

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2011-527565

(P2011-527565A)

(43) 公表日 平成23年11月4日(2011.11.4)

(51) Int.Cl.	F I	テーマコード (参考)
C 1 2 Q 1/68 (2006.01)	C 1 2 Q 1/68 Z N A A	4 B O 2 4
C 1 2 N 15/09 (2006.01)	C 1 2 N 15/00 A	4 B O 6 3
G O 1 N 33/53 (2006.01)	G O 1 N 33/53 M	

審査請求 未請求 予備審査請求 未請求 (全 114 頁)

(21) 出願番号	特願2011-517318 (P2011-517318)	(71) 出願人	509268233 ディコーダ ジェネテクス イーエイチエフ
(86) (22) 出願日	平成21年7月3日 (2009.7.3)		
(85) 翻訳文提出日	平成23年2月28日 (2011.2.28)		
(86) 国際出願番号	PCT/IS2009/000008		アイスランド国 レイキャピク スタール
(87) 国際公開番号	W02010/004591		ガータ 8
(87) 国際公開日	平成22年1月14日 (2010.1.14)	(74) 代理人	100106002 弁理士 正林 真之
(31) 優先権主張番号	8746	(74) 代理人	100120891 弁理士 林 一好
(32) 優先日	平成20年7月7日 (2008.7.7)	(72) 発明者	ステイシー サイモン アイスランド国 コーパヴォグル フェナ
(33) 優先権主張国	アイスランド (IS)	(72) 発明者	スレム パトリック アイスランド国 レイキャピク エスキー
			リッド 22

最終頁に続く

(54) 【発明の名称】 乳癌のリスクアセスメントのための遺伝的変異

(57) 【要約】

本発明は、乳癌の感受性変異体であると判定されたある遺伝的変異体に関する。乳癌に対する増加した感受性を診断することを含む疾病管理の方法、治療に対する反応の予測方法、およびこのような変異体を使用した予後診断の予測方法が記載されている。本発明はさらに、本発明の方法において有用なキットに関する。

【選択図】なし

【特許請求の範囲】

【請求項 1】

ヒト個体における乳癌に対する感受性の判定方法であって、少なくとも1つの多型マーカーの少なくとも1つの対立遺伝子が前記個体から入手した核酸試料の中にまたは前記個体由来の遺伝子型のデータセットの中に存在するかどうかを判定することを含み、前記少なくとも1つの多型マーカーは、rs 999737 (配列番号：6)、rs 2005154 (配列番号：1)、rs 2184380 (配列番号：2)、rs 2224696 (配列番号：3)、rs 2242503 (配列番号：4)、rs 12291026 (配列番号：5)、rs 9956546 (配列番号：7)、rs 11912922 (配列番号：8)、rs 6001954 (配列番号：9) およびそれらと連鎖不平衡のマーカーからなる群から選択され、前記少なくとも1つの対立遺伝子の存在の判定は、前記個体についての乳癌に対する感受性を示す、方法。

10

【請求項 2】

前記少なくとも1つの多型マーカーは、表4に記載されるマーカーから選択される、請求項1に記載の方法。

【請求項 3】

前記少なくとも1つの多型マーカーは、rs 999737 (配列番号：6)、rs 2005154 (配列番号：1)、rs 2184380 (配列番号：2)、rs 2224696 (配列番号：3)、rs 2242503 (配列番号：4)、rs 12291026 (配列番号：5)、rs 9956546 (配列番号：7)、rs 11912922 (配列番号：8)、rs 6001954 (配列番号：9) から選択される、請求項1または請求項2に記載の方法。

20

【請求項 4】

前記個体における少なくとも1つのハプロタイプの頻度を評価することをさらに含む、請求項1から請求項3のいずれか1項に記載の方法。

【請求項 5】

前記少なくとも1つの対立遺伝子またはハプロタイプの存在により与えられる感受性は増加した感受性である、請求項1から請求項4のいずれか1項に記載の方法。

【請求項 6】

rs 2005154における対立遺伝子T、rs 2184380における対立遺伝子G、rs 2224696における対立遺伝子T、rs 2242503における対立遺伝子C、rs 12291026における対立遺伝子G、rs 999737における対立遺伝子C、rs 9956546における対立遺伝子A、rs 11912922における対立遺伝子T、rs 6001954における対立遺伝子Gの存在は、前記個体における乳癌に対する増加した感受性を示す、請求項6に記載の方法。

30

【請求項 7】

前記少なくとも1つの対立遺伝子またはハプロタイプの存在は、少なくとも1.08の相対リスク(RR)またはオッズ比(OR)で乳癌に対する増加した感受性を示す、請求項5または請求項6に記載の方法。

【請求項 8】

前記少なくとも1つの対立遺伝子またはハプロタイプの存在は、少なくとも1.09 (少なくとも1.10、少なくとも1.11、少なくとも1.12、少なくとも1.13、少なくとも1.14、少なくとも1.15、少なくとも1.16、および少なくとも1.17を含む)の相対リスク(RR)またはオッズ比(OR)で、増加した感受性を示す、請求項5または請求項6に記載の方法。

40

【請求項 9】

前記少なくとも1つの対立遺伝子またはハプロタイプの存在によって与えられる感受性は減少した感受性である、請求項1から請求項4のいずれか1項に記載の方法。

【請求項 10】

表4に記載されるマーカーのうちのいずれか1つと連鎖不平衡ではない乳癌についての

50

少なくとも1つのリスクのある変異体の少なくとも1つのリスクのある対立遺伝子が、ヒト個体から得たゲノムDNAを含む試料またはヒト個体に由来する遺伝子型のデータセットの中に存在するかどうかを判定することをさらに含む、請求項1から請求項9のいずれか1項に記載の方法。

【請求項11】

少なくとも2つの多型マーカーのそれぞれの少なくとも1つの対立遺伝子が、ヒト個体から得たゲノムDNAを含む試料、またはヒト個体由来の遺伝子型のデータセットの中に存在するかどうかについての判定を含み、前記少なくとも2つの多型マーカーの前記少なくとも1つの対立遺伝子の存在は、乳癌に対する増加した感受性を示す、請求項1から請求項9のいずれか1項に記載の方法。

10

【請求項12】

前記個体から入手した核酸試料の、または前記個体由来の遺伝子型のデータセットの中の乳癌についての少なくとも1つの高い浸透性の遺伝要因の存在または不存在を判定する工程をさらに含む、請求項1から請求項11のいずれか1項に記載の方法。

【請求項13】

前記高い浸透性の遺伝要因は、BRCA1、BRCA2、TP53および/またはPTENの中の変異である、請求項12に記載の方法。

【請求項14】

前記個体のリスクアセスメント、診断、または予後診断を行うために、非遺伝的情報を分析することをさらに含む、請求項1から請求項13のいずれか1項に記載の方法。

20

【請求項15】

前記非遺伝的情報は、年齢、性別、民族性、社会経済的状況、以前の疾病診断、患者の病歴、乳癌の家族歴、生化学的測定結果、および臨床的測定結果から選択される、請求項14に記載の方法。

【請求項16】

合わせたリスクを算出することをさらに含む、請求項11から請求項15のいずれか1項に記載の方法。

【請求項17】

以前に乳癌と診断された個体の少なくとも続発性原発腫瘍の発生リスクの判定方法であって、前記方法は、少なくとも1つの多型マーカーの少なくとも1つの対立遺伝子が前記個体から入手した核酸試料または前記個体由来の遺伝子型のデータセットの中に存在するかどうかについての判定を含み、前記少なくとも1つの多型マーカーは、rs999737(配列番号:6)、rs2005154(配列番号:1)、rs2184380(配列番号:2)、rs2224696(配列番号:3)、rs2242503(配列番号:4)、rs12291026(配列番号:5)、rs9956546(配列番号:7)、rs11912922(配列番号:8)、rs6001954(配列番号:9)、およびそれらと連鎖不平衡のマーカーから選択され、前記少なくとも1つの対立遺伝子の存在は、少なくとも続発性原発腫瘍の発生のリスクを示す、方法。

30

【請求項18】

前記少なくとも1つの多型マーカーは、表4に記載されるマーカーから選択される、請求項17に記載の方法。

40

【請求項19】

ヒト個体において乳癌に対する感受性を評価するためのキットであって、前記キットは、

前記個体のゲノムにおいて少なくとも1つの多型マーカーの少なくとも1つの対立遺伝子を選択的に検出するための試薬であって、前記多型マーカーは、rs999737(配列番号:6)、rs2005154(配列番号:1)、rs2184380(配列番号:2)、rs2224696(配列番号:3)、rs2242503(配列番号:4)、rs12291026(配列番号:5)、rs9956546(配列番号:7)、rs11912922(配列番号:8)、rs6001954(配列番号:9)、およびそれらと

50

連鎖不平衡のマーカークからなる群から選択される、試薬、ならびに

前記少なくとも1つの多型性と乳癌に対する感受性との間の相関データを含むデータの収集物、を含むキット。

【請求項20】

前記少なくとも1つの多型マーカークは、表4に記載されるマーカークから選択される、請求項19に記載のキット。

【請求項21】

前記試薬は、前記少なくとも1つの多型マーカークを含む前記個体のゲノムの断片とハイブリダイズする少なくとも1つの近接するオリゴヌクレオチド、バッファークおよび検出可能な標識を含む、請求項19または請求項20に記載のキット。

10

【請求項22】

前記試薬は、被験者から入手したゲノム核酸分節の逆ストランドとハイブリダイズする少なくとも1つのオリゴヌクレオチド対を含み、各オリゴヌクレオチドプライマー対は、1つの多型マーカークを含む前記個体のゲノム断片を選択的に増幅するように設計され、前記断片は、大きさが少なくとも30塩基対である、請求項19から請求項21のいずれか1項に記載のキット。

【請求項23】

前記少なくとも1つのオリゴヌクレオチドは、前記個体の前記ゲノムに対して完全に相補的である、請求項21または請求項22に記載のキット。

20

【請求項24】

前記キットは、前記個体の前記ゲノムにおいて100以下の対立遺伝子を検出するための試薬を含む、請求項19から請求項22のいずれか1項に記載のキット。

【請求項25】

ヒト個体における乳癌に対する感受性を判定するためのコンピューターが実行可能な命令を有するコンピューター可読媒体であって、前記コンピューター可読媒体は、少なくとも1つの多型マーカークを示すデータ；

前記コンピューター可読媒体に保存され、かつ前記少なくとも1つの多型マーカークについて乳癌の発生のリスクを判定するためにプロセッサによって実行されるように適合されたルーチン；

30

を含み、

前記少なくとも1つの多型マーカークは、rs999737（配列番号：6）、rs2005154（配列番号：1）、rs2184380（配列番号：2）、rs2224696（配列番号：3）、rs2242503（配列番号：4）、rs12291026（配列番号：5）、rs9956546（配列番号：7）、rs11912922（配列番号：8）、rs6001954（配列番号：9）、およびそれらと連鎖不平衡のマーカークから選択される、コンピューター可読媒体。

【請求項26】

前記コンピューター可読媒体は、少なくとも2つの多型マーカークを示すデータを含む、請求項25に記載のコンピューター可読媒体。

40

【請求項27】

前記少なくとも1つの多型マーカークは、表4に記載されるマーカークから選択される、請求項25または請求項26に記載のコンピューター可読媒体。

【請求項28】

前記少なくとも1つの多型マーカークは、マーカークrs2005154（配列番号：1）、rs2184380（配列番号：2）、rs2224696（配列番号：3）、rs2242503（配列番号：4）、rs12291026（配列番号：5）、rs999737（配列番号：6）、rs9956546（配列番号：7）、rs11912922（配列番号：8）、およびrs6001954（配列番号：9）から選択される、請求項25から請求項27のいずれか1項に記載のコンピューター可読媒体。

50

【請求項 29】

2以上の多型マーカを含む少なくとも1つのハプロタイプを示すデータをさらに含む、請求項25から請求項28のいずれか1項に記載のコンピューター可読媒体。

【請求項 30】

ヒト個体において乳癌についての遺伝的指標を判定するための装置であって、プロセッサ、

コンピューター可読メモリであって、rs999737（配列番号：6）、rs2005154（配列番号：1）、rs2184380（配列番号：2）、rs2224696（配列番号：3）、rs2242503（配列番号：4）、rs12291026（配列番号：5）、rs9956546（配列番号：7）、rs11912922（配列番号：8）、rs6001954（配列番号：9）、およびそれらと連鎖不平衡のマーカから選択される少なくとも1つの多型マーカに関して少なくとも1人のヒト個体についてのマーカおよび/またはハプロタイプ情報を分析し、前記マーカまたはハプロタイプの情報に基づいて結果を生じるために前記プロセッサ上で実行されるように適合されたコンピューターが実行可能な命令を有し、前記結果は、前記ヒト個体についての乳癌の遺伝的指標として、前記少なくとも1つのマーカまたはハプロタイプのリスク判定基準を含む、コンピューター可読メモリを含む、装置。

10

【請求項 31】

前記コンピューター可読メモリは、前記少なくとも1つの多型マーカの少なくとも1つの対立遺伝子または少なくとも1つのハプロタイプに関連する乳癌の発生のリスクを示すデータをさらに含み、前記ヒト個体についてのリスク判定基準は、前記ヒト個体についての遺伝子型の状態と、前記少なくとも1つの対立遺伝子または前記少なくとも1つのハプロタイプに関連する乳癌の発生のリスクとの比較に基づく、請求項30に記載の装置。

20

【請求項 32】

前記コンピューター可読メモリは、乳癌と診断された複数の個体における前記少なくとも1つの多型マーカの少なくとも1つの対立遺伝子または前記少なくとも1つのハプロタイプの頻度を示すデータ、および複数の参照個体における少なくとも1つの多型マーカの少なくとも1つの対立遺伝子または少なくとも1つのハプロタイプの頻度を示すデータをさらに含み、前記少なくとも1つの対立遺伝子またはハプロタイプに関連する乳癌の発生のリスクは、乳癌と診断された個体および参照個体における前記少なくとも1つの対立遺伝子またはハプロタイプの頻度の比較に基づく、請求項31に記載の装置。

30

【請求項 33】

前記少なくとも1つのマーカまたはハプロタイプは表4に記載されるマーカから選択される少なくとも1つのマーカを含む、請求項30から請求項32のいずれか1項に記載の装置。

【請求項 34】

前記少なくとも1つのマーカまたはハプロタイプは、rs999737（配列番号：6）、rs2005154（配列番号：1）、rs2184380（配列番号：2）、rs2224696（配列番号：3）、rs2242503（配列番号：4）、rs12291026（配列番号：5）、rs9956546（配列番号：7）、rs11912922（配列番号：8）、rs6001954（配列番号：9）からなる群から選択される、請求項30から請求項33のいずれか1項に記載の装置。

40

【請求項 35】

前記リスク判定基準は、オッズ比（OR）または相対リスク（RR）によって特徴付けられる、請求項30から請求項34のいずれか1項に記載の装置。

【請求項 36】

乳癌に対する感受性を評価するために使用するためのマーカの同定方法であって、前記方法は、

a) rs999737（配列番号：6）、rs2005154（配列番号：1）、rs

50

2 1 8 4 3 8 0 (配列番号: 2)、rs 2 2 2 4 6 9 6 (配列番号: 3)、rs 2 2 4 2 5 0 3 (配列番号: 4)、rs 1 2 2 9 1 0 2 6 (配列番号: 5)、rs 9 9 5 6 5 4 6 (配列番号: 7)、rs 1 1 9 1 2 9 2 2 (配列番号: 8) および rs 6 0 0 1 9 5 4 (配列番号: 9) からなる群から選択される少なくとも1つのマーカーと連鎖不平衡の少なくとも1つの多型マーカーを同定すること;

b) 前記少なくとも1つの多型マーカーについて、乳癌と診断された個体または乳癌に対する感受性を有する個体の試料の遺伝子型状態を判定すること; ならびに

c) 前記少なくとも1つの多型マーカーについての対照個体の試料の遺伝子型状態を判定すること;

