2	a	9.90	3+4	positive Apex margin	N	2.00	TRUE	cat_1
13	TB08.0368_v14	71	T3	b		4+3	positive apex & circumferential margin	Y	34.00	FALSE	unknown
14	TB08.0429_v7	72	T3	b	9.20	3+4	positive circumferential margin	Y	3.00	FALSE	normal
15	TB08.0489_v5	62	T3	a	5.30	4+3	negative margins	Y	49.00	FALSE	normal
16	TB08.0489_v13	62	T3	a	5.30	4+3	negative margins	Y	49.00	FALSE	cat_2
17	TB08.0501_v8	64	T3	a	20.50	3+4	Negative	Y	76.00	FALSE	cat_3
18	TB08.0519_v14	55	T4	x	9.80	5+4	positive apex				cat_3
19	TB08.0533_v6	65	T3	a	5.80	3+4	negative margins	Y	62.00	FALSE	cat_1

【 0 2 8 5 】

10

20

30

40

50

【表 1 6 - 2】

20	TB08.0588_v1	55	T3	a	13.90	3+4	Negative	Y	55.00	TRUE	cat_2
21	TB08.0589_v1	66	T4	x	5.17	5+4	Positive base	Y	2.00	TRUE	cat_3
22	TB08.0589_v2	66	T4	x	5.17	5+4	Positive base	Y	2.00	TRUE	cat_3
23	TB08.0589_v4	66	T4	x	5.17	5+4	Positive base	Y	2.00	TRUE	cat_3
24	TB08.0589_v5	66	T4	x	5.17	5+4	Positive base	Y	2.00	TRUE	cat_3
25	TB08.0598_v12	65	T2	x	8.80	3+4	negative margins	N	61.00	FALSE	cat_1
26	TB08.0609_v11	66	T2	x	11.40	4+3	Negative	N	34.00	FALSE	cat_2
27	TB08.0667_v9	57	T2	x	7.80	3+3	negative margins	N	42.00	FALSE	cat_1
28	TB08.0667_v6	57	T2	x	7.80	3+3	negative margins	N	42.00	FALSE	normal
29	TB08.0689_v14	51	T2	x	8.80	3+3	Negative	N	60.00	FALSE	cat_1
30	TB08.0689_v15	51	T2	x	8.80	3+3	Negative	N	60.00	FALSE	cat_1
31	TB08.0689_v2	51	T2	x	8.80	3+3	Negative	N	60.00	FALSE	cat_1
32	TB08.0689_v8	51	T2	x	8.80	3+3	Negative	N	60.00	FALSE	cat_1
33	TB08.0691_v13	69	T3	a	9.40	3+4	Negative	Y	8.00	TRUE	cat_1
34	TB08.0716_v18	64	T3	a	8.90	3+4	negative margins	Y	60.00	FALSE	cat_1
35	TB08.0719_v11	62	T2	x	6.50	3+3	positive circumferential margin	N	60.00	FALSE	cat_1
36	TB08.0731_v13	59	T3	a	7.90	3+4	Negative	Y	61.00	FALSE	cat_1
37	TB08.0816_v2	63	T3	a	10.40	3+4	negative margins	Y	60.00	FALSE	cat_2
38	TB08.0817_v14	62	T3	a	10.40	3+4	Negative	Y	24.00	TRUE	cat_2
39	TB08.0848_v10	63	T3	a	4.90	4+3	negative margins	Y	55.00	FALSE	cat_2
40	TB08.0869_v4	58	T2	x	40.00	3+3	Negative	N	19.00	TRUE	cat_3
41	TB08.0869_v6	58	T2	x	40.00	3+3	Negative	N	19.00	TRUE	cat_3
42	TB08.0869_v7	58	T2	x	40.00	3+3	Negative	N	19.00	TRUE	cat_3
43	TB08.0870_v18	71	T3	b	8.20	3+4	Negative	Y	1.00	TRUE	cat_1
44	TB08.0872_v2	63	T2	c	7.50	3+3	Negative	N	56.00	FALSE	cat_1

【表 1 6 - 3】

45	TB08.0877_v19	61	T2	x	8.70	3+3	Negative	N	49.00	FALSE	cat_1
46	TB08.0879_v11	62	T3	a	8.40	4+3	Negative	Y	60.00	FALSE	cat_2
47	TB08.0884_v2	46	T2	x	2.40	3+4	negative margins	N	61.00	FALSE	cat_1
48	TB08.0927_v5	59	T2	c	9.30	3+3	positive Apex margin	N	59.00	FALSE	cat_1
49	TB08.0943_v7	56	T3	a	3.40	3+4	positive base margin	Y	53.00	FALSE	cat_1
50	TB08.0958_v12	42	T2	x	11.80	3+3	Negative	N	43.00	FALSE	cat_2
51	TB08.0958_v13	42	T2	x	11.80	3+3	Negative	N	43.00	FALSE	cat_2
52	TB08.0973_v9	68	T2	c	6.40	3+4	negative margins	N	58.00	FALSE	cat_1
53	TB08.0978_v7	64	T3	a	12.00	3+4	Negative	Y	58.00	FALSE	cat_2
54	TB08.0978_v8	64	T3	a	12.00	3+4	Negative	Y	58.00	FALSE	cat_2
55	TB08.0978_v9	64	T3	a	12.00	3+4	Negative	Y	58.00	FALSE	cat_2
56	TB08.0986_v2	56	T3	a	15.50	3+4	Positive base	Y	58.00	FALSE	cat_2
57	TB08.0987_v6	54	T3	a	12.00	3+4	positive circumferential margin	Y	58.00	FALSE	cat_2
58	TB08.0993_v12	66	T2	c	7.70	4+3	Negative	N	60.00	FALSE	cat_2
59	TB08.0997_v6	62	T3	a	7.00	4+3	positive Apex margin	Y	58.00	FALSE	normal
60	TB08.0999_v11	67	T3	a	9.20	3+4	Negative	Y	52.00	FALSE	cat_1
61	TB08.0999_v2	67	T3	a	9.20	3+4	Negative	Y	52.00	FALSE	cat_1
62	TB08.1015_v10	63	T3	a	8.00	3+5	Negative	Y	12.00	TRUE	cat_3
63	TB08.1015_v11	63	T3	a	8.00	3+5	Negative	Y	12.00	TRUE	cat_3
64	TB08.1015_v9	63	T3	a	8.00	3+5	Negative	Y	12.00	TRUE	cat_3
65	TB08.1019_v1	59	T3	a	5.00	3+4	Negative	Y	68.00	FALSE	cat_1
66	TB08.1019_v14	59	T3	a	5.00	3+4	Negative	Y	68.00	FALSE	cat_1
67	TB08.1019_v15	59	T3	a	5.00	3+4	Negative	Y	68.00	FALSE	cat_1
68	TB08.1019_v2	59	T3	a	5.00	3+4	Negative	Y	68.00	FALSE	cat_1
69	TB08.1026_v17	61	T3	a	8.40	3+4	negative margins	Y	57.00	TRUE	cat_1

【表 1 6 - 4】

70	TB08.1044_v7	71	T3	a	7.90	3+4	Positive base & circumferential	Y	59.00	FALSE	cat_1
71	TB08.1053_v5	71	T3	a	17.00	3+4	Negative	Y	57.00	FALSE	cat_2
72	TB08.1063_v16	67	T3	a	5.80	4+3	Positive base & circumferential	Y	38.00	TRUE	cat_2
73	TB08.1063_v8	67	T3	a	5.80	4+3	Positive base & circumferential	Y	38.00	TRUE	cat_2
74	TB08.1083_v3	64	T3	a	7.30	3+3	negative margins	Y	57.00	FALSE	cat_1
75	TB08.1116_v2	61	T3	a	6.00	3+4	Negative	Y	42.00	FALSE	cat_1
76	TB08.1116_v3	61	T3	a	6.00	3+4	Negative	Y	42.00	FALSE	cat_1
77	TB08.1116_v9	61	T3	a	6.00	3+4	Negative	Y	42.00	FALSE	cat_1
78	TB08.1159_v2	56	T2	a	7.90	3+3	negative margins	N	56.00	FALSE	cat_1
79	TB08.0601_v16	66	T2	x	8.28	3+3	negative margins	N	68.00	FALSE	normal
80	TB09.0217_v16	63	T3	a	11.50	3+4	Positive base	Y	12.00	TRUE	cat_2
81	TB09.0217_v7	63	T3	a	11.50	3+4	Positive base	Y	12.00	TRUE	cat_2
82	TB09.0219_v13	62	T3	a	17.30	3+4	Negative	Y	16.00	TRUE	cat_2
83	TB09.0219_v2	62	T3	a	17.30	3+4	Negative	Y	16.00	TRUE	cat_2
84	TB09.0219_v21	62	T3	a	17.30	3+4	Negative	Y	16.00	TRUE	cat_2
85	TB09.0219_v8	62	T3	a	17.30	3+4	Negative	Y	16.00	TRUE	cat_2
86	TB09.0238_v12	66	T3	a	9.60	3+4	Negative	Y	54.00	FALSE	cat_1
87	TB09.0238_v18	66	T3	a	9.60	3+4	Negative	Y	54.00	FALSE	cat_1
88	TB09.0238_v5	66	T3	a	9.60	3+4	Negative	Y	54.00	FALSE	cat_1
89	TB09.0272_v6	62	T3	a	12.00	3+4	Negative	Y	58.00	FALSE	cat_2
90	TB09.0272_v7	62	T3	a	12.00	3+4	Negative	Y	58.00	FALSE	cat_2
91	TB09.0295_v2	64	T3	b	22.60	3+4	positive apex				cat_3
92	TB09.0413_v11	48	T3	a	5.30	4+3	Negative	Y	45.00	TRUE	cat_2
93	TB09.0413_v8	48	T3	a	5.30	4+3	Negative	Y	45.00	TRUE	cat_2
94	TB09.0443_v3	41	T3	a	16.20	3+4	Negative	Y	51.00	FALSE	cat_2