を含み、前記対照試料中の少なくとも1つの対立遺伝子の頻度と比較した、乳癌と診断された個体または乳癌に対する感受性を有する個体における少なくとも1つの多型性の前記少なくとも1つの対立遺伝子の頻度の有意差は、前記少なくとも1つの多型性が、乳癌に対する感受性の評価に有用であることを示す、方法。

10

【請求項 3 7】

前記対照試料の前記少なくとも1つの対立遺伝子の頻度と比較した、乳癌と診断された個体または乳癌に対する感受性を有する個体における前記少なくとも1つの多型性における、前記少なくとも1つの対立遺伝子の頻度の増加は、前記少なくとも1つの多型性が、乳癌に対する増加した感受性の評価に有用であることを示す、請求項 3 6 に記載の方法。

【請求項 3 8】

前記対照試料中の前記少なくとも1つの対立遺伝子の頻度と比較した、乳癌と診断された個体または乳癌に対する感受性を有する個体における前記少なくとも1つの多型性における前記少なくとも1つの対立遺伝子の頻度の減少は、前記少なくとも1つの多型性が、乳癌に対する減少した感受性、または乳癌に対する保護の評価に有用であることを示す、請求項 3 6 または請求項 3 7 に記載の方法。

20

【請求項 3 9】

少なくとも1つの多型マーカーの少なくとも1つの対立遺伝子が個体試料由来の核酸試料の中に存在するかどうかを判定することを含む、ヒト個体から入手した核酸試料の遺伝子型を同定する方法であって、前記少なくとも1つのマーカーは、rs 9 9 9 7 3 7 (配列番号: 6)、rs 2 0 0 5 1 5 4 (配列番号: 1)、rs 2 1 8 4 3 8 0 (配列番号: 2)、rs 2 2 2 4 6 9 6 (配列番号: 3)、rs 2 2 4 2 5 0 3 (配列番号: 4)、rs 1 2 2 9 1 0 2 6 (配列番号: 5)、rs 9 9 5 6 5 4 6 (配列番号: 7)、rs 1 1 9 1 2 9 2 2 (配列番号: 8)、rs 6 0 0 1 9 5 4 (配列番号: 9)、およびそれらと連鎖不平衡のマーカーから選択され、前記試料中の前記少なくとも1つの対立遺伝子の存在の判定は、前記個体における乳癌に対する感受性を示す、方法。

30

【請求項 4 0】

rs 2 0 0 5 1 5 4 における対立遺伝子 T、rs 2 1 8 4 3 8 0 における対立遺伝子 G、rs 2 2 2 4 6 9 6 における対立遺伝子 T、rs 2 2 4 2 5 0 3 における対立遺伝子 C、rs 1 2 2 9 1 0 2 6 における対立遺伝子 G、rs 9 9 9 7 3 7 における対立遺伝子 C、rs 9 9 5 6 5 4 6 における対立遺伝子 A、rs 1 1 9 1 2 9 2 2 における対立遺伝子 T、または rs 6 0 0 1 9 5 4 における対立遺伝子 G の存在の判定は、前記個体における乳癌の増加した感受性を示す、請求項 3 9 に記載の方法。

40

【請求項 4 1】

乳癌治療薬に対する反応可能性について個体を評価する方法であって、少なくとも1つの多型マーカーの少なくとも1つの対立遺伝子が、前記個体から入手した核酸試料または前記個体由来の遺伝子型のデータセット中に存在しているかどうかを判定する工程を含み、前記少なくとも1つの多型マーカーは、rs 9 9 9 7 3 7 (配列番号: 6)、rs 2 0 0 5 1 5 4 (配列番号: 1)、rs 2 1 8 4 3 8 0 (配列番号: 2)、rs 2 2 2 4 6 9 6 (配列番号: 3)、rs 2 2 4 2 5 0 3 (配列番号: 4)、rs 1 2 2 9 1 0 2 6 (配列番号: 5)、rs 9 9 5 6 5 4 6 (配列番号: 7)、rs 1 1 9 1 2 9 2 2 (配列番号: 8)、rs 6 0 0 1 9 5 4 (配列番号: 9)、およびそれらと連鎖不平衡のマーカーか

50

ら選択され、前記少なくとも1つのマーカーの前記少なくとも1つの対立遺伝子の存在は、前記治療薬に対する陽性反応の可能性を示す、方法。

【請求項42】

乳癌と診断された個体の予後診断を予測する方法であって、前記方法は、少なくとも1つの多型マーカーの少なくとも1つの対立遺伝子が前記個体から入手した核酸試料または前記個体由来の遺伝子型データセットの中に存在するかどうかを判定することを含み、前記少なくとも1つの多型マーカーは、rs999737（配列番号：6）、rs2005154（配列番号：1）、rs2184380（配列番号：2）、rs2224696（配列番号：3）、rs2242503（配列番号：4）、rs12291026（配列番号：5）、rs9956546（配列番号：7）、rs11912922（配列番号：8）、rs6001954（配列番号：9）、およびそれらと連鎖不平衡のマーカーから選択され、前記少なくとも1つの対立遺伝子の存在は、前記個体における前記乳癌の悪化した予後診断を示す、方法。

10

【請求項43】

乳癌の治療を受けている個体の治療の進行をモニターする方法であって、前記方法は、少なくとも1つの多型マーカーの少なくとも1つの対立遺伝子が前記個体から入手した核酸試料または前記個体由来の遺伝子型データセットの中に存在するかどうかを判定することを含み、前記少なくとも1つの多型マーカーは、rs999737（配列番号：6）、rs2005154（配列番号：1）、rs2184380（配列番号：2）、rs2224696（配列番号：3）、rs2242503（配列番号：4）、rs12291026（配列番号：5）、rs9956546（配列番号：7）、rs11912922（配列番号：8）、rs6001954（配列番号：9）、およびそれらと連鎖不平衡のマーカーから選択され、前記少なくとも1つの対立遺伝子の存在は前記個体の治療成績を示す、方法。

20

【請求項44】

前記少なくとも1つの多型マーカーは表4に記載されるマーカーから選択される、請求項41から請求項43のいずれか1項に記載の方法。

【請求項45】

ヒト個体における乳癌に対する感受性の診断および/または評価のための試薬の製造におけるオリゴヌクレオチドプローブの使用であって、前記プローブは、配列番号：1～562のいずれか1つに記載されるヌクレオチド配列を有する核酸分節とハイブリダイズし、前記プローブは長さが15～500ヌクレオチドである、使用。

30

【請求項46】

ヒト個体における乳癌に対する感受性の判定方法であって、前記方法は、rs999737（配列番号：6）、rs2005154（配列番号：1）、rs2184380（配列番号：2）、rs2224696（配列番号：3）、rs2242503（配列番号：4）、rs12291026（配列番号：5）、rs9956546（配列番号：7）、rs11912922（配列番号：8）、rs6001954（配列番号：9）およびそれらと連鎖不平衡のマーカーから選択される少なくとも1つの多型マーカーの少なくとも1つの対立遺伝子を同定するヒト個体についての核酸配列データ入手することであって、前記少なくとも1つの多型マーカーの異なる対立遺伝子は、ヒトにおける乳癌に対する異なる感受性に関連すること、および

40

前記核酸配列データから乳癌に対する感受性を判定することを含む、方法。

【請求項47】

rs999737（配列番号：6）、rs2005154（配列番号：1）、rs2184380（配列番号：2）、rs2224696（配列番号：3）、rs2242503（配列番号：4）、rs12291026（配列番号：5）、rs9956546（配列番号：7）、rs11912922（配列番号：8）、rs6001954（配列番号：9）およびそれらと連鎖不平衡のマーカーから選択される少なくとも2つの多型マーカー

50

ーについての核酸配列データを手に入れたことを含む、請求項 4 6 に記載の方法。

【請求項 4 8】

感受性の判定は、前記核酸配列データを、前記少なくとも 1 つの多型マーカーと乳癌に対する感受性との間の相関データを含むデータベースと比較することを含む、請求項 4 6 または請求項 4 7 に記載の方法。

【請求項 4 9】

前記データベースは前記少なくとも 1 つの多型マーカーについての、乳癌に対する感受性の少なくとも 1 つのリスク判定基準を含む、請求項 4 8 に記載の方法。

【請求項 5 0】

前記データベースは、前記少なくとも 1 つの多型マーカーについての少なくとも 1 つの状態の少なくとも 1 つのリスク判定基準を含む検査表を含む、請求項 4 8 に記載の方法。

10

【請求項 5 1】

核酸配列データを手に入れたことは、前記ヒト個体から生物学的試料を手に入れたこと、および前記試料において核酸中の前記少なくとも 1 つの多型マーカーの配列を分析することを含む、請求項 4 6 から請求項 5 0 のいずれか 1 項に記載の方法。

【請求項 5 2】

前記少なくとも 1 つの多型マーカーの配列を分析することは、前記少なくとも 1 つの多型マーカーの少なくとも 1 つの対立遺伝子の存在または不存在を判定することを含む、請求項 5 1 に記載の方法。

【請求項 5 3】

前記核酸配列データを手に入れたことは、既存の記録から核酸配列情報を手に入れたことを含む、請求項 4 6 から請求項 5 0 のいずれか 1 項に記載の方法。

20

【請求項 5 4】

前記個体、前記個体の保護者、遺伝医療提供者、医師、医療機関、および医療保険会社からなる群から選択される少なくとも 1 つの実体に前記感受性を報告することをさらに含む、請求項 4 6 から請求項 5 3 のいずれか 1 項に記載の方法。

【請求項 5 5】

前記少なくとも 1 つの多型マーカーは表 4 に列挙されるマーカーからなる群から選択される、請求項 4 6 から請求項 5 4 のいずれか 1 項に記載の方法。

【請求項 5 6】

前記少なくとも 1 つの多型マーカーは、rs 9 9 9 7 3 7 (配列番号: 6)、rs 2 0 0 5 1 5 4 (配列番号: 1)、rs 2 1 8 4 3 8 0 (配列番号: 2)、rs 2 2 2 4 6 9 6 (配列番号: 3)、rs 2 2 4 2 5 0 3 (配列番号: 4)、rs 1 2 2 9 1 0 2 6 (配列番号: 5)、rs 9 9 5 6 5 4 6 (配列番号: 7)、rs 1 1 9 1 2 9 2 2 (配列番号: 8)、rs 6 0 0 1 9 5 4 (配列番号: 9) からなる群から選択される、請求項 4 6 から請求項 5 5 のいずれか 1 項に記載の方法。

30

【請求項 5 7】

乳癌対立遺伝子についての少なくとも 1 つのリスク対立遺伝子の存在の判定は、エストロゲン受容体陽性の乳癌またはプロゲステロン受容体陽性の乳癌の前兆となる、請求項 1 から請求項 5 6 のいずれか 1 項に記載の方法、キット、使用、媒体または装置。

40

【請求項 5 8】

マーカー間の連鎖不平衡は、前記連鎖不平衡の基準 r^2 および / または $|D'|$ の特定の数値によって特徴付けられる、請求項 1 から請求項 5 7 のいずれか 1 項に記載の方法、キット、使用、媒体または装置。

【請求項 5 9】

マーカー間の連鎖不平衡は少なくとも 0.2 の r^2 の値によって特徴付けられる、請求項 1 から請求項 5 8 のいずれか 1 項に記載の方法、キット、使用、媒体または装置。

【請求項 6 0】

前記ヒト個体は、ヨーロッパ人の系統を含む系統の個体である、請求項 1 から請求項 5 9 のいずれか 1 項に記載の方法、キット、使用、媒体または装置。

50

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、乳癌の感受性変異体であると判定されたある遺伝的変異体に関する。乳癌に対する増加した感受性を診断することを含む疾病管理の方法、治療に対する反応の予測方法、およびこのような変異体を使用した予後診断の予測方法が記載されている。本発明はさらに、本発明の方法において有用なキットに関する。

【背景技術】

【0002】

乳癌は、世界中の女性において間違いなく最も一般的な癌である。現在の、全世界での発生は、各年で診断された新しい症例において1,151,000を超える（非特許文献1）。乳癌発生率は先進国で最も高く、特に、北欧種族的出身の集団で高く、かつ、増えている。米国における年間の年齢調整発生率は、人口100,000人あたり約12.5症例であり、世界の平均値の3倍以上多い。北欧諸国での割合は同様に高い。2008年には、184,450の湿潤性乳癌の新しい症例が米国において診断され、40,930の人々が当該疾病によって死亡することが推定される（非特許文献2）。この数字に、さらに、2008年に予測される67,770の生体内原位置での腺管癌および小葉癌の診断が加えられなければならない。個々の観点からは、乳癌の発生の生涯可能性は、米国の女性において12.3%である（すなわち、8人中1人の女性は、その生涯の間に乳癌を発生する）。大部分の癌については、早期検出および適切な治療は重要な因子である。全体的に、乳癌の5年生存率は89%である。しかし、局所的浸潤性または転移性疾患を呈する個体においては、割合は、それぞれ84%および27%に減少する（非特許文献2）。

10

20

【0003】

次第に、原発性または再発性の乳癌について高いリスクのある個体の同定に重点が置かれている。このような個体は、より集中的なスクリーニング、予防的化学療法、ホルモン療法および、極度に高いリスクのある個体の場合は予防的手術によって管理され得る。集団検診プログラムは、予防的治療がリスクおよび生活の質の帰結に関連する一方で、公共医療に莫大な経済的負担となる。

【0004】

（乳癌の遺伝的素因）

乳癌の公知の危険因子の2つの主要な分類は、内分泌因子および遺伝的特徴である。後者については、約12%の乳癌患者が、乳癌の1つまたは複数の第一度近親者を有する（非特許文献3）。周知の、優性の乳癌素因遺伝子BRCA1およびBRCA2は、保有者に対して非常に増加した乳癌リスクをもたらし、生涯の浸透率は、40~80%の範囲であると推定される。BRCA1変異およびBRCA2変異の存在は、6またはそれ以上の乳癌の症例の家族の大部分、および乳癌および卵巣癌または男性の乳癌を含む家族の大部分について説明できる。しかし、このような家族は、実際は非常にまれである。BRCA1変異およびBRCA2変異は、より少ない症例の家族、または乳癌の症例のみによって特徴付けられる家族では、かなり少ないことが見出される。合わせて、BRCA1およびBRCA2の変異は、15~20%の家族性乳癌のリスクを説明できる。全ての共通のBRCA変異が検出できると仮定すると、非創始者集団では、2~3%の間の偶発的な乳癌患者が変異を抱えるものと予想される（非特許文献4；非特許文献5）。この低い「見つける機会」の統計量は、明らかな遺伝性素因の家族の外側のBRCA変異試験の責任ある利用を妨げる（非特許文献6）。まれに、高い浸透率の変異がTP53およびPTEN遺伝子において生じることが知られる。しかし、これらはともに、乳癌の全遺伝的危険率のうち、わずか5%しか説明しない（非特許文献7）。連鎖研究は、乳癌について高いリスクを与える広範な変異の同定については、もはや大部分がうまくいっていない（非特許文献8）。

30

40

【0005】

近年の疫学調査は、乳癌の症例の大部分が、集団のうち遺伝的な傾向があり、感受性の

50

ある少数において生じていることを示している（非特許文献 9；非特許文献 10）。一定の高い発生率の原発性乳癌で生存している患者の対側乳房における癌についての双生児研究および観察のデータは、未同定の乳癌のリスクの実質的な部分が、内在性要因に関連し、ほぼ確実に遺伝的に関連していることを示している（非特許文献 11；非特許文献 12）。この広範なリスクを支持する遺伝的要因の知識は、非常に限られる。分離比分析は、乳癌の未同定の遺伝的危険率が、低～中程度のリスクを与え、互いにおよびホルモンの危険因子と相互作用するかもしれないリスク対立遺伝子について、事実上多遺伝子性である可能性が最も高いことを予測する。それにもかかわらず、これらの研究は、これらの低～中程度のリスク対立遺伝子を捕捉する遺伝的特性によって定義され得る分布の五分位数の最も高いものおよび最も低いもの間の相対リスクにおいて、40倍もの量の相違があることを予測する（非特許文献 9；非特許文献 10）。全ての乳癌の症例の 88% は、遺伝的な傾向がある 50% の集団の間で生じることが予測され、12% の最も高いリスクの集団は、全ての乳癌の症例のうちの 50% を説明する（非特許文献 10；非特許文献 13）。従って、このような遺伝的に傾向がある個体の同定、および彼らのために個別化された医学的管理戦略の開発に向けて多くの焦点が向けられる。

10

20

30

40

50

【0006】

本発明者らおよび本発明者らの他も、アイスランドにおいて、少なくとも第 5 度近親者にまで及ぶ有意な乳癌の家族性リスクがあることを示してきた（非特許文献 14；非特許文献 15）。アイスランドにおける家族性リスクへの B R C A 1 変異の寄与は、最小であると考えられる（非特許文献 16；非特許文献 17）。B R C A 2 遺伝子（999 del 5）の単一の創始者変異は、0.6～0.8% の保有頻度で、一般的なアイスランド人の集団において存在し、7.7～8.6% の保有頻度で、女性の乳癌患者において存在する（非特許文献 18；非特許文献 19）。当該単一の変異は、約 40% の第一～三度近親者における遺伝する乳癌リスクを説明するものと推定される（非特許文献 15）。当該推定は、非創始者集団に組み込まれた全ての B R C A 1 および 2 変異に起因する家族性リスクの 15～25% よりも高いが、約 60% のアイスランド人の家族性乳癌リスクが依然として解明されていない。B R C A 2 999 del 5 に陰性と出た、患者の第一度近親者は、1.72 倍の乳癌の集団リスクにとどまる（95% C I 1.49～1.96）（非特許文献 15）。

【0007】

遺伝的危険率は、集団内の個体間の遺伝子におけるわずかな相違によって与えられる。個体間の遺伝子の違いは、他の変異もまた重要であるが、最も頻繁には、一塩基多型（SNP）が原因である。SNP は、ヒトゲノムにおいて、平均 1000 塩基対ごとに位置する。従って、250,000 塩基対を含む典型的なヒト遺伝子は、250 の異なる SNP を有するかも知れない。少数の SNP のみが、エクソンに位置し、当該遺伝子にコードされたタンパク質のアミノ酸配列を変化させる。他の SNP が、その遺伝子によってコードされた mRNA の転写、スプライシング、翻訳または安定性を変化させるかも知れない一方で、多くの SNP は、遺伝子機能に対してほとんど影響がないか、あるいは全く影響がないかも知れない。ヒトゲノムにおける、さらなる遺伝的多型性は、短い DNA の伸長または長い DNA の伸長のいずれかの挿入、欠損、転座または逆位によってもたらされる。疾病リスクを与える遺伝的多型性は、タンパク質のアミノ酸配列を直接に変化させるかも知れず、その遺伝子によって生じるタンパク質の量を増加させるかも知れず、または、その遺伝子によって生じるタンパク質の量を減少させるかも知れない。

【0008】

一般的な疾病のリスクを与える遺伝的多型性は明らかであるため、このような危険因子の遺伝的試験は、臨床医学において重要となってきた。例としては、アルツハイマー病の鑑別診断のために認知症患者における apo E 4 多型性の遺伝的保有者を同定するアポリタンパク質 E 試験、および深部静脈血栓症に対する素因を試験する第 5 因子ライデンの遺伝的保有者を同定するアポリタンパク質 E 試験がある。さらに重要なことには、癌治療においては、腫瘍細胞における遺伝的変異の診断は、個々の患者のための最も適切

な治療計画の選択に使用される。乳癌では、エストロゲン受容体の発現または2型ヘレグリン(Her2)受容体チロシンキナーゼ発現における遺伝的変異が、抗エストロゲン性薬剤(タモキシフェン)または抗Her2抗体(ハーセプチン)が治療計画に組み込まれるかどうかを決定する。フィラデルフィア染色体の慢性骨髄性白血病(CML)診断では、Bcr受容体およびAbl受容体チロシンキナーゼをコードする遺伝子と融合した遺伝的な転座は、Bcr-Ablキナーゼの特異的な阻害剤である、グリベック(STI571)が、当該癌の治療に使用されるべきであることを示す。このような遺伝的変化のあるCML患者においては、Bcr-Ablキナーゼの阻害は、腫瘍細胞の迅速な除去および白血病からの寛解を引き起こす。さらに、遺伝子検査サービスが今は利用でき、特定のSNPが一般的な疾病のうちの多くのリスクに関連付けられたという発見に基づいて、個人に、彼らの疾病のリスクについての情報を提供している。

10

【0009】

未解決の乳癌に対する遺伝的危険率に寄与する遺伝要因の理解は非常に限られる。2つの遺伝子における変異(CHEK2およびATM)は、低浸透率の乳癌リスク遺伝子として正確に確かめられている(非特許文献20;非特許文献21)。さらに、最近の報告は、染色体2q35および16q12上の変異とエストロゲン受容体陽性乳癌の増加したリスクとの間の関連性を確立している(特許文献22)。さらに、FGFR2、TNRC9、MAP3K1およびLSP1遺伝子(特許文献23)、ならびにFGFR2遺伝子(特許文献24)における変異またはそれらの近傍における変異が報告されている。多くの他の遺伝子が関連付けられているが、乳癌のリスクに対するこれらの遺伝子の寄与は、非常に大きいサンプル集合を用いた分析において確かめられていない(非特許文献25)。

20

【先行技術文献】

【非特許文献】

【0010】

【非特許文献1】Parkinら、CA Cancer J Clin、55、74-108、(2005)

【非特許文献2】Jemalら、CA Cancer J Clin、58、71-96、(2008)

【非特許文献3】Lancet、358、1389-99、(2001)

【非特許文献4】Gorskiら、Breast Cancer Res Treat、92、19-24、(2005)

30

【非特許文献5】Br J Cancer、83、1301-8、(2000)

【非特許文献6】Anon、J Clin Oncol、21、2397-406、(2003)

【非特許文献7】Easton、Breast Cancer Res、1、14-7、(1999)

【非特許文献8】Smithら、Genes Chromosomes Cancer、45、646-55、(2006)

【非特許文献9】Antoniouら、Br J Cancer、86、76-83、(2002)

40

【非特許文献10】Pharoahら、Nat Genet、31、33-6、(2002)

【非特許文献11】Lichtensteinら、N Engl J Med、343、78-85、(2000)

【非特許文献12】PetoおよびMack、Nat Genet、26、411-4、(2000)

【非特許文献13】Pharoah、Recent Results Cancer Res、163、7-18;discussion 264-6、(2003)

【非特許文献14】Amundadottirら、PLoS Med、1、e65、(2004)

50

【非特許文献15】Tulinus̄、J Med Genet、39、457-62、(2002)

【非特許文献16】Arason̄、J Med Genet、35、446-9、(1998)

【非特許文献17】Bergthorsson̄、Hum Mutat、Suppl 1、S195-7、(1998)

【非特許文献18】Thorlacius̄、Am J Hum Genet、60、1079-84、(1997)

【非特許文献19】Gudmundsson̄、Am J Hum Genet、58、749-56、(1996)

【非特許文献20】Renwick̄、Nat Genet、38、873-5、(2006)

【非特許文献21】Am J Hum Genet、74、1175-82、(2004)

【非特許文献22】Simon、SN.̄、Nat Genet 39:865-9(2007)

【非特許文献23】Easton、D.F.̄、Nature 447:1087-93、(2007)

【非特許文献24】Hunter、D.J.̄、Nat Genet 39:870-4、(2007)

【非特許文献25】Breast Cancer Association、J Natl Cancer Inst、98、1382-96、(2006)

【発明の概要】

【発明が解決しようとする課題】

【0011】

乳癌の予防または治療について普遍的に成功した方法であって現在利用できるもの、はない。乳癌の管理は、現在、一次予防、早期診断、適切な治療および再発予防の組み合わせに依存している。遺伝子検査をこれらの管理領域の全ての態様に統合させるための明確な臨床上の要請がある。癌感受性遺伝子の同定はまた、操作され得る重要な分子経路（例えば、小分子量または大分子量の薬剤の使用）および、より効果的な治療をもたらし得る重要な分子経路を明らかにするかも知れない。本発明は、乳癌についての予防プログラムに統合することができる、乳癌についてのさらなる遺伝的変異を提供する。

【課題を解決するための手段】

【0012】

本発明は、乳癌に対する感受性の評価方法に関する。本発明は、乳癌に対する増加した、または減少した感受性に関連することが見出されたあるマーカーまたはハプロタイプを評価することによる、乳癌に対する増加した感受性の診断方法、および乳癌に対する減少した感受性の診断方法もしくは癌に対する保護の診断方法、を含む。本発明はまた、乳癌と診断された個体の予後診断の評価方法、乳癌治療薬もしくは乳癌治療に対する反応の可能性の評価方法、および乳癌と診断された個体の治療の進行をモニターする方法に関する。

【0013】

1つの態様では、本発明は、ヒト個体において乳癌に対する感受性を診断する方法に関し、当該方法は、当該個体から入手した核酸試料の、rs999737（配列番号：6）、rs2005154（配列番号：1）、rs2184380（配列番号：2）、rs224696（配列番号：3）、rs2242503（配列番号：4）、rs12291026（配列番号：5）、rs9956546（配列番号：7）、rs11912922（配列番号：8）、rs6001954（配列番号：9）およびそれらと連鎖不平衡のマーカーからなる群から選択される少なくとも1つの多型マーカーの少なくとも1つの対立遺伝子の存在または不存在を判定することを含み、当該少なくとも1つの対立遺伝子の存

10

20

30

40

50

在は乳癌に対する感受性を示す。本発明はまた、rs2005154（配列番号：1）、rs2184380（配列番号：2）、rs2224696（配列番号：3）、rs2242503（配列番号：4）、rs12291026（配列番号：5）、rs999737（配列番号：6）、rs9956546（配列番号：7）、rs11912922（配列番号：8）、rs6001954（配列番号：9）およびそれらと連鎖不平衡のマーカールからなる群から選択される少なくとも1つの多型の少なくとも1つの対立遺伝子の存在または不存在を判定することにより、乳癌に対する感受性を判定する方法に関し、当該少なくとも1つの対立遺伝子の存在の判定は乳癌に対する感受性を示す。

【0014】

別の態様では、本発明はヒト個体における乳癌に対する感受性の判定方法に関し、当該方法は、少なくとも1つの多型マーカール中の少なくとも1つのリスクのある対立遺伝子が当該個体に由来する遺伝子型データセットの中に存在するかどうかを判定することを含み、当該少なくとも1つの多型マーカールは、rs2005154（配列番号：1）、rs2184380（配列番号：2）、rs2224696（配列番号：3）、rs2242503（配列番号：4）、rs12291026（配列番号：5）、rs999737（配列番号：6）、rs9956546（配列番号：7）、rs11912922（配列番号：8）、rs6001954（配列番号：9）およびそれらと連鎖不平衡のマーカールからなる群から選択されるマーカールから選択され、当該少なくとも1つのリスクのある対立遺伝子の存在の判定は、当該個体における乳癌に対する増加した感受性を示す。

【0015】

本発明はさらに、ヒト個体における乳癌に対する感受性の判定方法に関し、当該方法は、少なくとも1つの多型マーカールの少なくとも1つの対立遺伝子が当該個体から入手した核酸試料の中にまたは当該個体由来の遺伝子型データセットの中に存在するかどうかを判定することを含み、当該少なくとも1つの多型マーカールは、rs2005154（配列番号：1）、rs2184380（配列番号：2）、rs2224696（配列番号：3）、rs2242503（配列番号：4）、rs12291026（配列番号：5）、rs999737（配列番号：6）、rs9956546（配列番号：7）、rs11912922（配列番号：8）、rs6001954（配列番号：9）およびそれらと連鎖不平衡のマーカールから選択され、当該少なくとも1つの対立遺伝子の存在は、当該個体についての乳癌に対する感受性を示す。

【0016】

当該遺伝子型データセットは、1つの実施態様では、マーカール同一性、および当該個体の対立遺伝子の状態についての情報、すなわち当該マーカールについての当該個体が保有する当該2つの対立遺伝子の同一性についての情報を含む。当該遺伝子型データセットは、1以上のマーカール（2以上のマーカール、3以上のマーカール、5以上のマーカール、100以上のマーカールなどを含む）についての対立遺伝子の情報を含み得る。いくつかの実施態様では、当該遺伝子型データセットは、数十万ものマーカール、またはさらには百万以上のマーカールを含み得る当該個体の全ゲノム評価から得た遺伝子型情報を含む。

【0017】

ある実施態様では、当該少なくとも1つの多型マーカールは、PAX5（PAIRED BOX GENE 5、BSAPとしても知られている）遺伝子、TUB（Tubby homolog（マウス、rd5））遺伝子、SERPINH1（セルピンペプチダーゼ阻害剤、clade H（ヒートショックタンパク質47）、member 1、（コラーゲン結合タンパク質1））遺伝子、RAD51L1（REC2；R51H2；hREC2；RAD51B；MGC34245）遺伝子、FHOD3（formin homolog 2 domain containing 3、FHOS2としても知られている；Formactin2）遺伝子またはTNRC6B（trinucleotide repeat containing 6BまたはKIAA1093）遺伝子に関連する。

【0018】

10

20

30

40

50

特定のこのような実施態様では、当該少なくとも1つの多型マーカーは、上記の遺伝子のうちのいずれか1つ、すなわちPAX5 (PAIRED BOX GENE 5、BSAPとしても知られる) 遺伝子、TUB (Tubby homolog (マウス、rd5)) 遺伝子、SERPINH1 (セルピンペプチダーゼ阻害剤、clade H (ヒートショックタンパク質47)、member 1、(コラーゲン結合タンパク質1)) 遺伝子、RAD51L1 (REC2; R51H2; hREC2; RAD51B; MGC34245) 遺伝子、FHOD3 (formin homology 2 domain containing 3、FHOS2としても知られる; Formactin2) 遺伝子およびTNRC6B (trinucleotide repeat containing 6BまたはKIAA1093) 遺伝子のうちのいずれか1つと連鎖不平衡である。