【表 1 6 - 5】

95	TB09.0443_v8	41	T3	a	16.20	3+4	Negative		Y	51.00	FALSE	cat_2
96	TB09.0448_v8	70	T2	c	4.68	3+4	Negative		N	19.00	TRUE	cat_1
97	TB09.0462_v7	56	T3	a	5.80	3+4	Negative		Y	57.00	FALSE	cat_1
98	TB09.0471_v11	54	T2	c	5.80	3+3	negative margins		N	54.00	FALSE	cat_1
99	TB09.0504_v4	60	T2	a	5.10	3+5	Negative		N	51.00	FALSE	cat_3
100	TB09.0550_v15	47	T3	a	11.50	3+4	Negative		Y	51.00	FALSE	cat_2
101	TB09.0606_v3	64	T3	b	10.00	4+5	Negative		Y	15.00	TRUE	cat_3
102	TB09.0706_v5	63	T3	a	7.30	3+4	Negative		Y	17.00	TRUE	cat_1
103	TB09.0720_v19	67	T2	x	8.90	3+3	negative margins		N	50.00	FALSE	cat_1
104	TB09.0721_v14	58	T2	c	4.00	3+3	Negative		N	10.00	TRUE	cat_1
105	TB09.0721_v15	58	T2	c	4.00	3+3	Negative		N	10.00	TRUE	cat_1
106	TB09.0725_v9	64	T2	x	10.70	3+4	negative margins		N	49.00	FALSE	cat_2
107	TB09.0774_v1	64	T2	c	6.40	3+4	Negative		N	49.00	FALSE	cat_1
108	TB09.0774_v15	64	T2	c	6.40	3+4	Negative		N	49.00	FALSE	cat_1
109	TB09.0850_v2	21	T3	a	5.70	3+4	Negative		Y	56.00	FALSE	cat_1
110	TB09.0962_v13	65	T2	x	6.20	3+3	Negative		N	48.00	FALSE	cat_1
111	TB09.0962_v16	65	T2	x	6.20	3+3	Negative		N	48.00	FALSE	cat_1
112	NP1	66	T2	a	9.80	3+3	Negative		No	95.00	FALSE	cat_1
113	NP10	60	T3	b	15.00	4+3	Positive circumferential		Unknown	55.00	FALSE	normal
114	NP11	61	T2	c	6.90	3+4	Negative		No	72.00	FALSE	normal
115	NP12	65	T2	c	7.40	4+3	Negative		No	15.00	TRUE	normal
116	NP14	72	T2	x	11.10	3+4	Positive circumferential		Unknown	59.00	FALSE	normal
117	NP15	64	T2	c	6.10	3+4	Positive apex		No	51.00	FALSE	normal
118	NP16	53	T2	c	11.10	3+3	Negative		No	48.00	FALSE	normal
119	NP17	58	T2	c	4.70	3+4	Positive circumferential		No	60.00	FALSE	cat_1

【表 1 6 - 6】

120	NP18	60	T2	c	16.90	3+4	Positive circumferential	No	69.00	FALSE	normal
121	NP19	60	T3	b	8.00	4+5	Positive circumferential	Yes	6.00	TRUE	cat_3
122	NP2	53	T3	a		3+4	Positive circumferential	Yes	76.00	FALSE	unknown
123	NP20	68	T2	c	7.10	3+4	Negative	No	61.00	FALSE	normal
124	NP21	66	T2	c	6.10	3+4		No	63.00	FALSE	normal
125	NP4	60	T3	b	15.00	4+3	Positive circumferential	Unknown	55.00	FALSE	cat_2
126	NP5	61	T2	c	3.10	3+3	Negative	No	68.00	FALSE	cat_1
127	NP8	59	T2	c	7.70	3+4	Negative	No	55.00	FALSE	normal
128	NP9	62	T2	c	7.60	3+5	Negative	Unknown	68.00	FALSE	normal
129	PRC140	71	T2	b	6.30	4+3	Negative	No	57.00	FALSE	cat_2
130	PRC101	55	T2	c	4.75	3+4	Negative	Unknown	64.00	FALSE	cat_1
131	PRC102	59	T2	c	7.70	3+4	Negative	No	55.00	FALSE	cat_1
132	PRC103	61	T2	c	4.00	3+4	Negative	No	72.00	FALSE	cat_1
133	PRC105	62	T2	c	7.60	3+5	Negative	Unknown	68.00	FALSE	cat_3
134	PRC106	62	T2	c	7.60	3+5	Negative	Unknown	68.00	FALSE	cat_3
135	PRC109	62	T3	b	12.40	3+4	Positive apex & circumferential	No	47.00	FALSE	cat_2
136	PRC10	58	T2	c	6.60	3+3		Unknown	3.00	TRUE	cat_1
137	PRC110	62	T3	b	12.40	3+4	Positive apex & circumferential	No	47.00	FALSE	cat_2
138	PRC111	62	T3	b	12.40	3+4	Positive apex & circumferential	No	47.00	FALSE	cat_2
139	PRC112	68	T2	c	6.40	3+4	Positive apex	No	39.00	FALSE	cat_1
140	PRC113	49	T2	c	8.90	3+4	Negative	No	43.00	FALSE	cat_1
141	PRC114	40	T2	c	8.40	3+4	Negative	No	16.00	FALSE	cat_1
142	PRC115	40	T2	c	8.40	3+4	Negative	No	16.00	FALSE	cat_1
143	PRC116	61	T2	c	7.90	3+4	Negative	Unknown	69.00	FALSE	cat_1
144	PRC117	61	T2	c	7.90	3+4	Negative	Unknown	69.00	FALSE	cat_1

【 0 2 9 0】

10

20

30

40

50

【表 1 6 - 7】

145	PRC118	61	T2	c	6.90	3+4	Negative	No	72.00	FALSE	cat_1
146	PRC119	61	T2	c	3.10	3+3	Negative	No	68.00	FALSE	cat_1
147	PRC11	58	T2	x	4.10	3+3	Positive circumferential	No	93.00	FALSE	cat_1
148	PRC122	61	T2	c	7.90	3+4	Negative	Unknown	69.00	FALSE	cat_1
149	PRC123	55	T3	a	3.30	3+4	Positive complex	Yes	71.00	FALSE	cat_2
150	PRC124	61	T3	a	6.40	3+4	Negative	No	49.00	FALSE	cat_1
151	PRC125	55	T3	a	3.30	3+4	Positive complex	Yes	71.00	FALSE	cat_2
152	PRC126	72	T3	b		4+5	Positive apex & circumferential & base	Yes	56.00	TRUE	unknown
153	PRC127	72	T3	b		4+5	Positive apex & circumferential & base	Yes	56.00	TRUE	unknown
154	PRC128	70	T3	a	4.70	4+3	Positive circumferential	Yes	60.00	FALSE	cat_3
155	PRC129	70	T3	a	4.70	4+3	Positive circumferential	Yes	60.00	FALSE	cat_3
156	PRC12	63	T3	a	13.70	4+3	Positive circumferential	Yes	26.00	TRUE	cat_3
157	PRC130	65	T2	c	7.40	4+3	Negative	No	15.00	TRUE	cat_2
158	PRC133	72	T2	x	11.10	3+4	Positive circumferential	Unknown	59.00	FALSE	cat_2
159	PRC134	72	T2	x	11.10	3+4	Positive circumferential	Unknown	59.00	FALSE	cat_2
160	PRC135	71	T2	b	6.30	4+3	Negative	No	57.00	FALSE	cat_2
161	PRC136	51	T2	c	8.90	4+3	Positive circumferential	No	60.00	FALSE	cat_2
162	PRC137	64	T2	c	6.10	3+4	Positive apex	No	51.00	FALSE	cat_1
163	PRC138	53	T2	c	11.10	3+3	Negative	No	48.00	FALSE	cat_2
164	PRC139	53	T2	c	11.10	3+3	Negative	No	48.00	FALSE	cat_2
165	PRC13	63	T3	b	13.00	4+3	Positive complex	Unknown	9.00	TRUE	cat_2
166	PRC141	64	T2	c	15.20	3+4	Positive circumferential	No	27.00	FALSE	cat_2
167	PRC142	60	T3	b	8.00	4+5	Positive circumferential	Yes	6.00	TRUE	cat_3
168	PRC143	68	T2	c	7.10	3+4	Negative	No	61.00	FALSE	cat_1
169	PRC144	61	T2	c	7.80	3+4	Negative	No	23.00	TRUE	cat_1

【表 1 6 - 8】

170	PRC145	66	T2	c	6.10	3+4		No	63.00	FALSE	cat_1
171	PRC146	63	T2	c	5.60	3+4	Negative	No	55.00	FALSE	cat_1
172	PRC147	63	T2	c	5.60	3+4	Negative	No	55.00	FALSE	cat_1
173	PRC148	66	T2	c	6.70	3+4	Positive apex	No	66.00	FALSE	cat_1
174	PRC149	63	T2	c	11.50	4+3	Negative	No	53.00	FALSE	cat_2
175	PRC14	53	T2	x	8.00	3+3		Unknown	108.00	FALSE	normal
176	PRC150	50	T2	c	4.40	3+4	Negative	No	61.00	FALSE	cat_1
177	PRC151	56	T3	a	7.70	3+4	Negative	Yes	54.00	FALSE	cat_2
178	PRC152	58	T2	c	9.60	3+3	Negative	No	72.00	FALSE	cat_1
179	PRC153	58	T2	c	9.60	3+3	Negative	No	72.00	FALSE	cat_1
180	PRC154	69	T2	c	4.53	3+4	Positive apex	No	7.00	TRUE	cat_1
181	PRC155	61	T2	c	7.50	3+3	Negative	Unknown	70.00	FALSE	cat_1
182	PRC156	50	T3	a	3.60	3+4		Yes	44.00	FALSE	cat_2
183	PRC157	56	T1	c	10.50	3+3	Negative	No	34.00	FALSE	cat_2
184	PRC158	55	T3	a	7.00	4+5	Positive circumferential	Yes	44.00	TRUE	cat_3
185	PRC159	60	T3	a	5.40	4+3	Negative	Yes	32.00	FALSE	cat_3
186	PRC15	50	T3	b	16.20	3+4	Positive apex & circumferential	Unknown	4.00	TRUE	normal
187	PRC160	64	T3	b	7.20	4+3	Negative	No	34.00	FALSE	cat_2
188	PRC161	55	T3	a	7.00	4+5	Positive circumferential	Yes	44.00	TRUE	cat_3
189	PRC162	56	T3	b	9.28	3+4	Negative	Unknown	15.00	TRUE	cat_1
190	PRC163	60	T3	b	8.00	4+5	Positive circumferential	Yes	6.00	TRUE	cat_3
191	PRC164	62	T2	c	17.40	3+4	Negative	No	65.00	FALSE	cat_2
192	PRC165	64	T2	c	12.90	3+4		Unknown	9.00	TRUE	cat_2
193	PRC166	55	T2	c	12.40	3+4	Positive circumferential	No	73.00	FALSE	cat_2
194	PRC167	64	T3	a	3.80	3+4	Negative	Yes	67.00	FALSE	cat_2

【表 1 6 - 9】

195	PRC168	59	T2	c	8.70	3+3	Negative	No	39.00	FALSE	cat_1
196	PRC169	70	T2	c	8.10	3+4	Negative	No	32.00	FALSE	cat_1
197	PRC16	67	T3	a	16.00	3+3	Positive complex	Yes	66.00	TRUE	normal
198	PRC17	49	T3	b	7.50	3+4	Positive apex & circumferential	Unknown	17.00	TRUE	cat_1
199	PRC18	63	T3	b	13.00	4+3	Positive complex	Unknown	9.00	TRUE	cat_2
200	PRC19	65	T3	b	9.70	4+3	Positive complex	Unknown	45.00	TRUE	cat_2
201	PRC1	61	T2	c	9.30	3+3	Negative	No	110.00	FALSE	cat_1
202	PRC20	65	T2	x	9.80	3+4			56.00	FALSE	cat_1
203	PRC21	65	T2	x	9.80	3+4			56.00	FALSE	cat_1
204	PRC22	57	T2	c	7.10	3+4	Positive apex	No	83.00	FALSE	cat_1
205	PRC23	69	T2	x	5.60	3+4	Positive circumferential	Unknown	80.00	FALSE	cat_1
206	PRC24	56	T2	a	7.90	3+3	Positive complex	No	92.00	FALSE	cat_1
207	PRC25	58	T2	c	5.60	3+3	Positive apex & circumferential	No	19.00	TRUE	cat_1
208	PRC26	52	T2	c	3.40	3+3	Negative	No	94.00	FALSE	cat_1
209	PRC27	56	T3	b	8.00	3+3	Positive complex	No	98.00	FALSE	cat_1
210	PRC28	48	T2	c	3.70	3+3	Positive complex	Unknown	31.00	TRUE	cat_1
211	PRC29	64	T2	x	5.60	3+4	Positive complex	Unknown	90.00	FALSE	cat_1
212	PRC2	53	T2	x	8.00	3+3		Unknown	108.00	FALSE	cat_1
213	PRC30	58	T2	x	4.10	3+3	Positive circumferential	No	93.00	FALSE	cat_1
214	PRC31	56	T2	a	12.80	3+3	Negative	No	108.00	FALSE	cat_2
215	PRC32	52	T2	x	6.10	3+3	Positive complex	No	91.00	FALSE	cat_1
216	PRC34	56	T2	x	4.50	3+3	Positive apex & circumferential	Unknown	129.00	FALSE	cat_1
217	PRC35	56	T2	x	4.50	3+3	Positive apex & circumferential	Unknown	129.00	FALSE	cat_1
218	PRC36	56	T2	a	7.90	3+3	Positive complex	No	92.00	FALSE	cat_1
219	PRC38	55	T2	c	5.70	3+3	Positive base	Unknown	89.00	FALSE	cat_1