10

【0019】

本発明の別の態様はヒト個体において乳癌に対する感受性を判定する方法に関し、当該方法は、

rs2005154 (配列番号: 1)、rs2184380 (配列番号: 2)、rs2224696 (配列番号: 3)、rs2242503 (配列番号: 4)、rs1291026 (配列番号: 5)、rs999737 (配列番号: 6)、rs9956546 (配列番号: 7)、rs11912922 (配列番号: 8)、rs6001954 (配列番号: 9)、およびそれらと連鎖不平衡のマーカーから選択される少なくとも1つの多型マーカーの少なくとも1つの対立遺伝子を同定するヒト個体についての核酸配列データを入力することであって、当該少なくとも1つの多型マーカーの異なる対立遺伝子はヒトにおける乳癌に対する異なる感受性に関連する、こと; および

20

当該核酸配列データから乳癌に対する感受性を判定することを含む。

【0020】

一般的に、遺伝子マーカーは、核酸レベルで別の配列に導く。当該核酸マーカーが当該核酸によってコードされるポリペプチドのコドンを変える場合、当該マーカーはまた、コードされたポリペプチド (ポリペプチドマーカー) のアミノ酸レベルで別の配列を生じることになる。ポリペプチドマーカーにおける核酸または特定の対立遺伝子の中の、多型マーカーにおける特定の対立遺伝子の同一性の判定は、特定の対立遺伝子が当該配列の中の特定の位置に存在するかどうかを含む。マーカーで特定の対立遺伝子を同定する配列データは、ある対立遺伝子を検出するための十分な配列を含む。本願明細書に記載された一塩基多型 (SNP) またはアミノ酸多型については、配列データは、1つの位置での配列、すなわち配列内の1つの位置でのヌクレオチドまたはアミノ酸の同一性を含み得る。多型マーカーの特定の対立遺伝子を同定する核酸配列データは、遺伝子型データとも呼ばれることがある。核酸配列データは、例えば当該個体由来の生物学的試料において当該少なくとも1つの多型マーカーの配列を分析することにより、得られ得る。あるいは、核酸配列データは、当該ヒト個体由来の遺伝子型データセットの中で、および当該データセットの中で当該少なくとも1つの多型マーカーの配列を分析することで、得られ得る。このような分析は、ある実施態様では、特定の多型マーカーの特定の対立遺伝子の存在または不存在を判定することを含む。特定の対立遺伝子の同定は、一般論として、当該対立遺伝子 (1種または複数種) の存在または不存在の判定がなされることを意味すると解釈されるべきである。通常、個体のゲノムの中の両方の対立遺伝子のコピーの判定は、特定の個体の中の当該特定の多型性のすべての可能性のある対立遺伝子 (SNPについて、当該対立遺伝子の部位について可能性のある2つの可能性のあるヌクレオチドの各々) の存在を判定することによって行われる。特定の対立遺伝子のみが存在するかまたは存在しないかを判定することも可能である。例えば、ある実施態様では、緑内障のリスクと関連することが示された特定の対立遺伝子の存在または不存在の判定がなされるが、当該特定のマーカーの他の対立遺伝子の存在または不存在の判定は必ずしもなされず、感受性の判定はこのような判定に基づいてなされる。

30

40

【0021】

50

ある実施態様では、少なくとも2つの多型マーカーについての核酸配列を判定することが有用であり得る。他の実施態様では、少なくとも3つ、少なくとも4つまたは少なくとも5つまたはこれより多くの多型マーカーについての核酸配列が判定される。ハプロタイプ情報は、2以上の多型マーカーの分析から誘導され得る。従って、ある実施態様では、さらなる工程が実施され、それにより、少なくとも2つの多型マーカーについての配列データに基づいてハプロタイプ情報が誘導される。

【0022】

本発明はまた、ヒト個体における乳癌に対する感受性の判定方法を提供し、当該方法は、rs2005154（配列番号：1）、rs2184380（配列番号：2）、rs2224696（配列番号：3）、rs2242503（配列番号：4）、rs12291026（配列番号：5）、rs999737（配列番号：6）、rs9956546（配列番号：7）、rs11912922（配列番号：8）、rs6001954（配列番号：9）およびそれらと連鎖不平衡のマーカーから選択される少なくとも2つの多型マーカーの両方の対立遺伝子を同定するヒト個体についての核酸配列データを入手すること、当該配列データに基づいて少なくとも1つのハプロタイプの同一性を判定すること、ならびに当該ハプロタイプデータから乳癌に対する感受性を判定することを含む。

10

【0023】

ある実施態様では、感受性の判定は、当該核酸配列データを、当該少なくとも1つの多型マーカーと乳癌に対する感受性との間の相関データを含むデータベースと比較することを含む。いくつかの実施態様では、当該データベースは、当該少なくとも1つのマーカーについての乳癌に対する感受性の少なくとも1つのリスク判定基準を含む。当該配列データベースは、例えば、いずれか1つの、または複数の、特定の多型についての乳癌の感受性を示すデータを含む検査表として提供され得る。当該データベースはまた、少なくとも2つの多型マーカーを含む特定のハプロタイプについての感受性を示すデータを含み得る。

20

【0024】

核酸配列データを入手することは、ある実施態様では、当該ヒト個体から生物学的試料を入手すること、および当該試料において核酸中の当該少なくとも1つの多型マーカーの配列を分析することを含み得る。配列を分析することは、当該少なくとも1つの多型マーカーの少なくとも1つの対立遺伝子の存在または不存在を判定することを含み得る。特定の感受性対立遺伝子（例えば、リスクのある対立遺伝子）の存在の判定は、当該ヒト個体における乳癌に対する感受性を示す。特定の感受性対立遺伝子の不存在の判定は、当該特定の感受性が当該個体の中には存在しないことを示す。

30

【0025】

いくつかの実施態様では、核酸配列データを入手することは、既存の記録から核酸配列情報を入手することを含む。当該既存の記録は、例えば、当該ヒト個体についての、少なくとも1つの多型マーカーについての配列データ（遺伝子型データなど）を含むコンピューターファイルまたはデータベースであり得る。

【0026】

本発明の診断方法によって判定される感受性は、特定の实体に報告され得る。いくつかの実施態様では、当該少なくとも1つの实体は、当該個体、当該個体の保護者、遺伝医療提供者、医師、医療機関、および医療保険会社からなる群から選択される。

40

【0027】

本発明のある実施態様では、感受性の判定は、当該核酸配列データを、当該少なくとも1つの多型マーカーと乳癌に対する感受性との間の相関データを含むデータベースと比較することを含む。1つのこのような実施態様では、当該データベースは、当該少なくとも1つの多型マーカーについての乳癌に対する感受性の少なくとも1つのリスク判定基準を含む。別の実施態様では、当該データベースは、当該少なくとも1つの多型マーカーについての少なくとも1つの状態の少なくとも1つのリスク判定基準を含む検査表を含む。

【0028】

50

ある実施態様では、核酸配列データ入手することは、当該ヒト個体から生物学的試料入手すること、および当該試料において核酸中の当該少なくとも1つの多型マーカの配列を分析することを含む。当該少なくとも1つの多型マーカの配列を分析することは、当該少なくとも1つの多型マーカの少なくとも1つの対立遺伝子の存在または不存在を判定することを含み得る。核酸配列データ入手ことはまた、既存の記録から核酸配列情報入手することを含み得る。

【0029】

本発明の特定の実施態様は、rs2005154（配列番号：1）、rs2184380（配列番号：2）、rs2224696（配列番号：3）、rs2242503（配列番号：4）、rs12291026（配列番号：5）、rs999737（配列番号：6）、rs9956546（配列番号：7）、rs11912922（配列番号：8）、rs6001954（配列番号：9）およびそれらと連鎖不平衡のマーカから選択される少なくとも2つの多型マーカについての核酸配列データ入手することに関する。

10

【0030】

本発明のある実施態様では、当該少なくとも1つの多型マーカは、表4に記載されるマーカから選択される。1つの実施態様では、当該少なくとも1つの多型マーカは、配列番号：1～562に記載されるマーカから選択される。1つの実施態様では、当該少なくとも1つのマーカはrs2005154、rs2184380、rs2224696、rs2242503、rs12291026、rs999737、rs9956546、rs11912922およびrs6001954のうちの少なくとも1つと連鎖不平衡である。1つの実施態様では、rs999737と連鎖不平衡のマーカはrs999737、rs10134446、rs10138140、rs10146772、rs10467820、rs10483812、rs10483813、rs11158749、rs11158751、rs11621276、rs11624097、rs11624164、rs11624333、rs11628293、rs11846360、rs11847185、rs11849916、rs12878761、rs12879200、rs12886864、rs12889251、rs12894230、rs1468279、rs1468280、rs1547012、rs17105675、rs17755657、rs17755734、rs17755752、rs17755925、rs17756000、rs17828691、rs17828721、rs17828763、rs17828907、rs17828955、rs1956534、rs2074563、rs2074565、rs2097800、rs2107340、rs2145157、rs2158357、rs2189517、rs2253317、rs2257111、rs2257116、rs2257127、rs2331701、rs2331705、rs2331775、rs2525503、rs2525530、rs2842327、rs3784121、rs4531674、rs4899246、rs4902604、rs4902606、rs4902608、rs5004090、rs6573837、rs7140266、rs7146456、rs7153476、rs739874、rs746663、rs8007194、rs8010439、rs8012610、rs9323512、rs9323513、およびrs9323514からなる群から選択される。

20

30

40

【0031】

1つの実施態様では、rs2005154と連鎖不平衡のマーカは、rs2005154、rs4878662、およびrs4880019からなる群から選択される。

【0032】

1つの実施態様では、rs2184380と連鎖不平衡のマーカはrs2184380、rs10466295、rs10508363、rs10508364、rs10508365、rs10795670、rs10905411、rs10905414、rs10905415、rs10905430、rs10905437、rs10905439、rs10905440、rs10905443、rs10905444、rs10

50

905445、rs10905446、rs10905447、rs10905454、
rs11255764、rs11255776、rs11255777、rs11255
778、rs11255779、rs11255790、rs11255795、rs1
1255797、rs11255800、rs11255804、rs11255805
、rs11255820、rs11255821、rs11255822、rs1125
5832、rs11255836、rs11255840、rs11255858、rs
11255862、rs11255869、rs11255870、rs1125587
1、rs11255882、rs11255884、rs12049705、rs122
18610、rs12250379、rs12259226、rs1325874、rs
1334549、rs1334550、rs1334559、rs1360749、rs 10
1413678、rs1413683、rs1537601、rs1537602、rs
1537603、rs17407711、rs17407781、rs17407830
、rs17408204、rs17408337、rs17408580、rs1748
4150、rs17485426、rs17485998、rs17486082、rs
17486795、rs17486816、rs1970170、rs1999638、
rs2031561、rs2182292、rs2388821、rs2388825、
rs2388826、rs2892613、rs4112287、rs4112288、
rs4345867、rs4454616、rs4747806、rs4749805、
rs4749807、rs4749812、rs6602328、rs6602329、
rs7069110、rs7080765、rs7083359、rs7477023、 20
rs7904921、rs7912413、rs7912704、rs7912831、
rs827389、およびrs966562からなる群から選択される。

【0033】

1つの実施態様では、rs2224696と連鎖不平衡のマーカ-はrs222469
6、rs10905509、rs11256045、rs12761213、rs127
61461、rs12766048、rs12772042、rs12776383、r
s12778120、rs12780218、rs12781427、rs147518
9、rs1573109、rs1573110、rs17145088、rs17145
095、rs17145118、rs17145120、rs17145151、rs1
7145164、rs17145169、rs17145188、rs17145193 30
、rs17145221、rs17363338、rs1775559、rs18572
30、rs1891532、rs1935813、rs2013364、rs20252
89、rs2057442、rs2093625、rs2093626、rs21465
98、rs2185817、rs2397336、rs2760204、rs27972
66、rs391733、rs4550140、rs7081544、rs852273
、rs860418、rs861172、rs962993、およびrs965307か
らなる群から選択される。

【0034】

1つの実施態様では、rs2242503と連鎖不平衡のマーカ-はrs224250
3、rs10431029、rs1055233、rs10734629、rs1074 40
3052、rs10743053、rs10743054、rs10743055、rs
10769872、rs10769873、rs10769878、rs1076988
2、rs10839976、rs10839984、rs11041740、rs110
41742、rs11041788、rs11041791、rs11041794、r
s1108277、rs12146654、rs12808387、rs1528125
、rs1569128、rs1970880、rs1997262、rs2049684
、rs2141321、rs2242501、rs2272383、rs3750955
、rs3752898、rs3849986、rs3849990、rs3911309
、rs3911310、rs4340037、rs4343012、rs4385931
、rs4575312、rs4578424、rs4636658、rs4758040 50

、rs4758042、rs4758287、rs4758309、rs4758310
 、rs7103334、rs7112519、rs7115706、およびrs7122
 690、rs7127738、rs7358396、rs7479156、rs7479
 738、rs7480804、rs7481667、rs7481683、rs7482
 611、rs7927368、およびrs7940668からなる群から選択される。

【0035】

1つの実施態様では、rs12291026と連鎖不平衡のマーカはrs12291
 026、rs1004856、rs10899091、rs11236449、rs11
 236452、rs11236454、rs12362081、rs1540210、r
 s1540211、rs1557471、rs1631470、rs1783551、r
 s1783556、rs1783559、rs1790144、rs1790152、r
 s1790307、rs1793396、rs1793397、rs1793398；r
 s1793399、rs1793414、rs1938800、rs2853066、r
 s499613、rs504793、rs514477、rs549034、rs550
 881、rs581007、rs589724、rs600387、rs606460、
 rs617617、rs618202、rs628972、rs640649、rs66
 2279、rs667410、rs667531、rs670100、rs670491
 、rs682292、rs7128888、rs7129014、rs7129150、
 およびrs947844からなる群から選択される。

10

【0036】

1つの実施態様では、rs11912922と連鎖不平衡のマーカは、rs1191
 2922、rs11089967、rs11704971、rs11705454、rs
 17406386、rs17406434、rs2071771、rs2958650、
 rs2958651、rs2958659、rs7284488、rs7285507、
 rs7291782、rs739145、rs9611246、およびrs961126
 5からなる群から選択される。

20

【0037】

1つの実施態様では、rs6001954と連鎖不平衡のマーカはrs600195
 4、rs10483203、rs10483204、rs10483205、rs104
 83206、rs1106673、rs11913132、rs12158399、rs
 12158872、rs12159200、rs12159970、rs1248469
 7、rs12627881、rs133036、rs133038、rs1698589
 9、rs17001846、rs17001868、rs17001943、rs170
 01974、rs17001977、rs17001993、rs17001994、r
 s17001997、rs17002019、rs17002020、rs170020
 26、rs17002027、rs17002034、rs17002036、rs17
 002038、rs17002069、rs2075764、rs2187832、rs
 2235318、rs2280790、rs2294348、rs2294350、rs
 2294352、rs2413624、rs3788577、rs3788578、rs
 3788579、rs3827381、rs3827382、rs4140512、rs
 470113、rs5750957、rs5750960、rs5750966、rs5
 757976、rs5757998、rs5758001、rs5995849、rs5
 995856、rs5995870、rs5995871、rs5995886、rs6
 001900、rs6001910、rs6001911、rs6001912、rs6
 001913、rs6001930、rs6001931、rs6001932、rs6
 001935、rs6001950、rs6001974、rs6001980、rs6
 001990、rs6002000、rs718193、rs7292804、rs72
 93100、rs742140、rs760700、rs760701、rs93063
 45、rs932379、rs9607721、rs9611310、rs961131
 1、rs9611312、rs9611316、rs9611318、rs961132

30

40

50

4、rs9611325、rs9611328、およびrs9611329からなる群から選択される。

【0038】

本発明のある実施態様では、当該個体における少なくとも1つのハプロタイプの頻度を評価するというさらなる工程が実施される。このような実施態様では、2以上のマーカー（3、4、5、6、7、8、9または10またはこれより多いマーカーを含む）が当該ハプロタイプの中に含まれ得る。ある実施態様では、当該少なくとも1つのハプロタイプは、すべてrs2005154（配列番号：1）、rs2184380（配列番号：2）、rs2224696（配列番号：3）、rs2242503（配列番号：4）、rs12291026（配列番号：5）、rs999737（配列番号：6）、rs9956546（配列番号：7）、rs11912922（配列番号：8）およびrs6001954（配列番号：9）のうちの少なくとも1つとLDであるマーカーを含む。このような実施態様では、当該少なくとも1つのハプロタイプは、上記のマーカーのうちのいずれか1つが属する特定のゲノムの領域（LDブロック）のゲノムの構造を代表する。1つの実施態様では、当該ハプロタイプはrs999737と連鎖不平衡のマーカーを含む。

10

【0039】

本願明細書に記載された乳癌のリスクを与えるマーカーは、乳癌についての他の遺伝子マーカーと組み合わせられ得る。このようなマーカーは、典型的には、本願明細書に記載されたマーカー、特にマーカーrs2005154（配列番号：1）、rs2184380（配列番号：2）、rs2224696（配列番号：3）、rs2242503（配列番号：4）、rs12291026（配列番号：5）、rs999737（配列番号：6）、rs9956546（配列番号：7）、rs11912922（配列番号：8）、rs6001954（配列番号：9）およびそれらと連鎖不平衡のマーカーのいずれの1つとも連鎖不平衡でない。本願明細書に記載された方法のいずれも、本願明細書に記載された遺伝的危険因子を乳癌についてのさらなる遺伝的危険因子と併用することにより実施され得る。

20

【0040】

従って、ある実施態様では、表1および/または表4に記載されるマーカーのうちのいずれか1つと連鎖不平衡ではない乳癌についての少なくとも1つのリスクのある変異体の少なくとも1つのリスクのある対立遺伝子が個体中に存在するかどうかを判定することを含む、さらなる工程が含まれる。換言すると、ゲノム中の他の位置の遺伝子マーカーは、複数の遺伝的要因に基づく乳癌の全体のリスクを判定するために、本発明のマーカーとの併用において有用であり得る。連鎖不平衡でない（LDでない）マーカーの選択は、さらに本願明細書に記載されるように、連鎖不平衡についての適した基準に基づき得る。ある実施態様では、連鎖不平衡でないマーカーは、0.2未満のマーカー間のLD基準 r^2 の値を有する。ある他の実施態様では、LDでないマーカーは、0.10未満、0.05未満、0.02未満および0.01未満を含む、0.15未満のマーカー間の r^2 値を有する。これらの値のいずれかをつなぐ値を含む、マーカーがLDでないことを定める他の適したカットオフ値が考えられる。このような遺伝的危険因子の例としては、染色体5p12および染色体10q26上のマーカー、例えば、マーカーrs10941679およびマーカーrs1219648が挙げられる。あるいは、これらのマーカーのうちのいずれか1つとLDのマーカーは評価され得る。乳癌のリスクを与えることが公知の他のマーカーもまた、本願明細書に記載されたマーカーと一緒に評価され得、その例としては、染色体2q14上のマーカー（例えば、マーカーrs4848543またはそれらと連鎖不平衡のマーカー）、2q35上のマーカー（例えば、マーカーrs13387042、またはそれらと連鎖不平衡のマーカー）、および染色体16上のマーカー（例えば、マーカーrs3803662、またはそれらと連鎖不平衡のマーカー）が挙げられる。

30

40

【0041】

ある実施態様では、本願明細書に記載される複数のマーカーは、乳癌の全体のリスクを判定するために決定される。従って、ある実施態様では、さらなる工程が含まれ、当該工

50

程は、少なくとも2つの多型マーカーのそれぞれの少なくとも1つの対立遺伝子が、ヒト個体から得たゲノムDNAを含む試料、またはヒト個体由来の遺伝子型データセットに存在するかどうかについての判定を含み、当該少なくとも2つの多型マーカーの少なくとも1つの対立遺伝子の存在は、乳癌に対する増加した感受性を示す。1つの実施態様では、当該マーカーは、rs2005154（配列番号：1）、rs2184380（配列番号：2）、rs2224696（配列番号：3）、rs2242503（配列番号：4）、rs12291026（配列番号：5）、rs999737（配列番号：6）、rs9956546（配列番号：7）、rs11912922（配列番号：8）およびrs6001954（配列番号：9）ならびにそれらと連鎖不平衡のマーカーから選択される。

【0042】

本発明のマーカーに基づくリスクアセスメントはまた、当該個体から入手した核酸試料または当該個体由来の遺伝子型データセット中の少なくとも1つの高い浸透性の乳癌の遺伝要因の存在または不存在に対する評価と併用され得る。当該高い浸透性の乳癌の遺伝要因は、例えば、BRCA1変異、BRCA2変異、TP53変異またはPTEN変異であり得る。合わせて、BRCA1変異およびBRCA2変異は、家族性乳癌のリスクの15～20%を説明することができ、これらは、乳癌の患者の発生の2～3%を説明できる[Gorskiら、(2005)、Breast Cancer Res Treat、92、19-24；(2000)、Br J Cancer、83、1301-8]。TP53遺伝子およびPTEN遺伝子の既知の変異は、乳癌の全体の遺伝的危険率の約5%を説明する[Easton、(1999)、Breast Cancer Res、1、14-7]。1つの実施態様では、当該高い浸透性の遺伝要因は、BRCA2 999del5である。

【0043】

本発明の遺伝子マーカーはまた、個体の全体のリスクを定めるために、非遺伝的情報と併用され得る。従って、ある実施態様では、当該個体のリスクアセスメント、診断、または予後診断を行うために、非遺伝的情報を分析することを含む、さらなる工程が含まれる。当該非遺伝的情報は、当該個体の疾病の状態に関連するいずれの情報、または当該個体の乳癌の全体のリスクの評価に影響し得る他の情報でもあり得る。1つの実施態様では、当該非遺伝的情報は、年齢、性別、民族性、社会経済的状況、以前の疾病診断、患者の病歴、乳癌の家族歴、生化学的測定結果、および臨床的測定結果から選択される。

【0044】

別の態様では、本発明は、以前に乳癌と診断された個体の少なくとも続発性原発腫瘍の発生リスクの評価方法に関し、当該方法は、当該個体から入手した核酸試料中の少なくとも1つの多型マーカーの少なくとも1つの対立遺伝子の存在または不存在の判定を含み、当該少なくとも1つの多型マーカーは、表1および表4に記載された多型マーカー、ならびにそれらと連鎖不平衡のマーカーからなる群から選択され、当該少なくとも1つの対立遺伝子の存在は、少なくとも続発性原発腫瘍の発生リスクを示す。あるいは、本発明は、以前に乳癌と診断された個体の少なくとも続発性原発腫瘍の発生リスクの判定方法に関し、当該方法は、少なくとも1つの多型マーカーの少なくとも1つの対立遺伝子が当該個体から入手した核酸試料または当該個体由来の遺伝子型データセット中に存在するかどうかについての判定を含み、当該少なくとも1つの多型マーカーは、rs2005154（配列番号：1）、rs2184380（配列番号：2）、rs2224696（配列番号：3）、rs2242503（配列番号：4）、rs12291026（配列番号：5）、rs999737（配列番号：6）、rs9956546（配列番号：7）、rs11912922（配列番号：8）およびrs6001954（配列番号：9）ならびにそれらと連鎖不平衡のマーカーから選択され、当該少なくとも1つの対立遺伝子の存在は少なくとも続発性原発腫瘍の発生リスクを示す。このような実施態様の1つでは、当該少なくとも1つの多型マーカーは、表4に記載されるマーカーから選択される。

【0045】

本発明はまた、コンピューターで実施される態様を提供する。1つのこのような態様で

10

20

30

40

50

は、本発明は、個体における乳癌に対する感受性を判定するためのコンピューターが実行可能な命令を有するコンピューター可読媒体を提供し、このコンピューター可読媒体は、

少なくとも1つの多型マーカ-を表すデータ；および

コンピューター可読媒体に保存され、かつ個体の中の当該少なくとも1つの多型マーカ-の少なくとも1つの対立遺伝子の対立遺伝子の状態に基づいて当該個体における乳癌に対する感受性を判定するためにプロセッサによって実行されるように適合されたルーチンを含む。

【0046】

1つの実施態様では、少なくとも1つの多型マーカ-を表す当該データは、当該少なくとも1つの多型マーカ-と関係がある乳癌に対する感受性を示す少なくとも1つのパラメータを含む。別の実施態様では、少なくとも1つの多型マーカ-を表す当該データは、当該個体の中の当該少なくとも1つの対立遺伝子のマーカ-の少なくとも1つの対立遺伝子の対立遺伝子の状態を示すデータを含む。別の実施態様では、当該ルーチンは、当該個体の中の当該少なくとも1つの対立遺伝子のマーカ-の少なくとも1つの対立遺伝子についての当該対立遺伝子の状態を示す入力データを受信するように適合されている。好ましい実施態様では、当該少なくとも1つのマーカ-は、rs2005154（配列番号：1）、rs2184380（配列番号：2）、rs2224696（配列番号：3）、rs2242503（配列番号：4）、rs12291026（配列番号：5）、rs999737（配列番号：6）、rs9956546（配列番号：7）、rs11912922（配列番号：8）、rs6001954（配列番号：9）およびそれらと連鎖不平衡のマーカ-から選択される。別の好ましい実施態様では、当該少なくとも1つの多型マーカ-は表4に記載されるマーカ-から選択される。

10

20

【0047】

本発明はさらに、ヒト個体において乳癌についての遺伝的指標を判定するための装置を提供し、当該装置は、

プロセッサ、

コンピューターが実行可能な命令を有するコンピューター可読メモリであって、乳癌に関して少なくとも1人のヒト個体のマーカ-および/またはハプロタイプの情報を分析し、当該マーカ-またはハプロタイプの情報に基づいて結果を生じるためにプロセッサ上で実行されるように適合されたコンピューターが実行可能な命令を有し、当該結果は、当該ヒト個体についての乳癌の遺伝的指標として、少なくとも1つのマーカ-またはハプロタイプのリスク判定基準を含む、コンピューター可読メモリ

30

を含む。1つの実施態様では、当該コンピューター可読メモリは、乳癌と診断された複数の個体における少なくとも1つの多型マーカ-の少なくとも1つの対立遺伝子または少なくとも1つのハプロタイプの頻度を示すデータ、および複数の参照個体における少なくとも1つの多型マーカ-の少なくとも1つの対立遺伝子または少なくとも1つのハプロタイプの頻度を示すデータを含み、リスク判定基準は、当該ヒト個体についての少なくとも1つのマーカ-および/またはハプロタイプ状態と、乳癌と診断された複数の個体についての少なくとも1つのマーカ-および/またはハプロタイプ情報の頻度を示すデータとの比較に基づく。別の実施態様では、当該コンピューター可読メモリはさらに、少なくとも1つの多型マーカ-の少なくとも1つの対立遺伝子または少なくとも1つのハプロタイプに関連する乳癌の発生のリスクを示すデータを含み、当該ヒト個体についてのリスク判定基準は、当該ヒト個体についての遺伝子型の状態と、少なくとも1つの多型マーカ-の少なくとも1つの対立遺伝子または当該少なくとも1つのハプロタイプに関連するリスクとの比較に基づく。別の実施態様では、当該コンピューター可読メモリはさらに、乳癌と診断された複数の個体における少なくとも1つの多型マーカ-の少なくとも1つの対立遺伝子または少なくとも1つのハプロタイプの頻度を示すデータ、および複数の参照個体における少なくとも1つの多型マーカ-の少なくとも1つの対立遺伝子または少なくとも1つのハプロタイプの頻度を示すデータを含み、乳癌の発生のリスクは、乳癌と診断された個体、および参照個体における当該少なくとも1つの対立遺伝子またはハプロタイプの頻度の

40

50

比較に基づく。好ましい実施態様では、当該少なくとも1つのマーカーは、rs2005154（配列番号：1）、rs2184380（配列番号：2）、rs2224696（配列番号：3）、rs2242503（配列番号：4）、rs12291026（配列番号：5）、rs999737（配列番号：6）、rs9956546（配列番号：7）、rs11912922（配列番号：8）、rs6001954（配列番号：9）およびそれらと連鎖不平衡のマーカーから選択される。別の好ましい実施態様では、当該少なくとも1つの多型マーカーは、表4に記載されるマーカーから選択される。

【0048】

別の態様では、本発明は、乳癌に対する感受性を評価するために使用するためのマーカーの同定方法に関し、当該方法は、rs2005154（配列番号：1）、rs2184380（配列番号：2）、rs2224696（配列番号：3）、rs2242503（配列番号：4）、rs12291026（配列番号：5）、rs999737（配列番号：6）、rs9956546（配列番号：7）、rs11912922（配列番号：8）およびrs6001954（配列番号：9）のうち少なくとも1つと連鎖不平衡の少なくとも1つの多型マーカーを同定すること；乳癌と診断された個体または乳癌に対する感受性を有する個体の試料の遺伝子型の状態を判定すること；および対照個体の試料の遺伝子型の状態を判定することを含み；対照試料中の少なくとも1つの対立遺伝子の頻度と比較した、乳癌と診断された個体または乳癌に対する感受性を有する個体における少なくとも1つの多型性の少なくとも1つの対立遺伝子の頻度の有意差は、当該少なくとも1つの多型性が、乳癌に対する感受性の評価に有用であることを示す。有意差は、乳癌患者および対照のある多型マーカーにおける対立遺伝子数の統計解析上で評価され得る。1つの実施態様では、有意差は、0.05未満の、乳癌患者および対照間の算出されたP値に基づく。他の実施態様では、有意差は、算出されたP値のより低い値、例えば0.005未満、0.0005未満、または0.00005未満などに基づく。1つの実施態様では、対照試料の少なくとも1つの対立遺伝子の頻度と比較した、乳癌と診断された個体または乳癌に対する感受性を有する個体における少なくとも1つの多型性における、少なくとも1つの対立遺伝子の頻度の増加は、当該少なくとも1つの多型性が、乳癌に対する増加した感受性の評価に有用であることを示す。別の実施態様では、対照試料中の少なくとも1つの対立遺伝子の頻度と比較した、乳癌と診断された個体または乳癌に対する感受性を有する個体における少なくとも1つの多型性における少なくとも1つの対立遺伝子の頻度の減少は、当該少なくとも1つの多型性が、乳癌に対する減少した感受性、または乳癌に対する保護の評価に有用であることを示す。