【表 1 6 - 1 0】

220	PRC39	62	T3	b	22.30	3+4	Positive apex	No	84.00	FALSE	cat_3
221	PRC3	50	T3	b	16.20	3+4	Positive apex & circumferential	Unknown	4.00	TRUE	cat_2
222	PRC40	61	T2	c	9.30	3+3	Negative	No	110.00	FALSE	cat_1
223	PRC42	68	T2	x	9.80	3+3	Positive circumferential	Unknown	39.00	TRUE	cat_1
224	PRC45	56	T2	a	12.80	3+3	Negative	No	108.00	FALSE	cat_2
225	PRC4	54	T3	a	11.40	3+3	Negative	Yes	123.00	FALSE	cat_2
226	PRC5	41	T2	x	4.00	3+3	Negative	Unknown	74.00	FALSE	cat_1
227	PRC6	67	T3	a	16.00	3+3	Positive complex	Yes	66.00	TRUE	cat_2
228	PRC7	68	T2	x	9.80	3+3	Positive circumferential	Unknown	39.00	TRUE	cat_1
229	PRC8	67	T3	a	16.00	3+3	Positive complex	Yes	66.00	TRUE	cat_2
230	PRC9	67	T2	c	13.90	4+5	Negative	No	11.00	TRUE	cat_3
231	ST1	72	T3	b		4+5	Positive apex & circumferential & base	Yes	56.00	TRUE	unknown
232	ST2	63	T2	b	4.78	3+4	Positive circumferential	No	67.00	FALSE	cat_1
233	ST3	63	T2	c	5.00	3+3	Negative	No	60.00	FALSE	cat_1
234	ST4	60	T2	c	16.90	3+4	Positive circumferential	No	69.00	FALSE	cat_2
235	ST5	64	T3	b	7.20	4+3	Negative	No	34.00	FALSE	normal

【 0 2 9 4】

上記の表中、マージンの欄における、「negative margins」、「Positive circumferential」、「Negative」、「Positive base」、「positive Apex margin」、「positive apex & circumferential margin」、「positive circumferential margin」、「positive apex」、「positive base margin」、「Positive base & circumferential」、「Positive complex」は、それぞれ「ネガティブマージン」、「ポジティブサーカムフェレンシャル」、「ネガティブ」、「ポジティブベース」、「ポジティブアペックスマージン」、「ポジティブアペックスアンドサーカムフェレンシャルマージン」、「ポジティブサーカムフェレンシャルマージン」、「ポジティブアペック」、「ポジ

「ポジティブベースマージン」、「ポジティブベースアンドサーカムフェレンシャル」、「ポジティブコンプレックス」であり、囊外伸展の欄の「Unknown」は、「未知」であり、BCRイベントの欄の「FALSE」、「TRUE」、はそれぞれ「偽」、「真」であり、ICGCカテゴリーの欄の、「normal」、「unknown」はそれぞれ、「正常」、「未知」である。

本発明は以下の実施態様を含む。

実施態様第1項：以下を含む、癌を分類する方法、または癌の進行を予測する方法：

- a) 患者から得られたサンプル中の複数の遺伝子の発現ステータスを決定し、患者の発現プロファイルを提供すること；
- b) 統計のベイズのクラスタリング分析あるいは他のクラスタリング分析を、患者の発現プロファイルおよび異なる患者からの同じ複数の遺伝子についての参照データセットについて行うこと；
- c) 任意に複数回、分析ステップb)を繰り返すこと；および
- d) 癌を分類すること、または癌の進行を予測すること。

10

実施態様第2項：前記ステップb)が少なくとも2回、少なくとも3回、少なくとも5回、少なくとも20回、少なくとも50回、あるいは少なくとも100回繰り返される、実施態様第1項記載の方法。

実施態様第3項：異なるランダムなシードが各クラスタリング分析に使用される、実施態様第2項記載の方法。

実施態様第4項：前記の複数の遺伝子の発現ステータスを決定することが、複数の遺伝子の発現レベルを決定することを含む、実施態様第1から3のいずれか1項記載の方法。

20

実施態様第5項：少なくとも50、少なくとも100、少なくとも200あるいは最も好ましくは少なくとも500の遺伝子について発現ステータスが決定される、実施態様第1から4のいずれか1項記載の方法。

実施態様第6項：統計分析を行なう前に、参照データセットに対して患者の発現プロファイルを正規化することをさらに含む、実施態様第1から5のいずれか1項記載の方法。

実施態様第7項：ステップa)の遺伝子が表1にリストされた遺伝子から選ばれる、実施態様第1から6のいずれか1項記載の方法。

実施態様第8項：前記ステップa)は、少なくとも20、少なくとも50、少なくとも100、少なくとも200、少なくとも500または少なくとも1000の遺伝子の発現ステータスを決定することを含む、実施態様第1から7のいずれか1項記載の方法。

30

実施態様第9項：前記ステップa)は、表1にリストされた遺伝子から選ばれる少なくとも100の遺伝子の発現ステータスを決定することを含む、実施態様第1から8のいずれか1項記載の方法。

実施態様第10項：前記ステップa)は、表1にリストされたすべての500の遺伝子の発現ステータスを決定することを含む、実施態様第1から9のいずれか1項記載の方法。

実施態様第11項：その発現ステータスが統計分析のために決定されている遺伝子の部分集合を選択するステップをさらに含む、実施態様第1から10のいずれか1項記載の方法。

実施態様第12項：遺伝子の部分集合内のそれぞれの遺伝子の発現ステータスが、癌患者のサンプルにより変化することが知られている、実施態様第11項記載の方法。

40

実施態様第13項：統計分析の前に、患者の発現プロファイルにユニークなラベルを付することをさらに含む、実施態様第1から12のいずれか1項記載の方法。

実施態様第14項：前記ステップd)が、患者の癌サンプルのDESENTステータスを決定することを含む、実施態様第1から13のいずれか1項記載の方法。

実施態様第15項：前記ステップd)が、患者の癌サンプルのDESENTステータスを決定することを含む、実施態様第1から14のいずれか1項記載の方法。

実施態様第16項：該参照データセット内のそれぞれの発現プロファイルのDESENTステータスが知られている、実施態様第1から15のいずれか1項記載の方法。

実施態様第17項：該患者の発現プロファイルが、統計分析の前に少なくとも2つの参照データセットと組み合わせられる、実施態様第1から16のいずれか1項記載の方法。

50

実施態様第 18 項：統計分析が L P D 分析である、実施態様第 1 から 17 のいずれか 1 項記載の方法。

実施態様第 19 項：L P D 分析が個々の患者の発現プロファイルをグループに分ける、実施態様第 18 項記載の方法。

実施態様第 20 項：個々の患者の発現プロファイルをグループに分けることが、それぞれの発現プロファイルについて、L D P 分析を使用して、個々の患者の発現プロファイルについて全体の発現プロファイルへの各グループの貢献 (p_i) を決定することを含む、実施態様第 19 項記載の方法。

実施態様第 21 項：該患者の発現プロファイルが、全体の発現プロファイルに最も貢献するグループへ、個々のグループが分けられる、実施態様第 20 項記載の方法。

10

実施態様第 22 項：患者の癌進行が、全体の発現プロファイルへの D E S N T プロセスへの貢献 (p_i 値) により予測される、実施態様第 20 または 21 項記載の方法。

実施態様第 23 項：全体の発現プロファイルへの D E S N T プロセスへの貢献が、全体の発現プロファイルへの他のプロセスの貢献よりも大きい場合に、D E S N T 癌が予測される、実施態様第 22 項記載の方法。

実施態様第 24 項：全体の発現プロファイルへの D E S N T プロセスへの貢献、患者の腫瘍のステージ、患者のグリーソンスコアおよび / または患者の P S A スコアにより、D E S N T 癌が予測される、実施態様第 22 項記載の方法。

実施態様第 25 項：患者の癌サンプルの D E S N T プロセスに対する p_i 値が、少なくとも 0 . 1、少なくとも 0 . 2、少なくとも 0 . 3、少なくとも 0 . 4 あるいは少なくとも 0 . 5 である場合、D E S N T 癌が予測される、実施態様第 22 項記載の方法。

20

実施態様第 26 項：グループが D E S N T または非 D E S N T ステータスに分けられる、実施態様第 19 項記載の方法。

実施態様第 27 項：1 グループのみが D E S N T ステータスに分けられる、実施態様第 26 項記載の方法。

実施態様第 28 項：1 または複数の参照データセットから、患者のサンプルグループが D E S N T 癌を有するとき、患者の癌が侵襲性と分類されるか、癌進行が予測される、実施態様第 27 項記載の方法。

実施態様第 29 項：統計分析が複数回行なわれ、統計分析の少なくとも 60 % の検定で、1 または複数の参照データセットから患者のサンプルグループが D E S N T 癌を有するとき、患者の癌が侵襲性として分類され、癌進行が予測される、実施態様第 27 項記載の方法。

30

実施態様第 30 項：癌は前立腺癌である、実施態様第 1 から 29 のいずれか 1 項記載の方法。

実施態様第 31 項：以下を含む、癌を分類する方法、または癌の進行を予測する方法：

a) データセット中の個々の患者のサンプルの癌進行ステータスが知られている参照データセットを提供すること；

b) このデータセットから複数の遺伝子を選択すること、ここで複数の遺伝子が、表 2 にリストされたグループから選ばれた少なくとも 5、少なくとも 10、少なくとも 20、少なくとも 30、少なくとも 40、あるいは少なくとも 45 の遺伝子を含むか、または表 3 にリストされたグループから選ばれた少なくとも 5、少なくとも 10、少なくとも 15 あるいは少なくとも 20 の遺伝子を含む；

40

c) 任意に：

(i) コントロールとして患者のサンプル中の少なくとも 1 つの異なる遺伝子の発現ステータスをさらに決定すること、ここでコントロール遺伝子は表 2 あるいは表 3 にリストされた遺伝子ではない；