【0049】

本発明はまた、ヒト個体から入手した核酸試料の遺伝子型を同定する方法であって、少なくとも1つの多型マーカーの少なくとも1つの対立遺伝子が、当該個体試料から得た核酸試料中に存在するかどうかの判定をする工程を含み、当該少なくとも1つのマーカーは、rs2005154（配列番号：1）、rs2184380（配列番号：2）、rs2224696（配列番号：3）、rs2242503（配列番号：4）、rs12291026（配列番号：5）、rs999737（配列番号：6）、rs9956546（配列番号：7）、rs11912922（配列番号：8）およびrs6001954（配列番号：9）、ならびにそれらと連鎖不平衡のマーカーから選択され、試料中の少なくとも1つの対立遺伝子の存在の判定は、当該個体における乳癌に対する感受性を示す、方法に関する。1つの実施態様では、rs2005154対立遺伝子T、rs2184380対立遺伝子G、rs2224696対立遺伝子T、rs2242503対立遺伝子C、rs12291026対立遺伝子G、rs999737対立遺伝子C、rs9956546対立遺伝子A、rs11912922対立遺伝子Tおよびrs6001954対立遺伝子Gの存在の判定は、当該個体における乳癌の増加した感受性を示す。1つの実施態様では、遺伝子型の同定は、当該少なくとも1つの多型マーカーに隣接するヌクレオチドプライマー対を使用した、ポリメラーゼ連鎖反応（PCR）による、少なくとも1つの多型マーカーを含む核酸分節の増幅を含む。別の実施態様では、遺伝子型の同定は、対立遺伝子特異

的なプローブハイブリダイゼーション、対立遺伝子特異的なプライマー伸張、対立遺伝子特異的な増幅、核酸配列決定、5'-エキソヌクレアーゼ消化、分子ビーコンアッセイ、オリゴヌクレオチドライゲーションアッセイ、大きさ分析、一本鎖DNA高次構造解析およびマイクロアレイ技術から選択される工程を使用して行われる。1つの実施態様では、当該マイクロアレイ技術は、分子反転プローブアレイ技術またはビーズアレイ技術である。1つの実施態様では、当該工程は、対立遺伝子特異的なプローブハイブリダイゼーションを含む。別の実施態様では、当該工程は、マイクロアレイ技術を含む。1つの好ましい実施態様は、(1)当該核酸とのオリゴヌクレオチドプローブの特異的なハイブリダイゼーションのための条件下で、核酸のコピーを、検出オリゴヌクレオチドプローブおよびエンハンサーオリゴヌクレオチドプローブに接触させる工程であって、(a)当該検出オリゴヌクレオチドプローブは、長さが5~100ヌクレオチドであり、ヌクレオチド配列が配列番号:1~562のいずれか1つによって与えられる第1の核酸分節に特異的にハイブリダイズし、(b)当該検出オリゴヌクレオチドプローブは、その3'末端に検出可能な標識を、およびその5'末端に消光部分を含み、(c)当該エンハンサーオリゴヌクレオチドは、長さが5~100ヌクレオチドであり、オリゴヌクレオチドプローブに対して5'のヌクレオチド配列の第2の分節に相補的であり、そのため、当該エンハンサーオリゴヌクレオチドは、両方のオリゴヌクレオチドが当該核酸にハイブリダイズされると、当該検出オリゴヌクレオチドプローブに対して3'に位置し、かつ、(d)1つの塩基間隙が、当該第1の分節および当該第2の分節の間に存在し、そのため、当該オリゴヌクレオチドプローブおよび当該エンハンサーオリゴヌクレオチドプローブが当該核酸に両方ハイブリダイズされると、1つの塩基間隙が当該オリゴヌクレオチド間に存在する工程と、(2)当該検出プローブが当該核酸にハイブリダイズされると当該検出プローブの3'末端から当該検出可能な標識を切断しフリーな検出可能な標識を放出するエンドヌクレアーゼで当該核酸を処理する工程と、(3)フリーな検出可能な標識を測定する工程であって、当該フリーな検出可能な標識の存在は、当該検出プローブが、当該核酸の当該第1の分節に特異的にハイブリダイズすることを示し、かつ、当該多形性部位の配列が当該検出プローブの補完物であることを示す工程と、を含む。

【0050】

本発明のさらなる態様は、乳癌治療薬に対する反応可能性について個体を評価する方法であって、少なくとも1つの多型マーカーの少なくとも1つの対立遺伝子が、当該個体から入手した核酸試料または当該個体由来の遺伝子型データセット中に存在しているかどうかを判定する工程を含み、当該少なくとも1つの多型マーカーは、rs2005154(配列番号:1)、rs2184380(配列番号:2)、rs2224696(配列番号:3)、rs2242503(配列番号:4)、rs12291026(配列番号:5)、rs999737(配列番号:6)、rs9956546(配列番号:7)、rs11912922(配列番号:8)およびrs6001954(配列番号:9)、ならびにそれらと連鎖不平衡のマーカーから選択され、当該少なくとも1つのマーカーの少なくとも1つの対立遺伝子の存在は、当該治療薬に対する陽性反応の可能性を示す方法に関する。

【0051】

別の態様における本発明は、乳癌と診断された個体の予後診断を予測する方法に関し、当該方法は、少なくとも1つの多型マーカーの少なくとも1つの対立遺伝子が、当該個体から入手した核酸試料、または当該個体由来の遺伝子型のデータセット中に存在するかどうかの判定を含み、当該少なくとも1つの多型マーカーは、rs2005154(配列番号:1)、rs2184380(配列番号:2)、rs2224696(配列番号:3)、rs2242503(配列番号:4)、rs12291026(配列番号:5)、rs999737(配列番号:6)、rs9956546(配列番号:7)、rs11912922(配列番号:8)およびrs6001954(配列番号:9)、ならびにそれらと連鎖不平衡のマーカーから選択され、当該少なくとも1つの対立遺伝子の存在は、当該個体の乳癌の悪化した予後診断を示す。

【0052】

10

20

30

40

50

本発明のさらに別の態様は、乳癌の治療を受けている個体の治療の進行をモニターする方法に関し、当該方法は、少なくとも1つの多型マーカーの少なくとも1つの対立遺伝子が、当該個体から入手した核酸試料、または当該個体由来の遺伝子型のデータセットに存在しているかどうかについての判定を含み、当該少なくとも1つの多型マーカーは、rs 2005154 (配列番号: 1)、rs 2184380 (配列番号: 2)、rs 2224696 (配列番号: 3)、rs 2242503 (配列番号: 4)、rs 12291026 (配列番号: 5)、rs 999737 (配列番号: 6)、rs 9956546 (配列番号: 7)、rs 11912922 (配列番号: 8) および rs 6001954 (配列番号: 9)、ならびにそれらと連鎖不平衡のマーカーから選択され、当該少なくとも1つの対立遺伝子の存在は、当該個体の治療結果を示す。1つの実施態様では、当該治療は、手術による治療、放射線治療による治療、または薬剤投与による治療である。

10

【0053】

本発明はまた、ヒト個体における乳癌に対する感受性の診断および/または評価のための試薬の製造におけるオリゴヌクレオチドプローブの使用に関し、当該プローブは、配列番号: 1 ~ 562のいずれか1つに記載されるヌクレオチド配列を有する核酸分節とハイブリダイズし、当該プローブは長さが15 ~ 500ヌクレオチドである。ある実施態様では、当該プローブは、長さが約16 ~ 約100ヌクレオチドである。ある他の実施態様では、当該プローブは、長さが約20 ~ 約50ヌクレオチドである。ある他の実施態様では、当該プローブは、長さが約20 ~ 約30ヌクレオチドである。

【0054】

乳癌の表現型の様々な診断および分類は、本発明の範囲内である。その最も広い意味で、本発明はいずれの乳癌の表現型にも関する。乳癌は、ある実施態様では、下記を含む(しかしこれらに限定されない)乳癌のいずれの臨床的診断をも含む: 浸潤性乳管癌、浸潤性小葉癌、管状癌、または他の浸潤性癌もしくは混合型浸潤性癌、髄様癌、DCIS (非浸潤性乳管癌)、LCIS (非浸潤性小葉癌)、または他の非浸潤性癌; 湿潤性乳癌 (ステージ0、ステージ1、ステージ2 (ステージ2aおよびステージ2bを含む)、ステージ3 (ステージ3a、ステージ3bおよびステージ3cを含む) およびステージ4の乳癌を含む)。ある実施態様では、当該乳癌の表現型は、全ての乳癌、多発性原発性乳癌、および若年性乳癌から選択される。いくつかの実施態様では、本発明のマーカーは、乳癌の家族歴がある個体の乳癌のリスクに関連する。このような実施態様の1つでは、合計された家族歴 (summed family history, FHS) は、乳癌に関連する表現型である。別の実施態様では、本発明の変異体に関連する乳癌は、エストロゲン受容体 (ER) 陽性の乳癌および/またはプロゲステロン受容体 (PR) 陽性の乳癌である。1つの実施態様では、本発明の変異体に関連する乳癌は、エストロゲン受容体 (ER) 陽性である。別の実施態様では、本発明の変異体に関連する乳癌は、プロゲステロン受容体 (ER) 陽性である。このような実施態様の1つは、乳癌の増加したリスクまたは感受性に関連する、本願明細書に記載されたマーカーは、ER陽性の乳癌および/またはPR陽性の乳癌の増加したリスクまたは感受性を与える。従って、ある実施態様では、少なくとも1つのリスクのある本発明の変異体の存在は、個体におけるER陽性の乳癌またはPR陽性の乳癌の前兆となる。

20

30

40

【0055】

本発明の方法のいくつかの実施態様では、当該方法で判定された感受性は、増加した感受性である。このような実施態様の1つでは、当該増加した感受性は、少なくとも1.10の相対リスク (RR) によって特徴付けられる。別の実施態様では、当該増加した感受性は、少なくとも1.20の相対リスクによって特徴付けられる。別の実施態様では、当該増加した感受性は、少なくとも1.30の相対リスクによって特徴付けられる。別の実施態様では、当該増加した感受性は、少なくとも1.40の相対リスクによって特徴付けられる。さらに別の実施態様では、当該増加した感受性は、少なくとも1.50の相対リスクによって特徴付けられる。さらなる実施態様では、当該増加した感受性は、少なくとも1.70の相対リスクによって特徴付けられる。さらに別の実施態様では、当該増加し

50

た感受性は、少なくとも2.0の相対リスクによって特徴付けられる。他の実施態様は、少なくとも1.10、1.11、1.12、1.13、1.14、1.15、1.16、1.17、1.18、1.19、1.20、1.21、1.22、1.23、1.24、1.25、1.26、1.27、1.28、1.29、1.30、1.31、1.32、1.33、1.34、1.35の相対リスクによって特徴付けられる。これらの上記値のいずれの隙間をも埋める、リスクについての他の数値もまた可能であり、これらもまた本発明の範囲内である。

【0056】

本発明の方法のいくつかの実施態様では、当該方法で判定された感受性は、減少した感受性である。このような実施態様の1つでは、当該減少した感受性は、0.9未満の相対リスク(RR)によって特徴付けられる。別の実施態様では、当該減少した感受性は、0.8未満の相対リスク(RR)によって特徴付けられる。別の実施態様では、当該減少した感受性は、0.7未満の相対リスク(RR)によって特徴付けられる。さらに別の実施態様では、当該減少した感受性は、0.5未満の相対リスク(RR)によって特徴付けられる。他のカットオフ、例えば、0.89、0.88、0.87、0.86、0.85、0.84、0.83、0.82、0.81、0.80、0.79、0.78、0.77、0.76、0.75、0.74、0.73、0.72、0.71、0.70など未満の相対リスクも、本発明の範囲内である。

10

【0057】

本発明はまた、キットに関する。1つのこのような態様では、本発明は、ヒト個体において乳癌に対する感受性を評価するためのキットに関し、当該キットは、当該個体のゲノムにおいてrs2005154(配列番号:1)、rs2184380(配列番号:2)、rs2224696(配列番号:3)、rs2242503(配列番号:4)、rs12291026(配列番号:5)、rs999737(配列番号:6)、rs9956546(配列番号:7)、rs11912922(配列番号:8)、rs6001954(配列番号:9)およびそれらと連鎖不平衡のマーカから選択される少なくとも1つの多型マーカの少なくとも1つの対立遺伝子を選択的に検出するために必要な試薬を含み、当該少なくとも1つの対立遺伝子の存在は、乳癌に対する増加した感受性を示す。1つの実施態様では、当該キットはさらに、当該少なくとも1つの多型性と乳癌に対する感受性との間の相関データを含むデータの収集物を含む。当該相関データは、例えば相対リスク判定基準(RR)、オッズ比(OR)、または当業者に公知の他の簡便な尺度として、いずれの適した形式であり得る。1つの実施態様では、当該データの収集物は、コンピューター可読媒体上にある。別の態様では、本発明は、ヒト個体において乳癌に対する感受性を評価するためのキットに関し、当該キットは、当該個体のゲノムの中の少なくとも1つの多型マーカの少なくとも1つの対立遺伝子を選択的に検出するための試薬を含み、当該多型マーカは、rs2005154(配列番号:1)、rs2184380(配列番号:2)、rs2224696(配列番号:3)、rs2242503(配列番号:4)、rs12291026(配列番号:5)、rs999737(配列番号:6)、rs9956546(配列番号:7)、rs11912922(配列番号:8)およびrs6001954(配列番号:9)から選択され、当該少なくとも1つの対立遺伝子の存在は乳癌に対する感受性を示す。1つの実施態様では、当該少なくとも1つの多型マーカは、表4に記載されるマーカから選択される。

20

30

40

【0058】

キット試薬は、1つの実施態様では、当該少なくとも1つの多型マーカを含む個体のゲノムの断片とハイブリダイズする少なくとも1つの近接するオリゴヌクレオチドを含んでいてもよい。別の実施態様では、当該キットは、当該被験者から入手したゲノム分節の逆ストランドとハイブリダイズする少なくとも1つのオリゴヌクレオチド対を含み、各オリゴヌクレオチドプライマー対は、1つの多型性を含む個体のゲノム断片を選択的に増幅するように設計され、当該多型性は、表4で定められる多型性からなる群から選択され、当該断片は、大きさが少なくとも20塩基対である。1つの実施態様では、当該オリゴヌ

50

クレオチドは、当該個体のゲノムに完全に相補的である。別の実施態様では、当該キットは、さらに、当該分節を増幅するためのバッファーおよび酵素を含む。別の実施態様では、当該試薬は、さらに、当該断片を検出するための標識を含む。

【0059】

1つの好ましい実施態様では、当該キットは、長さが5～100ヌクレオチドの検出オリゴヌクレオチドプローブ、長さが5～100ヌクレオチドのエンハンサーオリゴヌクレオチドプローブ、およびエンドヌクレアーゼ酵素を含み、当該検出オリゴヌクレオチドプローブは、ヌクレオチド配列が配列番号：1～562のいずれか1つに記載される核酸の第1の分節に特異的にハイブリダイズし、当該検出オリゴヌクレオチドプローブは、その3'末端に検出可能な標識およびその5'末端に消光部分を含み、当該エンハンサーオリゴヌクレオチドは、長さが5～100ヌクレオチドであり、当該オリゴヌクレオチドプローブに対して5'のヌクレオチド配列の第2の分節に相補的であり、そのため、当該エンハンサーオリゴヌクレオチドは、両方のオリゴヌクレオチドが当該核酸にハイブリダイズされると、当該検出オリゴヌクレオチドプローブに対して3'に位置し、1つの塩基間隙が当該第1の分節と当該第2の分節との間に存在し、そのため、当該オリゴヌクレオチドプローブおよび当該エンハンサーオリゴヌクレオチドプローブが両方とも当該核酸にハイブリダイズされると、1つの塩基間隙が当該オリゴヌクレオチド間に存在し、当該検出プローブが当該核酸にハイブリダイズされる場合、当該核酸を当該エンドヌクレアーゼで処理すると、当該検出可能な標識を当該検出プローブの3'末端から切断して、フリーな検出可能な標識を放出する。