(i i) 複数の遺伝子およびコントロール遺伝子の発現の相対的なレベルを決定すること；および

d) これらの選択された遺伝子の発現ステータスを使用し、データセットにスーパーバイズされた機械学習アルゴリズムを適用し、癌の進行の予測を得ること；

50

e) 患者から得られたサンプル中の同じ複数の遺伝子の発現ステータスを決定し、患者の発現プロフィールを提供すること；

f) 任意に、参照データセットに対して患者の発現プロフィールを正規化すること；および

g) 患者の発現プロフィールへ予測を適用し、癌を分類するか、または癌の進行を予測すること。

実施態様第32項：該複数の遺伝子が、表2にリストされたグループから選ばれた少なくとも5、少なくとも10、少なくとも20、少なくとも30、少なくとも40、あるいは少なくとも45の遺伝子を含む、実施態様第31項記載の方法。

実施態様第33項：該複数の遺伝子が、表3にリストされたグループから選ばれた少なくとも5、少なくとも10、少なくとも15あるいは少なくとも20の遺伝子を含む、実施態様第31項記載の方法。

実施態様第34項：前記の発現の相対的なレベルを決定することが、患者のデータセットと参照データセットのそれぞれの遺伝子対の発現の比率を決定することを含む、実施態様第31から33のいずれか1項記載の方法。

実施態様第35項：該機械学習アルゴリズムがランダムフォレスト分析である、実施態様第31から34のいずれか1項記載の方法。

実施態様第36項：データセット中の個々の患者のサンプルの癌進行ステータスは、癌のDESENTステータスによって知られる、実施態様第31から35のいずれか1項記載の方法。

実施態様第37項：癌のDESENTステータスはLPDを含む分析を使用して、あらかじめ決定されている、実施態様第31から36のいずれか1項記載の方法。

実施態様第38項：癌のDESENTステータスは、実施態様第1から23のいずれか1項記載の方法を使用してあらかじめ決定されている、実施態様第37項記載の方法。

実施態様第39項：癌の分類または癌進行の予測は、癌のDESENTステータスを決定することを含む、実施態様第31から36のいずれか1項記載の方法。

実施態様第40項：少なくとも1つのコントロール遺伝子は表6あるいは表7にリストされた遺伝子である、実施態様第31から39のいずれか1項記載の方法。

実施態様第41項：少なくとも2つのコントロール遺伝子の発現ステータスが決定される、実施態様第31から40のいずれか1項記載の方法。

実施態様第42項：以下を含む、癌を分類する方法、または癌の進行を予測する方法：

a) データセット中の個々の患者の癌進行ステータスが知られている参照データセット（例えばLPD分析によって決定される）を提供すること；

b) このデータセットから複数の遺伝子を選択すること；

c) これらの選択された遺伝子の発現ステータスを使用し、データセットにスーパーバイズされた機械学習アルゴリズムを適用し、癌の進行の予測を得ること；

d) 患者から得られたサンプル中の同じ複数の遺伝子の発現ステータスを決定し、患者の発現プロフィールを提供すること；

e) 任意に参照データセットへ患者の発現プロフィールを正規化すること；および

f) 患者の発現プロフィールへ予測を適用し、癌を分類するか、または癌の進行を予測すること。

実施態様第43項：データセット中の個々の患者のサンプルの癌進行ステータスが、癌のDESENTステータスによって知られている、実施態様第42項記載の方法。

実施態様第44項：癌のDESENTステータスはLPDを含む分析を使用してあらかじめ決定されている、実施態様第42または43項記載の方法。

実施態様第45項：癌のDESENTステータスは、実施態様第1から23のいずれか1項記載の方法を使用してあらかじめ決定されている、実施態様第44項記載の方法。

実施態様第46項：該スーパーバイズされた機械学習アルゴリズムがランダムフォレスト分析である、実施態様第42から45のいずれか1項記載の方法。

実施態様第47項：癌の分類または癌進行の予測は、癌のDESENTステータスを決定す

10

20

30

40

50

ることを含む、実施態様第 4 2 から 4 6 のいずれか 1 項記載の方法。

実施態様第 4 8 項：少なくとも 1 0、少なくとも 2 0、少なくとも 3 0、少なくとも 4 0 あるいは少なくとも 5 0 の遺伝子が、ステップ b) で選択される、実施態様第 4 2 から 4 7 のいずれか 1 項記載の方法。

実施態様第 4 9 項：ステップ b) で選択された遺伝子が、進行しているかまたは進行するだろう癌内でダウンレギュレートされる、実施態様第 4 2 から 4 8 のいずれか 1 項記載の方法。

実施態様第 5 0 項：ステップ b) で選択された遺伝子が、表 2 にリストされたグループから選ばれた少なくとも 5、少なくとも 1 0、少なくとも 1 5、少なくとも 2 0、少なくとも 3 0、少なくとも 4 0、あるいは 4 5 のすべての遺伝子を含む、実施態様第 4 2 から 4 9 のいずれか 1 項記載の方法。

10

実施態様第 5 1 項：以下を含む、癌を分類する方法、または癌の進行を予測する方法：

a) データセット中の個々の患者のサンプルの D E S N T ステータスが知られている 1 以上の参照データセット（例えば L P D 分析によって決定される）を提供すること；

b) このデータセットから、発現ステータスが癌が進行しているかまたは進行するだろう癌と、癌が進行していないかまたは進行しないだろう癌との間で発現ステータスが変わることが知られている複数の遺伝子を選択すること、；

c) 選択された遺伝子に L A S S O ロジスティック回帰モデル分析を適用し、癌進行が予測された選択された遺伝子の部分集合を同定すること；

d) 選択された遺伝子のこの部分集合の発現ステータスを使用し、スーパーバイズされた機械学習アルゴリズムをデータセットに適用し、D E S N T 癌についての予測を得ること；

20

e) 患者から得られたサンプル中の選択された遺伝子の部分集合の発現ステータスを決定し、患者の発現プロファイルを提供すること；

f) 任意に 1 または複数の参照データセットへ患者の発現プロファイルを正規化すること；および

g) 患者の発現プロファイルへ予測を適用し、癌を分類するか、または癌の進行を予測すること。

実施態様第 5 2 項：データセット中の個々の患者のサンプルの癌進行ステータスが、癌の D E S N T ステータスによって知られている、実施態様第 5 1 項記載の方法。

30

実施態様第 5 3 項：癌の D E S N T ステータスは L P D を含む分析を使用してあらかじめ決定されている、実施態様第 5 1 または 5 2 項記載の方法。

実施態様第 5 4 項：癌の D E S N T ステータスは、実施態様第 1 から 2 3 のいずれか 1 項記載の方法を使用してあらかじめ決定されている、実施態様第 5 3 項記載の方法。

実施態様第 5 5 項：複数の遺伝子が表 4 にリストされた少なくとも 1 0 0、少なくとも 2 0 0、少なくとも 3 0 0、少なくとも 4 0 0、少なくとも 5 0 0 あるいは少なくとも 1 0 0 0 の遺伝子である、実施態様第 5 1 から 5 4 のいずれか 1 項記載の方法。

実施態様第 5 6 項：癌の D E S N T ステータスによって、遺伝子が癌進行を予測する、実施態様第 5 1 から 5 5 のいずれか 1 項記載の方法。

実施態様第 5 7 項：該スーパーバイズされた機械学習アルゴリズムがランダムフォレスト分析である、実施態様第 5 1 から 5 6 のいずれか 1 項記載の方法。

40

実施態様第 5 8 項：癌の分類または癌進行の予測は、癌の D E S N T ステータスを決定することを含む、実施態様第 5 1 から 5 7 のいずれか 1 項記載の方法。

実施態様第 5 9 項：癌は前立腺癌である、実施態様第 1 から 5 8 のいずれか 1 項記載の方法。

実施態様第 6 0 項：サンプルは、尿サンプル、精液サンプル、前立腺の滲出液サンプル、あるいは前立腺に由来する高分子または細胞を含んでいる任意のサンプル、全血サンプル、漿液サンプル、唾液、あるいは生検試料である、実施態様第 1 から 5 9 のいずれか 1 項記載の方法。

実施態様第 6 1 項：サンプルは前立腺生検、前立腺摘除あるいは T U R P サンプルである

50

、実施態様第 6 0 項記載の方法。

実施態様第 6 2 項：患者からサンプルを得ることをさらに含む、実施態様第 1 から 6 1 のいずれか 1 項記載の方法。

実施態様第 6 3 項：少なくとも 2、少なくとも 3、または少なくとも 5 のサンプルに対して方法が実行される、実施態様第 1 から 6 2 のいずれか 1 項記載の方法。

実施態様第 6 4 項：方法は多数の患者サンプルに対して同時に行われる、実施態様第 6 3 項記載の方法。

実施態様第 6 5 項：方法は多数の患者サンプルに対して同時に行われる、実施態様第 6 3 項記載の方法。

実施態様第 6 6 項：1 または複数のデータセットが複数の腫瘍、あるいは患者の発現プロフィールを含む、実施態様第 1 から 6 5 のいずれか 1 項記載の方法。

10

実施態様第 6 7 項：データセットは各々少なくとも 2 0、少なくとも 5 0、少なくとも 1 0 0、少なくとも 2 0 0、少なくとも 3 0 0、少なくとも 4 0 0 あるいは少なくとも 5 0 0 の患者もしくは腫瘍発現プロフィールを含む、実施態様第 6 6 項記載の方法。

実施態様第 6 8 項：患者または腫瘍の発現プロフィールは、少なくとも 1 0、少なくとも 4 0、少なくとも 1 0 0、少なくとも 5 0 0、少なくとも 1 0 0 0、少なくとも 1 5 0 0、少なくとも 2 0 0 0、少なくとも 5 0 0 0 あるいは少なくとも 1 0 0 0 0 の遺伝子の発現ステータスについての情報を含む、実施態様第 6 6 または 6 7 項記載の方法。

実施態様第 6 9 項：患者または腫瘍の発現プロフィールは、少なくとも 1 0、少なくとも 4 0、少なくとも 1 0 0、少なくとも 5 0 0、少なくとも 1 0 0 0、少なくとも 1 5 0 0、少なくとも 2 0 0 0、少なくとも 5 0 0 0 あるいは少なくとも 1 0 0 0 0 の遺伝子の発現レベルについての情報を含む、実施態様第 6 6 または 6 7 項記載の方法。

20

実施態様第 7 0 項：実施態様第 1 から 6 9 のいずれか 1 項記載の方法によって診断がされた患者へ、治療薬を投与することを含む癌を治療する方法。

実施態様第 7 1 項：a . 患者のサンプルを提供すること；

b . 実施態様第 1 から 6 9 のいずれか 1 項で定義される方法による癌進行の予測をすること；および

c . ステップ b の予測の結果により、癌進行が予測されるか、検知されるか、疑われる場合に、患者へ癌の治療薬を投与することを含む、実施態様第 7 0 項記載の方法。

実施態様第 7 2 項：実施態様第 1 から 6 9 のいずれか 1 項で定義される方法による癌進行の予測をすることを含む、癌の診断方法。

30

実施態様第 7 3 項：実施態様第 1 から 7 2 のいずれか 1 項記載の方法を実行するように構成されたコンピュータ装置。

実施態様第 7 4 項：実施態様第 1 から 7 2 のいずれか 1 項記載の方法を実行するようにプログラムされたコンピュータ読み取り可能なメディア。

実施態様第 7 5 項：表 2 にリストされた遺伝子を少なくとも 5、少なくとも 1 0、少なくとも 2 0、少なくとも 3 0、少なくとも 4 0 あるいは 4 5 のすべてを含むバイオマーカー・パネル。

実施態様第 7 6 項：表 2 にリストされた遺伝子を少なくとも 4 0 含む、実施態様第 7 5 項記載のバイオマーカー・パネル。

40

実施態様第 7 7 項：表 2 にリストされた 4 5 遺伝子のすべてを含む、実施態様第 7 5 項記載のバイオマーカー・パネル。

実施態様第 7 8 項：表 3 にリストされた遺伝子を少なくとも 5、少なくとも 1 0、少なくとも 1 5、あるいは 2 0 のすべてを含むバイオマーカー・パネル。

実施態様第 7 9 項：表 3 にリストされた遺伝子を少なくとも 1 5 含む、実施態様第 7 8 項記載のバイオマーカー・パネル。

実施態様第 8 0 項：表 3 にリストされた 2 0 遺伝子のすべてを含む、実施態様第 7 8 項記載のバイオマーカー・パネル。

実施態様第 8 1 項：表 1 にリストされた遺伝子を 少なくとも 1 0、少なくとも 1 5、少なくとも 2 0、少なくとも 5 0、少なくとも 1 0 0、少なくとも 2 0 0、少なくとも 3 0

50

0、少なくとも400あるいは500のすべてを含むバイオマーカー・パネル。

実施態様第82項：表1にリストされた遺伝子を少なくとも400含む、実施態様第81項記載のバイオマーカー・パネル。

実施態様第83項：表1にリストされた500遺伝子のすべてを含む、実施態様第81項記載のバイオマーカー・パネル。

実施態様第84項：表5にリストされた遺伝子を少なくとも5、少なくとも10、少なくとも20、少なくとも30、あるいは35のすべてを含むバイオマーカー・パネル。

実施態様第85項：表5にリストされた遺伝子を少なくとも30含む、実施態様第84項記載のバイオマーカー・パネル。

実施態様第86項：表5にリストされた35遺伝子のすべてを含む、実施態様第84項記載のバイオマーカー・パネル。

10

実施態様第87項：実施態様第1から69のいずれか1項記載の方法によって調製されたバイオマーカー・パネル。

実施態様第88項：癌の診断での使用、あるいは癌進行の予測での使用のための、実施態様第75から87のいずれか1項記載のバイオマーカー・パネル。

実施態様第89項：癌の診断または予後を予測する方法、癌進行の予測方法、または癌の分類方法における、実施態様第75から87のいずれか1項記載のバイオマーカー・パネル。

実施態様第90項：実施態様第75から87のいずれか1項記載のバイオマーカー・パネルにおいて、遺伝子の発現レベルまたは発現ステータスを決定することを含む、癌の診断または予後を予測する方法、癌進行の予測方法、または癌の分類方法。

20

実施態様第91項：測定されたバイオマーカーの発現レベルまたは発現ステータスを1つ以上の参照遺伝子と比較することをさらに含む、実施態様第90項記載の方法。

実施態様第92項：1つ以上の参照遺伝子がハウスキーピング遺伝子である、実施態様第91項記載の方法。

実施態様第93項：ハウスキーピング遺伝子が表6あるいは表7の中の遺伝子から選ばれる、実施態様第92項記載の方法。

実施態様第94項：測定されたバイオマーカーの発現レベルまたは発現ステータスを比較することを含む、実施態様第93項記載の方法。

実施態様第95項：バイオマーカー・パネルは、表2にリストされた遺伝子の少なくとも5、少なくとも10、少なくとも20、少なくとも30、少なくとも40あるいは45のすべてを含むか、または表5にリストされた遺伝子の少なくとも5、少なくとも10、少なくとも20、少なくとも30あるいは35のすべてを含み、

30

参照との発現レベルまたは発現ステータスの比較は、健康な人または、進行しているかもしくは進行するだろう癌を持っていない患者からのサンプル内の1または複数の同じ遺伝子の発現レベルまたは発現ステータスを比較することを含む、実施態様第94項記載の方法。

実施態様第96項：遺伝子の発現レベルまたは発現ステータスの比較は、前記の患者からのものを使用する、実施態様第90から95のいずれか1項記載の方法。

実施態様第97項：実施態様第1から69のいずれか1項で定義された方法により統計分析を行う、実施態様第90項記載の方法。

40

実施態様第98項：実施態様第75から87のいずれか1項で定義されたバイオマーカー・パネルからの少なくとも5つの遺伝子の発現レベルまたは発現ステータスを検知するための手段を含むキット。

実施態様第99項：実施態様第75から87のいずれか1項で定義されたバイオマーカー・パネルのすべての遺伝子の発現レベルまたは発現ステータスを検知するための手段を含む、実施態様第98項記載のキット。

実施態様第100項：1以上のコントロールあるいは参照遺伝子の発現レベルまたは発現ステータスを検知するための手段をさらに含む、実施態様第98または99項記載のキット。

50

実施態様第 101 項：使用のための指示をさらに含む、実施態様第 98 から 100 のいずれか 1 項記載のキット。

実施態様第 102 項：実施態様第 74 で定義されたコンピュータ読み取り可能なメディアをさらに含む、実施態様第 98 から 101 のいずれか 1 項記載のキット。

実施態様第 103 項：実施態様第 75 から 87 のいずれか 1 項で定義されたバイオマーカー・パネルからの少なくとも 5 つの遺伝子の発現レベルまたは発現ステータスを検知するように構成されたバイオセンサー。

実施態様第 104 項：実施態様第 75 から 87 のいずれか 1 項で定義されたバイオマーカー・パネルのすべての遺伝子の発現レベルまたは発現ステータスを検知するように構成された、実施態様第 103 項記載のバイオセンサー。

実施態様第 105 項：1 以上のコントロールあるいは参照遺伝子の発現レベルまたは発現ステータスを検知するように構成される、実施態様第 103 または 104 項記載のバイオセンサー。

実施態様第 106 項：バイオセンサーはマイクロアレイである、実施態様第 103 から 105 のいずれか 1 項記載のバイオセンサー。

実施態様第 107 項：実施態様第 103 から 106 のいずれか 1 項で定義されたバイオセンサーを含む、実施態様第 98 から 102 のいずれか 1 項記載のキット。

【0295】

参考文献

1. D'Amico, A. V., Moul, J., Carroll, P. R. & Sun, L. Cancer-specific mortality after surgery or radiation for patients with clinically localized prostate cancer managed during the prostate-specific antigen era. *J. Clin. Oncol.* 21, 2163-2172 (2003).
2. Buyyounouski, M. K., Pickles, T. & Kestin, L. L. Validating the interval to biochemical failure for the identification of potentially lethal prostate cancer. *J. Clin. Oncol.* 30, 1857-1863 (2012).
3. Draisma, G. et al. Lead time and overdiagnosis in prostate-specific antigen screening: importance of methods and context. *J. Natl. Cancer Inst.* 101, 374-383 (2009).
4. Etzioni, R., Gulati, R., Mallinger, L. & Mandelblatt, J. Influence of study features and methods on overdiagnosis estimates in breast and prostate cancer screening. *Ann. Intern. Med.* 158, 831-838 (2013).
5. Sorlie, T. et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc. Natl. Acad. Sci. U.S.A.* 100, 8418-8423 (2003).
6. Taylor, B. S. et al. Integrative genomic profiling of human prostate cancer. *Cancer Cell* 18, 11-22 (2010).
7. Ross-Adams, H. et al. Integration of copy number and transcriptomics provides risk stratification in prostate cancer: A discovery and validation cohort study. *EBioMedicine* 2, 1133-1144 (2015).
8. Carrivick, L. et al. Identification of prognostic signatures in breast cancer microarray data using Bayesian techniques. *J R Soc Interface* 3, 367-381 (2006).
9. Glinsky, G. V., Glinskii, A. B., Stephenson, A. J., Hoffman, R. M. & Gerald, W. L. Gene expression profiling predicts clinical outcome of prostate cancer. *J. Clin. Invest.* 113, 913-923 (2004).
10. Erho, N. et al. Discovery and validation of a prostate cancer genomic classifier that predicts early metastasis following radical prostatectomy. *PLoS ONE* 8, e66855 (2013).
11. Cuzick, J. et al. Prognostic value of an RNA expression signature derived from cell cycle proliferation genes in patients with prostate cancer: a retrospect

10

20

30

40

50

- ective study. *Lancet Oncol.* 12, 245-255 (2011).
12. Ramos-Montoya, A. et al. HES6 drives a critical AR transcriptional programme to induce castration-resistant prostate cancer through activation of an E2F1-mediated cell cycle network. *EMBO Molecular Medicine* n/a-n/a (2014). doi:10.1002/emmm.201303581
 13. Tomlins, S. A. et al. Characterization of 1577 primary prostate cancers reveals novel biological and clinicopathologic insights into molecular subtypes. *Eur. Urol.* 68, 555-567 (2015).
 14. You, S. et al. Integrated classification of prostate cancer reveals a novel luminal subtype with poor outcome. *Cancer Res.* (2016). doi:10.1158/0008-5472.CAN-16-0902 10
 15. Blei, D. M., Ng, A. Y. & Jordan, M. I. Latent dirichlet allocation. *the Journal of machine Learning research* (2003).
 16. Boutros, P. C. et al. Spatial genomic heterogeneity within localized, multifocal prostate cancer. *Nat. Genet.* (2015). doi:10.1038/ng.3315
 17. Clark, J. et al. Complex patterns of ETS gene alteration arise during cancer development in the human prostate. *Oncogene* 27, 1993-2003 (2008).
 18. Cooper, C. S. et al. Analysis of the genetic phylogeny of multifocal prostate cancer identifies multiple independent clonal expansions in neoplastic and morphologically normal prostate tissue. *Nat. Genet.* 47, 367-372 (2015). 20
 19. Svensson, M. A. et al. Testing mutual exclusivity of ETS rearranged prostate cancer. *Lab. Invest.* 91, 404-412 (2011).
 20. Cancer Genome Atlas Research Network. The Molecular Taxonomy of Primary Prostate Cancer. *Cell* 163, 1011-1025 (2015).
 21. Olmos, D. et al. Prognostic value of blood mRNA expression signatures in castration-resistant prostate cancer: a prospective, two-stage study. *Lancet Oncol.* 13, 1114-1124 (2012).
 22. Klein, E. A., Yousefi, K., Haddad, Z., Choeurng, V. & Buerki, C. A genomic classifier improves prediction of metastatic disease within 5 years after surgery in node-negative high-risk prostate cancer patients managed by *Eur. Urol.* (2015). 30
 23. Stephenson, A. J. et al. Integration of gene expression profiling and clinical variables to predict prostate carcinoma recurrence after radical prostatectomy. *Cancer* 104, 290-298 (2005).
 24. Ramaswamy, S., Ross, K. N., Lander, E. S. & Golub, T. R. A molecular signature of metastasis in primary solid tumors. *Nat. Genet.* 33, 49-54 (2003).
 25. Klein, E. A. et al. A 17-gene assay to predict prostate cancer aggressiveness in the context of Gleason grade heterogeneity, tumor multifocality, and biopsy undersampling. *Eur. Urol.* 66, 550-560 (2014).
 26. Tomlins, S. A. et al. Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. *Science* 310, 644-648 (2005). 40
 27. Weischenfeldt, J. et al. Integrative genomic analyses reveal an androgen-driven somatic alteration landscape in early-onset prostate cancer. *Cancer Cell* 23, 159-170 (2013).
 28. Park, K. et al. TMPRSS2:ERG Gene Fusion Predicts Subsequent Detection of Prostate Cancer in Patients With High-Grade Prostatic Intraepithelial Neoplasia. *J. Clin. Oncol.* 32, 206-211 (2014).
 29. Friedl, P., Locker, J., Sahai, E. & Segall, J. E. Classifying collective cancer cell invasion. *Nat. Cell Biol.* 14, 777-783 (2012).
 30. Schroder, F. H. et al. Screening and prostate cancer mortality: results of 50

the European Randomised Study of Screening for Prostate Cancer (ERSPC) at 13 years of follow-up. *Lancet* 384, 2027-2035 (2014).