10

20

【0060】

本発明におけるキットはまた、以前に乳癌と診断された個体における少なくとも続発性原発腫瘍の発生リスクの評価方法、乳癌治療薬に対する反応可能性についての個体の評価方法、および乳癌と診断され、当該疾病について治療を受けた個体の治療の進行をモニターする方法を含む、本発明の他の方法に使用されてもよい。

【0061】

乳癌と関連すると本願明細書に記載されたマーカーは、全て、本願明細書に記載された方法、キット、使用、装置、手順を含む、本発明の様々な態様で使用され得る。一般的な意味では、本発明は、本願明細書で定められるLDブロックC09、LDブロックC10A、LDブロック10B、LDブロックC11A、LDブロックC11B、LDブロックC14、LDブロックC18、LDブロックC22A、およびLDブロックC22Bのうちのいずれか1つに関連するマーカーに関する。ある実施態様では、本発明は表1または表4に記載されるマーカー、およびそれらと連鎖不平衡のマーカーに関する。ある他の実施態様では、本発明は表4に記載されるマーカーに関する。ある他の実施態様では、本発明はマーカーrs2005154、rs2184380、rs2224696、rs2242503、rs12291026、rs999737、rs9956546、rs11912922およびrs6001954、およびそれらと連鎖不平衡のマーカーに関する。いくつかの他の好ましい実施態様では、本発明は、rs2005154（配列番号：1）、rs2184380（配列番号：2）、rs2224696（配列番号：3）、rs2242503（配列番号：4）、rs12291026（配列番号：5）、rs999737（配列番号：6）、rs9956546（配列番号：7）、rs11912922（配列番号：8）およびrs6001954（配列番号：9）のうちのいずれか1つに関する。

30

40

【0062】

ある実施態様では、乳癌の増加したリスクを与える少なくとも1つのマーカー対立遺伝子は、rs2005154対立遺伝子T、rs2184380対立遺伝子G、rs2224696対立遺伝子T、rs2242503対立遺伝子C、rs12291026対立遺伝子G、rs999737対立遺伝子C、rs9956546対立遺伝子A、rs11912922対立遺伝子Tおよびrs6001954対立遺伝子Gから選択される。これらの実施態様では、当該対立遺伝子（当該リスクのある対立遺伝子）の存在は、乳癌の増加

50

したリスクを示す。

【0063】

本発明のある実施態様では、連鎖不平衡は、連鎖不平衡の基準 r^2 および $|D'|$ を使用して測定され、これらは2つの遺伝的要素（例えば、多型マーカー）間の連鎖不平衡（LD）についての程度の定量的尺度を与える。特定のマーカーについてのこれらの尺度についてのある数値は、さらに本願明細書に記載されるように、当該マーカーが連鎖不平衡であることを示す。本発明の1つの実施態様では、マーカー間の連鎖不平衡（すなわち、当該マーカーが連鎖不平衡であることを示すLD値）は、 $r^2 > 0.1$ であるとして定められる。別の実施態様では、連鎖不平衡は、 $r^2 > 0.2$ であるとして定められる。他の実施態様は、連鎖不平衡の他の定義（ $r^2 > 0.25$ 、 $r^2 > 0.3$ 、 $r^2 > 0.35$ 、 $r^2 > 0.4$ 、 $r^2 > 0.45$ 、 $r^2 > 0.5$ 、 $r^2 > 0.55$ 、 $r^2 > 0.6$ 、 $r^2 > 0.65$ 、 $r^2 > 0.7$ 、 $r^2 > 0.75$ 、 $r^2 > 0.8$ 、 $r^2 > 0.85$ 、 $r^2 > 0.9$ 、 $r^2 > 0.95$ 、 $r^2 > 0.96$ 、 $r^2 > 0.97$ 、 $r^2 > 0.98$ 、または $r^2 > 0.99$ など）を含むことができる。連鎖不平衡はまた、ある実施態様では、 $|D'| > 0.2$ として、または $|D'| > 0.3$ 、 $|D'| > 0.4$ 、 $|D'| > 0.5$ 、 $|D'| > 0.6$ 、 $|D'| > 0.7$ 、 $|D'| > 0.8$ 、 $|D'| > 0.9$ 、 $|D'| > 0.95$ 、 $|D'| > 0.98$ 、もしくは $|D'| > 0.99$ としても定められ得る。ある実施態様では、連鎖不平衡は、 r^2 および $|D'|$ の2つの判断基準（ $r^2 > 0.2$ および $|D'| > 0.8$ など）を満たすように定められる。 r^2 および $|D'|$ の値の他の組み合わせもまた可能、かつ、本発明の範囲内であり、上記のこれらのパラメータの値を含むが、これらに限定されない。

10

20

【0064】

たとえ特徴の組み合わせが本願明細書中の同じ文章または段落に具体的に見出されないとしても、本願明細書に記載された特徴のすべての組み合わせが企図されているということを理解されたい。これは、特に、本願明細書に記載された本発明のすべての態様における、個体のまたはハプロタイプの分析のための、単独または組合せての本願明細書に開示されるすべてのマーカーの使用を含む。

【0065】

本発明の前述のおよび他の目的、特徴および利点は、本発明の好ましい実施態様についての下記のより詳細な説明で明らかとなるだろう。

30

【図面の簡単な説明】

【0066】

【図1】本願明細書に記載されたリスク変異体を利用するコンピューターで実施されるシステムを図示する略図を提供する。

【0067】

本発明は、乳癌に関連することが見出された多型の変異体およびハプロタイプを開示する。染色体9、10、11、14、18および22上の多型マーカーの特定の対立遺伝子は、乳癌の発生のリスクと関連することが見出された。このようなマーカーおよびハプロタイプは、診断目的、薬剤反応の予測方法および治療進行の予測方法に対して、本願明細書にさらに詳細に記載されるように、有用である。本発明のさらなる適用は、本発明の多型マーカーを利用した、手術または放射線による乳癌治療に対する反応を評価するための方法、および本発明の方法の使用のためのキットを含む。

40

【0068】

（定義）

別段の表示がない限り、核酸配列は、左から右に、5'から3'方向に書かれる。本願明細書で挙げられる数値的範囲は、範囲を定める数字を含み、定義された範囲内の各整数またはいずれの非整数有理数を含む。別段の定義がない限り、本願明細書で使用される全ての技術用語および科学用語は、本発明が関与する技術分野の当業者によって通常理解されるものと同義である。

【0069】

50

下記の用語は、本願の構成では、下記に示される意味を有するものとする。

【0070】

「多型マーカー」は、「マーカー」と呼ばれることもあり、本願明細書に記載される場合、ゲノム多形性部位を意味する。各多型マーカーは、多形性部位の特定の対立遺伝子に特徴のある少なくとも2つの配列多様性を有する。従って、多型マーカーの遺伝的関連は、その特定の多型マーカーの少なくとも1つの特定の対立遺伝子に関連があることを意味する。当該マーカーは、一塩基多型(SNP)、ミニサテライトまたはマイクロサテライト、転座およびコピー数の変異(挿入、欠失、重複)を含む、ゲノム中に見出されるいずれの変異型のいずれの対立遺伝子をも含み得る。多型マーカーは、集団における測定可能な頻度のいずれであってもよい。疾病遺伝子のマッピングのためには、集団頻度が5~10%よりも高い多型マーカーが一般的に最も有用である。しかし、多型マーカーもまた、特定のコピー数変異(CNV)について、より低い集団頻度(1~5%頻度、またはさらに低い頻度など)を有するかも知れない。当該用語は、本願の構成においては、いずれの集団頻度の多型マーカーを含むように解釈されるものとする。

10

【0071】

「対立遺伝子」は、染色体の所定の遺伝子座(位置)のヌクレオチド配列を意味する。多型マーカー対立遺伝子は、従って、染色体上のマーカーの構成(すなわち、配列)を意味する。個体のゲノムDNAは、各染色体上のマーカーの各コピーを代表する、いずれの所定の多型マーカーの2つの対立遺伝子を含む。本願明細書で使用されるヌクレオチドの配列コードは、A=1、C=2、G=3、T=4である。マイクロサテライト対立遺伝子のために、CEPH試料(ヒト多型研究センター、ゲノム貯蔵所、CEPH試料1347-02)を参照として使用し、試料中の各マイクロサテライトのより短い対立遺伝子を0と設定し、他の試料中の全ての他の対立遺伝子に、当該参照に関連して番号を付けた。従って、例えば、対立遺伝子1は、CEPH試料中のより短い対立遺伝子より1bp長く、対立遺伝子2は、CEPH試料中のより短い対立遺伝子より2bp長く、対立遺伝子3はCEPH試料中の下位の対立遺伝子より3bp長いなどであり、対立遺伝子-1は、CEPH試料中のより短い対立遺伝子よりも1bp短く、対立遺伝子-2は、CEPH試料中のより短い対立遺伝子より2bp短いなどである。

20

【0072】

「一塩基多型」または「SNP」は、ゲノムの特異的な部位の1つのヌクレオチドが種のメンバー間または個体の対の染色体間で異なる場合に生じる、DNA配列の多様性である。大部分のSNP多型は、2つの対立遺伝子を有する。当該例では、各個体は、多型の1つの対立遺伝子のホモ接合(すなわち、個体の染色体のコピーの両方が、SNP部位の同一のヌクレオチドを有する)であるか、または、個体はヘテロ接合(すなわち、異なるヌクレオチドを含む個体の2つの姉妹染色体)である。本願明細書で報告されたSNPの命名は、国立バイオテクノロジー情報センター(NCBI)によって各特有のSNPに割り当てられた公式の参照SNP(rs)ID同定タグを意味する。

30

【0073】

本願明細書記載の配列コヌクレオチド(conucleotide)の多義性は、IUPAC-IUBによって提唱される通りである。これらのコードは、EMBL、GenBank、およびPIRデータベースによって使用されるコードと両立する。

40

【0074】

【表 1】

IUB コード	意味
A	アデノシン
C	シチジン
G	グアニン
T	チミジン
R	GまたはA
Y	TまたはC
K	GまたはT
M	AまたはC
S	GまたはC
W	AまたはT
B	C GまたはT
D	A GまたはT
H	A CまたはT
V	A CまたはG
N	A C GまたはT (いずれかの塩基)

10

【0075】

本願明細書中に与えられる配列表は、表 1 ~ 4 の中で本願明細書に示される多型マーカーについての隣接配列を提供する。多形性部位は、上記の配列コヌクレオチドの多義コードを使用して当該配列の中に示されている。

20

【0076】

1 より大きい配列が集団（自然の集団または合成の集団（例えば、合成分子のライブラリー）のいずれか）において可能性のあるヌクレオチドの位置は、本願明細書では「多形性部位」と呼ばれる。

【0077】

「変異」は、本願明細書に記載される場合、参照 DNA と異なる DNA の分節を意味する。「マーカー」または「多型マーカー」は、本願明細書に定義されるように、変異である。参照と異なる対立遺伝子は、「変異」対立遺伝子と呼ばれる。

30

【0078】

「マイクロサテライト」は、特定の部位の長さが 2 ~ 8 ヌクレオチドの塩基の複数の小さい反復（CA 反復など）を有する多型マーカーであり、反復長の数は一般的な集団において変化する。

【0079】

「インデル」は、典型的にはわずか数ヌクレオチド長の小さな挿入または欠損を含む、多型性の一般的な形態である。

【0080】

「ハプロタイプ」は、本願明細書に記載される場合、分節に沿って並んだ対立遺伝子の特定の組み合わせによって特徴付けられる一本鎖 DNA 内のゲノム DNA の分節を意味する。ヒトなどの 2 倍体の生物では、ハプロタイプは各多型マーカーまたは遺伝子座の対立遺伝子の 1 つのメンバーを含む。ある実施態様では、当該ハプロタイプは、2 つ以上の対立遺伝子、3 つ以上の対立遺伝子、4 つ以上の対立遺伝子、または 5 つ以上の対立遺伝子を含み得る。

40

【0081】

用語「感受性」は、本願明細書に記載されるように、増加した感受性および減少した感受性の両方を包含する。従って、本発明の特定の多型マーカーおよび/またはハプロタイプは、1 よりも大きい相対リスク（RR）によって特徴付けられるように、または 1 よりも大きいオッズ比（OR）であるように、乳癌の増加した感受性（すなわち、増加したリスク）の特性を示していてもよい。あるいは、本発明のマーカーおよび/またはハプロタ

50

イブは、1未満の相対リスクまたは1未満のオッズ比によって特徴付けられるように、乳癌の減少した感受性（すなわち、減少したリスク）の特徴を表わす。ハプロタイプは、本願明細書では、マーカー名およびそのハプロタイプのマーカーの対立遺伝子に関連して記載され、例えば、「A r s 9 9 5 6 5 4 6」は、そのハプロタイプ中のマーカー r s 9 9 5 6 5 4 の A 対立遺伝子を意味し、この命名は、「r s 9 9 5 6 5 4 6 対立遺伝子 A」および「A - r s 9 9 5 6 5 4 6」に等しい。さらに、ハプロタイプにおける対立形質のコードは、個体マーカーについてのものと同じようであり、すなわち、1 = A、2 = C、3 = G および 4 = T である。

【0082】

用語「感受性」は、本願明細書に記載される場合、ある状態（例えば、ある形質、表現型または疾病（例えば、乳癌））の発生または、平均的な個体よりも特定の状態に抵抗できないことに対する個体の易発性を意味する。当該用語は、増加した感受性および減少した感受性の両方を包含する。従って、本発明の多型マーカーにおける特定の対立遺伝子および/またはハプロタイプは、本願明細書に記載されるように、1より大きいその特定の対立遺伝子またはハプロタイプの相対リスク（RR）、またはオッズ比（OR）によって特徴付けられるように、乳癌の増加した感受性（すなわち、増加したリスク）に特徴があってもよい。あるいは、本発明のマーカーおよび/またはハプロタイプは、1未満の相対リスクによって特徴付けられるように、乳癌の減少した感受性（すなわち、減少したリスク）に特徴がある。

10

【0083】

用語「および（ならびに、かつ）/または（もしくは）」は、本願の構成では、当該用語によって結合された項目のいずれかまたは両方が関連することを示すことが理解されるだろう。換言すれば、本願明細書におけるこの用語は、「いずれかまたは両方」を意味することが捉えられるだろう。

20

【0084】

用語「検査表」は、本願明細書に記載される場合、データの形態を別の形態と関連付ける表、または、1もしくは複数の形態のデータを、当該データが関連する予測される成果（表現型または形質など）と関連付ける表である。例えば、検査表は、少なくとも1つの多型マーカーの対立遺伝子のデータと、特定の対立遺伝子のデータを含む個体が示しやすいまたは特定の対立遺伝子のデータを含まない個体よりも示しやすい特定の形質もしくは表現型（特定の疾病診断など）との間の相関を含み得る。検査表は、多次元的であり得、すなわち、単一のマーカーの複数の対立遺伝子についての情報を同時に含み得、または、複数のマーカーについての情報を含み得、他の因子（疾病診断の特徴、人種の情報、バイオマーカー、生化学的測定、治療方法または薬剤など）も含んでいてもよい。