31. Warren, A. Y. et al. Method for sampling tissue for research which preserves pathological data in radical prostatectomy. *Prostate* 73, 194-202 (2013).

32. Jhavar, S. et al. Detection of TMPRSS2-ERG translocations in human prostate cancer by expression profiling using GeneChip Human Exon 1.0 ST arrays. *J Mol Diagn* 10, 50-57 (2008).

33. Irizarry, R. A. et al. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 4, 249-264 (2003).

34. Johnson, W. E., Li, C. & Rabinovic, A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* (2007).

35. Rogers, S., Girolami, M., Campbell, C. & Breitling, R. The latent process decomposition of cDNA microarray data sets. *IEEE/ACM Trans Comput Biol Bioinform* 2, 143-156 (2005).

36. Ritchie, M. E. et al. limma powers differential expression analyses for RNA-seq and microarray studies. *Nucleic Acids Res.* 43, e47 (2015).

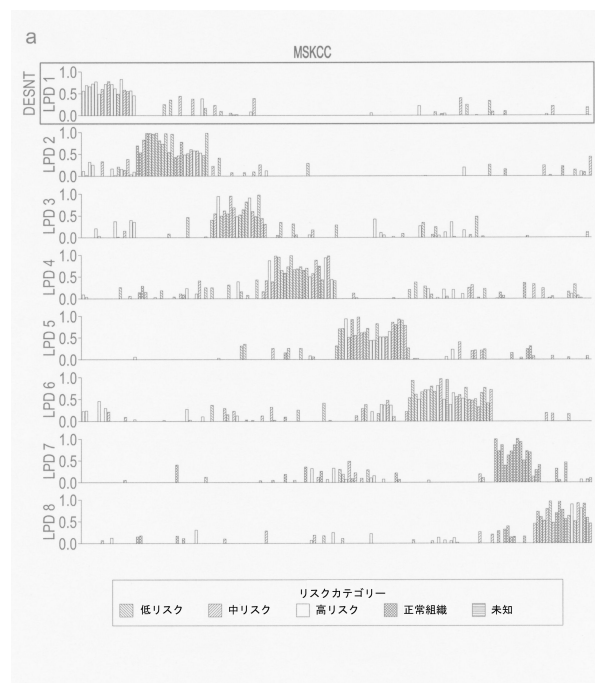
37. Friedman, J., Hastie, T. & Tibshirani, R. Regularization paths for generalized linear models via coordinate descent. *Journal of statistical software* (2010).

38. Breiman, L. Random forests. *Machine learning* (2001).

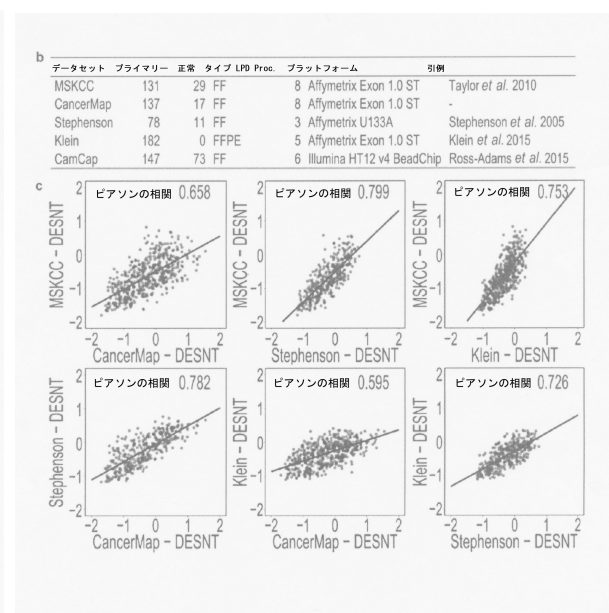
39. Liaw, A. & Wiener, M. Classification and regression by randomForest. *R news* (2002).

【図面】

【図 1 - 1】



【図 1 - 2】



10

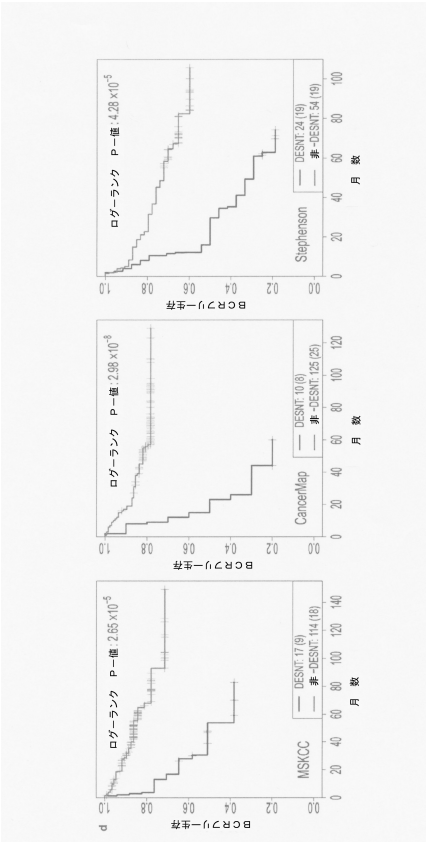
20

30

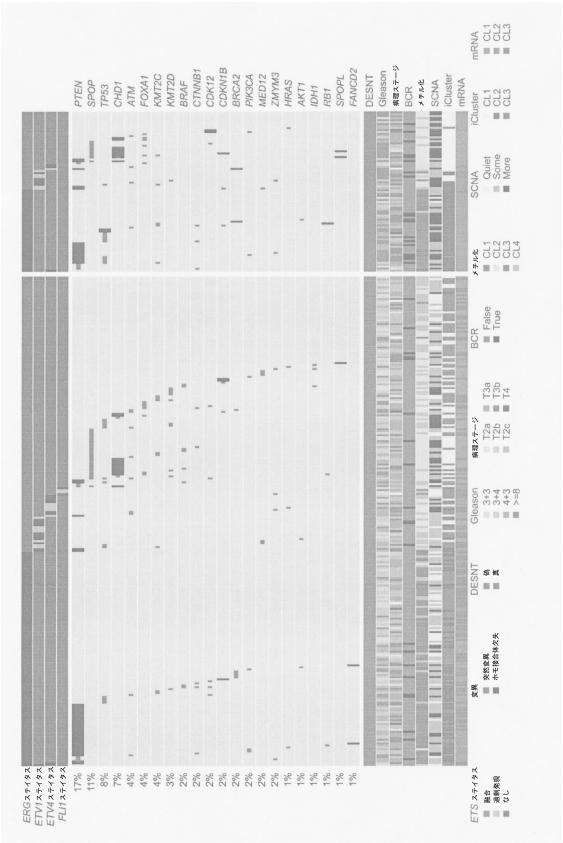
40

50

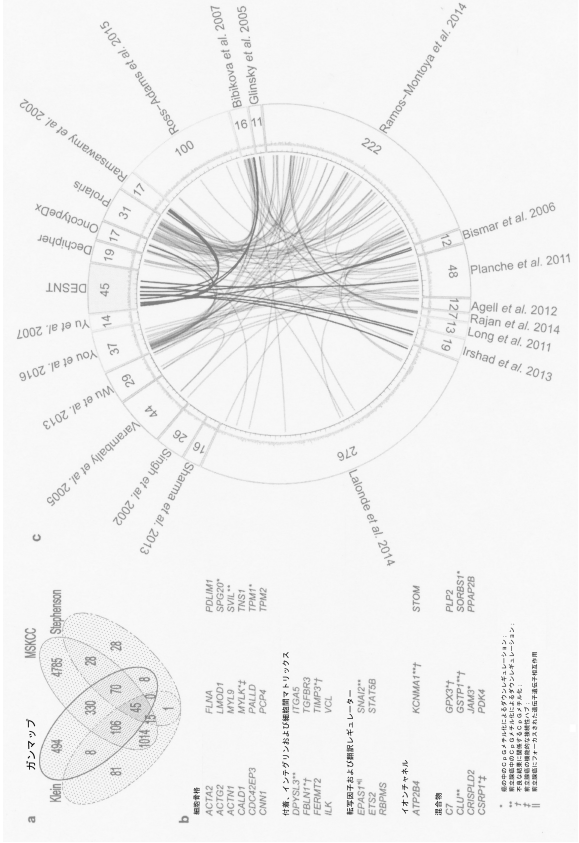
【図 1 - 3】



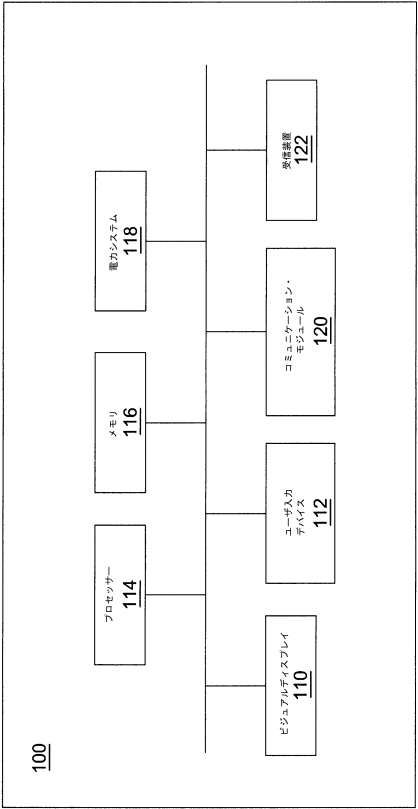
【図 3】



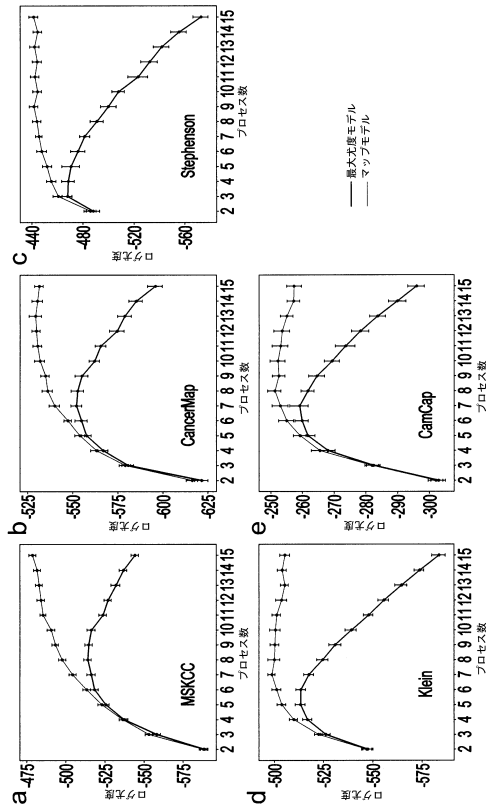
【図 2】



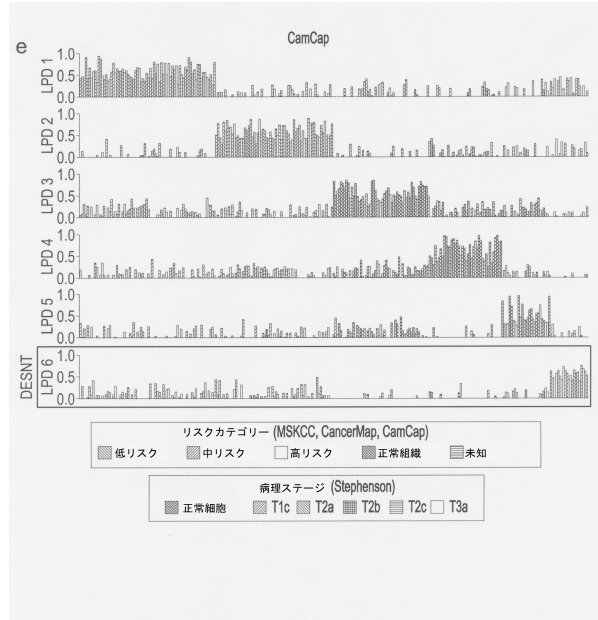
【図 4】



【図 5】



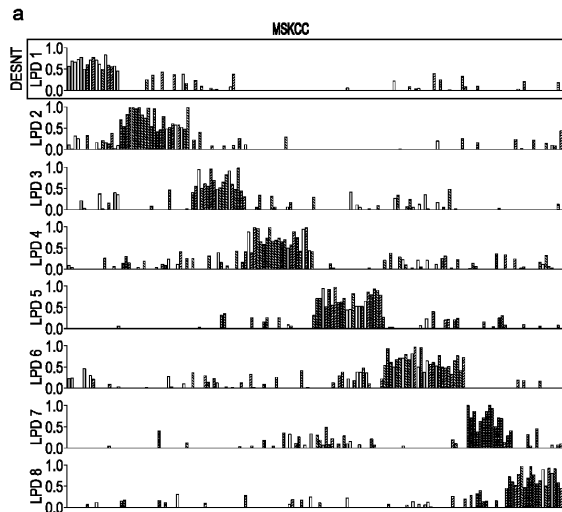
【図 6】



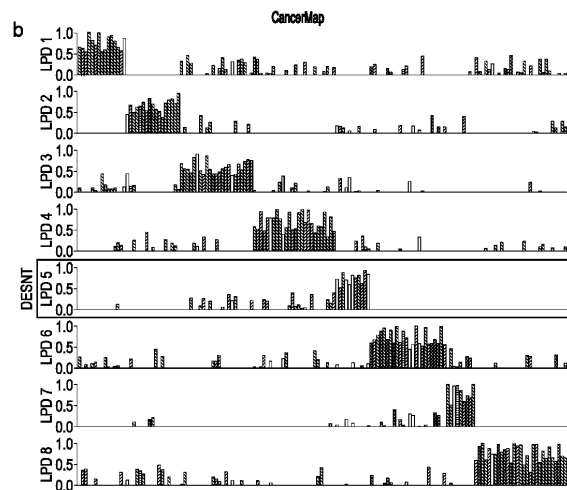
10

20

【図 6 a】



【図 6 b】

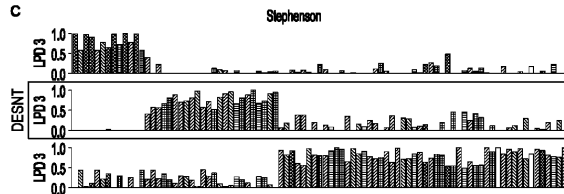


30

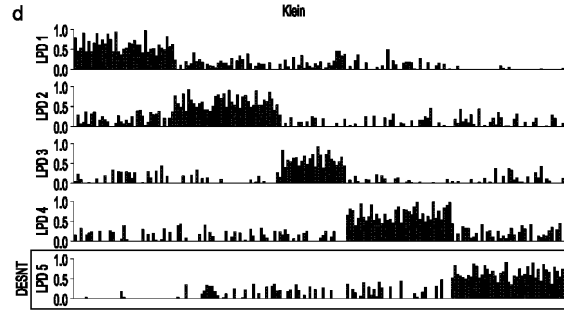
40

50

【図 6 c】

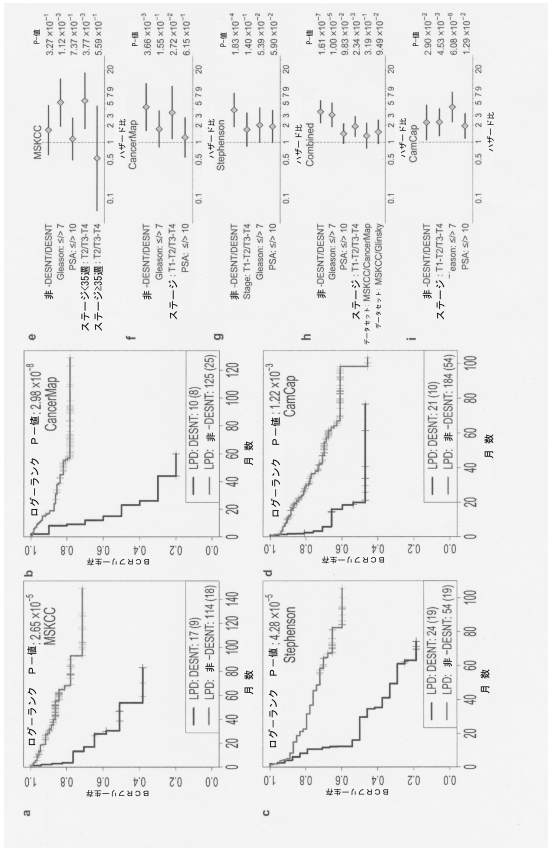


【図 6 d】

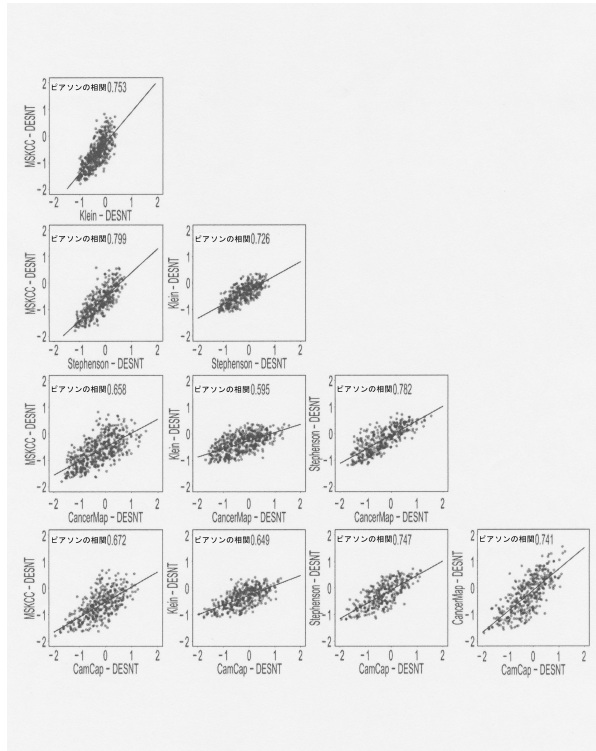


10

【図 7】



【図 8】



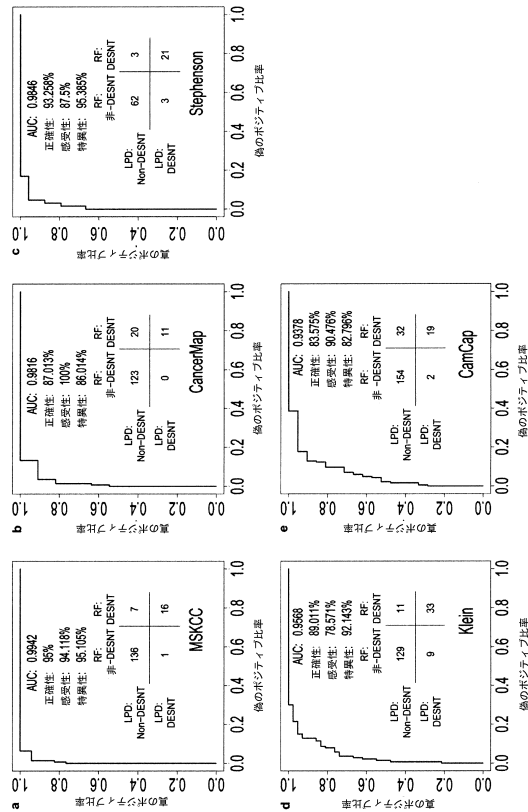
20

30

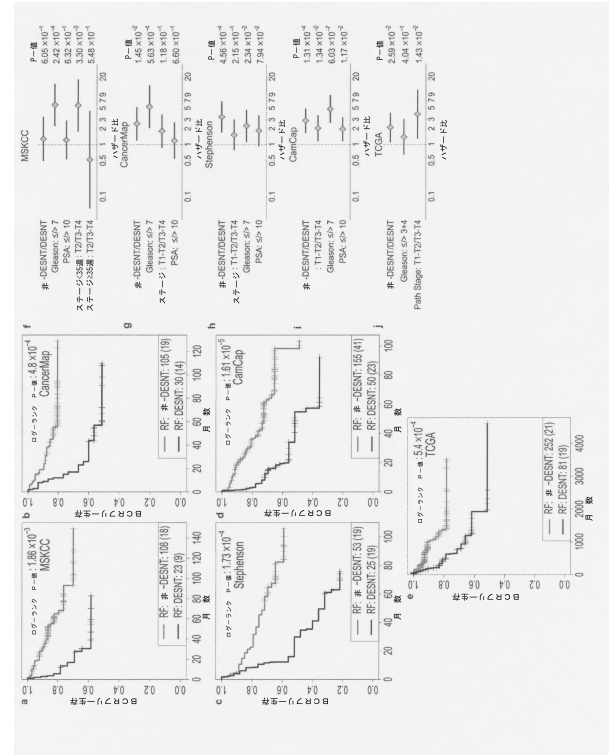
40

50

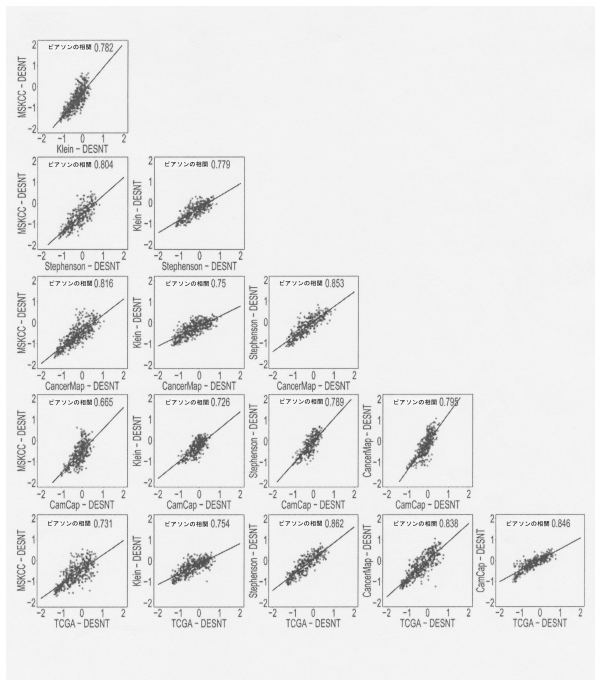
【図 9】



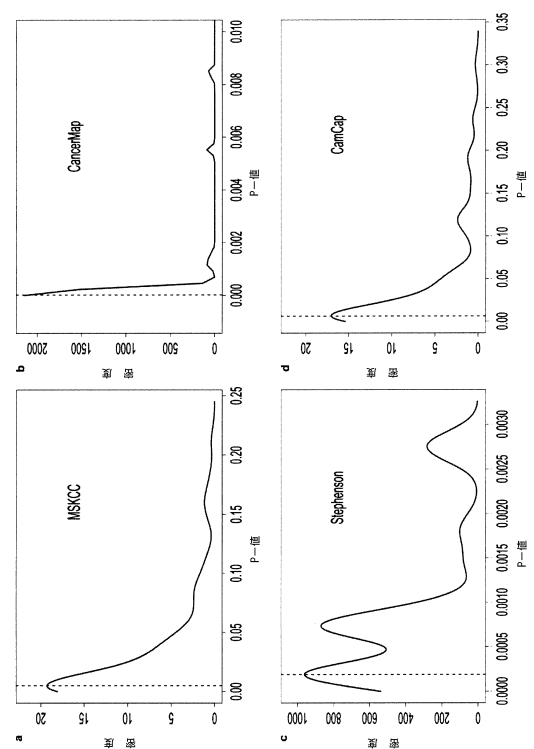
【図 10】



【図 11】



【図 12】



10

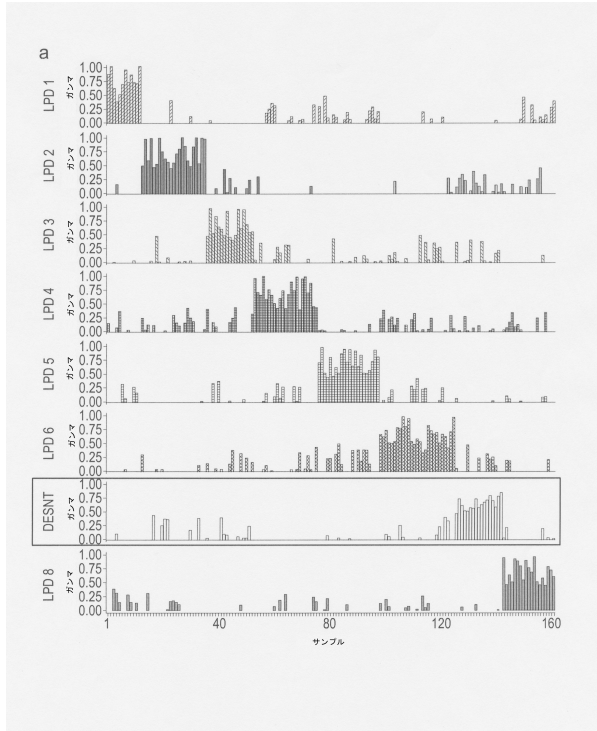
20

30

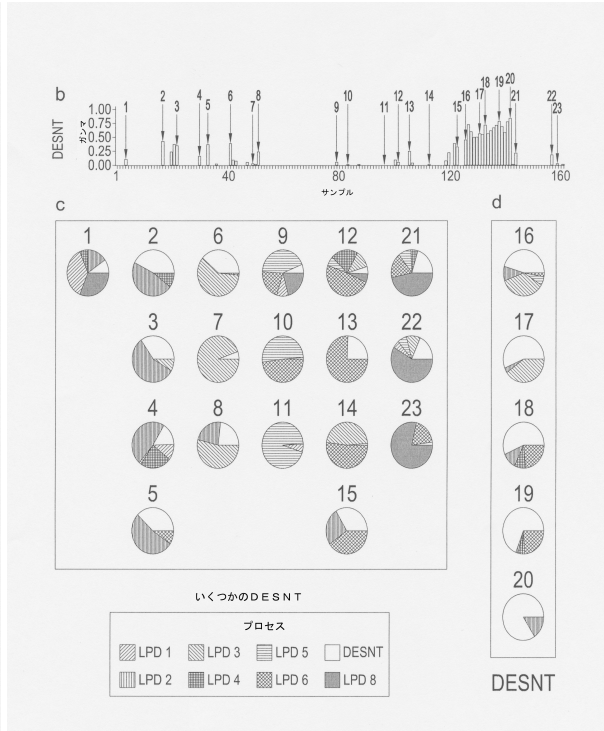
40

50

【図 13 - 1】



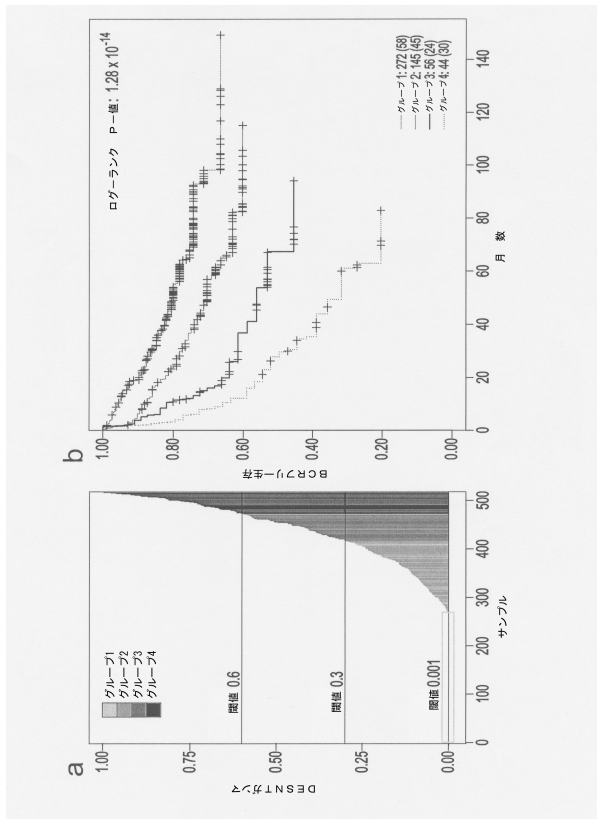
【図 13 - 2】



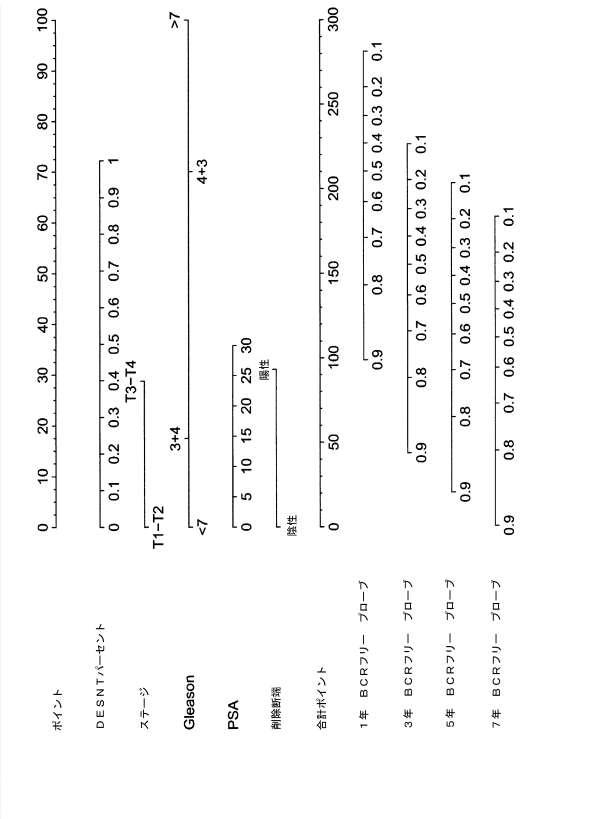
10

20

【図 14】



【図 15】

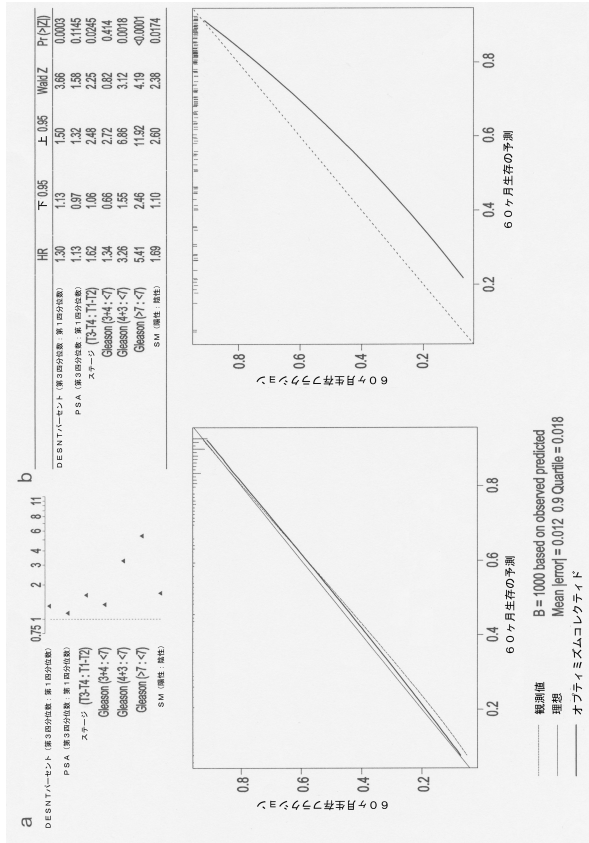


30

40

50

【図 16】



10

20

30

40

50

フロントページの続き

- 英国(GB)
ルウィッチ ノルフォーク ユニバーシティ オブ イースト アングリア シ - / オー
(72)発明者 ブルーワー ダニエル サイモン
英国 エヌアール4 7ティージェイ ノルウィッチ ノルフォーク ユニバーシティ オブ イースト
アングリア シ - / オー
(72)発明者 クーパー コリン ステファン
英国 エヌアール4 7ティージェイ ノルウィッチ ノルフォーク ユニバーシティ オブ イースト
アングリア シ - / オー
合議体
審判長 伏本 正典
審判官 佐藤 智康
審判官 松田 直也
(56)参考文献 特表2016-521966(JP,A)
David Olmos et al., Prognostic value of blood mRNA expression signatures in castration-resistant prostate cancer: a prospective, two-stage study, online, PMC, 2016年5月24日, PMID: PMC4878433 DOI: 10.1016/S1470-2045(12)70372-8, [2022年4月19日検索], インターネット, <URL: <https://pubmed.ncbi.nlm.nih.gov/23059046/>>
(58)調査した分野 (Int.Cl., DB名)
G16B 5/00-99/00