30

【0085】

「コンピューター可読媒体」は、市販または特注のインターフェースを使用した、コンピューターによって接続され得る情報記憶媒体である。典型的なコンピューター可読媒体は、メモリ（例えば、RAM、ROM、フラッシュメモリなど）、光学式記憶媒体（例えば、CD-ROM）、磁気記憶媒体（例えば、コンピューターのハードドライブ、フロッピー（登録商標）ディスクなど）、パンチカードまたは他の市販の媒体を含む。情報は、興味のあるシステムと媒体との間、コンピューター間、またはコンピューターと、記憶した情報を保存または接続するコンピューター可読媒体との間で転送されてもよい。このような転送は、電気的であり得、または他の利用可能な方法（IRリンク、無線接続など）によってもよい。

40

【0086】

「核酸試料」は、核酸（DNAまたはRNA）を含む個体から入手した試料である。ある実施態様、すなわち、特異的な多型マーカーの検出および/またはハプロタイプの検出では、当該核酸試料はゲノムDNAを含む。このような核酸試料は、ゲノムDNAを含むいずれの供給源（血液試料、羊水試料、脳脊髄液試料、または皮膚、筋肉、頬側粘膜もしくは結膜粘膜、胎盤、消化管もしくは他の器官から得た組織試料を含む）からも得られ得

50

る。

【0087】

用語「乳癌の治療薬」は、乳癌に関連する症候の寛解または予防に使用され得る薬剤を意味する。

【0088】

用語「乳癌に関連した核酸」は、本願明細書に記載される場合、乳癌に関連することが見出された核酸を意味する。これは、本願明細書記載のマーカーおよびハプロタイプ、ならびにそれらと強い連鎖不平衡(LD)のマーカーおよびハプロタイプを含むが、これらに限定されない。

【0089】

用語「乳癌」は、本願明細書に記載されるように、乳癌のいずれかの臨床的診断を意味し、乳癌の特定の亜表現型のいずれかおよび全てを含む。例えば、乳癌は、エストロゲン受容体(ER)陽性の乳癌またはエストロゲン受容体陰性の乳癌として分類されることがある。乳癌はまた、プロゲステロン受容体(PR)陽性または陰性として分類されることもある。乳癌は、さらに、浸潤性乳管癌として、浸潤性小葉癌として、管状癌として、または他の浸潤性乳癌または混合型浸潤性乳癌として診断されることがある。乳癌はまた、髄様DCIS(非浸潤性乳管癌)またはLCIS(非浸潤性小葉癌)、または他の非浸潤性乳癌として分類され得る。浸潤性乳癌はまた、ステージ0、ステージ1、ステージ2(ステージ2aおよびステージ2bを含む)、ステージ3(ステージ3a、ステージ3bおよびステージ3cを含む)またはステージ4乳癌として定められ得る。本願の構成においては、「乳癌」は、乳癌のこれらの亜表現型のいずれをも含むことができ、また、いずれの他の臨床的に適用できる乳癌の亜表現型をも含む。

【0090】

用語「全ての乳癌」または「全てのBC」は、乳癌と診断された全ての個体を意味する。

【0091】

用語「媒体素因の」乳癌または「MedPreの」乳癌は、乳癌の亜表現型を意味する。当該表現型の定義は、発端者が少なくとも1つの下記の判断基準を満たすことを必要とする。

当該発端者は、3減数分裂事象(meiotic events)(3M)の遺伝的距離内の、3人以上の罹患した近親者を含む乳癌の症例の集団の一員である。

当該発端者は、3M以内に関連する、罹患したペアの一員であり、ペアの一人は50歳以下のときに診断された。

当該発端者は、3M以内に関連する、罹患したペアの一員であり、ペアの一人はいずれかのタイプの続発性原発腫瘍であると診断された。

当該発端者は、いずれかのタイプの続発性原発腫瘍であると診断された。

【0092】

用語「多発性原発性乳癌」、または「MPBC」は、本願明細書に記載される場合、最初の乳癌の診断に加えて少なくとも1つの原発性腫瘍が診断され、2つの腫瘍が臨床的に、かつ組織学的検査によって、別個の原発性腫瘍であることが確かめられ、最初の乳癌と同時に、またはその後生じており、対側乳房または同側乳房に生じている症例を意味する。

【0093】

用語「家族歴スコア」または「FHS」は、本願明細書に記載される場合、疾病の発端者の乳癌に罹患した近親者の数に基づいて定義される。各発端者について、スコア1は各罹患した第一度近親者に割り当てられ、0.5は各罹患した第二度近親者に割り当てられ、0.25は各第三度近親者に割り当てられる。従って、全ての罹患した近親者にわたって得られた累計は、合計された家族歴スコアまたはFHSを表す。

【0094】

用語「エストロゲン受容体陽性の乳癌」、または「ER陽性の乳癌」は、本願明細書で

10

20

30

40

50

定義される場合、エストロゲン受容体について陽性と判定された腫瘍を意味する。本願の構成においては、ERレベルが10 fmol/mg以上、および/または免疫組織化学的な知見で陽性の核が10%以上であるときは、ER陽性であると考えられる。ER陽性であるという判断基準を満たさない乳癌は、本願明細書では、「ER陰性」または「エストロゲン受容体陰性」と定義される。

【0095】

用語「プロゲステロン受容体陽性の乳癌」、または「PR陽性の乳癌」は、本願明細書に記載されるように、プロゲステロン受容体が陽性と判定された腫瘍を意味する。本願の構成においては、10 fmol/mg以上のPRレベルおよび/または10%以上の陽性の核についての免疫組織化学的な知見は、PR陽性であると考えられる。PR陽性であるという判断基準を満たさない乳癌は、本願明細書では、「PR陰性」または「プロゲステロン受容体陰性」として定められる。

10

【0096】

本願明細書に記載される用語「PAX5」または「PAX5遺伝子」は、ヒト染色体9p13上のBSAPとしても知られるPAIRED BOX GENE 5遺伝子を指す。

【0097】

本願明細書に記載される用語「TUB」または「rd5遺伝子」は、ヒト染色体11p15.5上のTubby homolog (マウス) 遺伝子を指す。

【0098】

本願明細書に記載される用語「SERPINH1」は、ヒト染色体11上のセルピンペプチダーゼ阻害剤、clade H (ヒートショックタンパク質47)、member 1、(コラーゲン結合タンパク質1) 遺伝子を指す。

20

【0099】

また、本願明細書に記載される、「REC2; R51H2; hREC2; RAD51B またはMGC34245」とも呼ばれる用語「RAD51L1 X遺伝子」は、ヒト染色体14q23-24上に位置するタンパク質コード遺伝子を指す。

【0100】

さらに、本願明細書に記載される、「FHOS2遺伝子」または「Formactin 2」としても知られる用語「FHOD3遺伝子」は、ヒト染色体18q12上に位置するタンパク質コード遺伝子であるformin homolog 2 domain containing 3遺伝子を表す。

30

【0101】

最後に、本願明細書に記載される、「KIAA1093」としても知られる用語「TNRC6B」または「TNRC6B遺伝子」は、ヒト染色体22q13上に位置するタンパク質コード遺伝子である、trinucleotide repeat containing 6Bを指す。

【0102】

本発明の乳癌と診断された個体の集団の関連分析を通じて、染色体9、10、11、14、18および22上のある多型マーカーのある対立遺伝子は、乳癌と関連することが発見された。癌と関連する変異体についてのゲノム全体での分析は、乳癌と、9つの染色体の領域、すなわち

40

染色体9、位置36,806,001~36,859,001の間(LDブロックC09) ;

染色体10、位置8,643,001~8,817,001の間(LDブロックC10A) ;

染色体10、位置9,077,001~9,264,001の間(LDブロックC10B) ;

染色体11、位置8,053,258~8,191,268の間(LDブロックC11A) ;

50

染色体 1 1、位置 7 4, 8 8 6, 3 4 1 ~ 7 4, 9 7 1, 3 4 1 の間 (LD ブロック C 1 1 B) ;

染色体 1 4、位置 6 8, 0 3 5, 7 1 2 ~ 6 8, 1 3 0, 7 1 2 の間 (LD ブロック C 1 4) ;

染色体 1 8、位置 3 2, 1 1 0, 0 1 2 ~ 3 2, 1 4 5, 0 1 2 の間 (LD ブロック C 1 8) ;

染色体 2 2、位置 3 8, 7 0 4, 9 0 7 ~ 3 8, 8 5 9, 9 0 7 の間 (LD ブロック C 2 2 A) ;

染色体 2 2、位置 3 8, 8 5 9, 9 0 7 ~ 3 9, 4 1 1, 9 0 7 の間 (LD ブロック C 2 2 B) ;

(すべての位置は N C B I ビルド 3 6 の座標に対応する) との関連を明らかにした。

【0103】

これらの領域内の特定のマーカーは、これらの部位における乳癌の増加したリスクに関連することが見出された。

【0104】

イルミナ ヒト H a p 3 0 0 マイクロアレイ技術を使用した、約 1 8 4 0 のアイスランド人の乳癌患者および 3 0, 3 5 0 の対照の平均の遺伝子型の同定を通じて、いくつかの染色体部位上の多数のマーカーが、乳癌との関連を示すことが見出された (表 1)。特に、9 個の SNP ; r s 2 0 0 5 1 5 4、r s 2 1 8 4 3 8 0、r s 2 2 2 4 6 9 6、r s 2 2 4 2 5 0 3、r s 1 2 2 9 1 0 2 6、r s 9 9 9 7 3 7、r s 9 9 5 6 5 4 6、r s 1 1 9 1 2 9 2 2 および r s 6 0 0 1 9 5 4 は、乳癌の増加したリスクに関連することが見出された。

【0105】

アイスランド、オランダ、スペインおよびスウェーデンからのさらなるコホートにおける追跡調査解析は、当該 9 つのマーカーの関連シグナルが実際有意であるということを示した (表 3)。これらのマーカー、およびこれらのマーカーのうちのいずれか 1 つと連鎖不平衡の代用マーカーは、従って、個体における乳癌のリスクを予測するために有用である。代表的な代用マーカーは、本願明細書中の表 4 に与えられる。

【0106】

従ってこれらのマーカーは、上記のマーカーのうちのいずれか 1 つと LD であるため乳癌を予測するマーカーを含むと特に予想される 9 つの染色体の領域を同定する。これらの領域は、本願明細書中では、LD ブロック C 0 9、LD ブロック C 1 0 A、LD ブロック C 1 0 B、LD ブロック C 1 1 A、LD ブロック C 1 1 B、LD ブロック C 1 4、LD ブロック C 1 8、LD ブロック C 2 2 A、および LD ブロック C 2 2 B と呼ばれる。

【0107】

当業者なら、当該 9 つのアンカーマーカー r s 2 0 0 5 1 5 4、r s 2 1 8 4 3 8 0、r s 2 2 2 4 6 9 6、r s 2 2 4 2 5 0 3、r s 1 2 2 9 1 0 2 6、r s 9 9 9 7 3 7、r s 9 9 5 6 5 4 6、r s 1 1 9 1 2 9 2 2 および r s 6 0 0 1 9 5 4 のうちのいずれかと LD のマーカーは、本願明細書で定義される LD ブロックの外側に位置し得るということはわかるであろう。これは、LD が、通常定められる LD ブロックの見かけ上の物理的な境界 (高組み換え率の領域によって通常定められる境界) を超えてひろがり得るという事実の結果である。これらの 9 つのマーカーと LD のこのような代用マーカーもまた、本発明のために有用であると特に企図され、従って本発明の範囲内にある。

【0108】

(マーカーおよびハプロタイプの評価)

集団内のゲノム配列は、個体を比較すると同一ではない。ある程度、ゲノムは、ゲノム中の多くの部位で、個体間の配列変異性を呈する。このような配列の変異は、通常、多型性と呼ばれ、各ゲノム内に多くのこのような部位がある。例えば、ヒトゲノムは、平均 5 0 0 塩基対ごとに生じる配列多様性を呈する。最も一般的な配列変異は、ゲノム中の単一の塩基位置の塩基変異からなり、このような配列変異、または多型性は、通常、一塩基多

10

20

30

40

50

型（「SNP」）と呼ばれる。これらのSNPは、単一の突然変異事象によって生じていると考えられ、従って、通常、各SNP部位において可能性を有する、2つの可能性のある対立遺伝子がある（本来の対立遺伝子および変異した（一方の）対立遺伝子）。自然の遺伝的浮動、および、おそらくまた、選択圧によって、最初の変異は、いずれの所定の集団におけるその対立遺伝子の特定の頻度によって特徴付けられる多型性をもたらす。配列変異の多くの他の型は、ヒトゲノムにおいて見出され、ミニサテライトおよびマイクロサテライト、ならびに挿入、欠損、逆位（コピー数多型（CNV）とも呼ばれる）を含む。多型のマイクロサテライトは、複数の小塩基反復（CA反復、相補鎖上のTGなど）を、反復長数が一般的な集団において異なる特定の部位に有する。一般論では、多形性部位に関する配列の各バージョンは、多形性部位の特定の対立遺伝子を表す。全ての配列変異は、問題になっている配列変異を特徴付ける特定の多形性部位に生じる、多型性と呼ばれ得る。各ヒト個体は、各多形性部位に2つの対立遺伝子 - 1つの母系対立遺伝子および1つの父系対立遺伝子 - を有するが、一般に、多型性は、当該集団内のあらゆる数の特定の対立遺伝子を含み得る。従って、本発明の1つの実施態様では、当該多型性は、集団における2つ以上の対立遺伝子の存在によって特徴付けられる。別の実施態様では、当該多型性は、3つ以上の対立遺伝子の存在によって特徴付けられる。他の実施態様では、当該多型性は、4つ以上の対立遺伝子、5つ以上の対立遺伝子、6つ以上の対立遺伝子、7つ以上の対立遺伝子、9つ以上の対立遺伝子、または10以上の対立遺伝子によって特徴付けられる。全てのこのような多型性は、本発明の方法およびキットに使用され得、従って、本発明の範囲内である。

10

20

【0109】

その存在量によって、SNPは、ヒトゲノムの配列変異性の大部分について説明する。600万超のヒトのSNPが、現在までに確認されている（http://www.ncbi.nlm.nih.gov/projects/SNP/snp_summary.cgi）。しかし、CNVは、ますます注目されている。これらの大規模な多型性（典型的には、1kb以上）は、構築されたヒトゲノムの実質的な割合に影響する多型の変異性を説明する。公知のCNVはヒトゲノム配列の15%超にわたる（covery）（Estivill, X Armengol; L., PLoS Genetics 3:1787-99 (2007)）。<http://projects.tcag.ca/variation/>）。しかし、これらの多型性のほとんどは、非常にまれであり、平均して、各個体のゲノム配列の一部分のみに影響する。CNVは、遺伝子量の攪乱による遺伝子発現、表現型の変異性および適応に影響することが知られ、また、疾病（微小欠失疾患および微小重複疾患）の原因となることも知られ、一般的な複合疾患（HIV-1感染および糸球体腎炎を含む）のリスクを与える（Redon, R.ら、Nature 23:444-454 (2006)）。従って、前述のCNVまたは未知のCNVのいずれかが、当該疾病と関連のある本願明細書に記載されたマーカーと連鎖不平衡の、原因となる変異体を表わす可能性がある。CNVの検出方法は、比較ゲノムハイブリダイゼーション（CGH）および遺伝子型の同定（Carter (Nature Genetics 39:S16~S21 (2007)）によって記載されたように、遺伝子型の同定アレイの使用を含む）を含む。ゲノムの変異体のデータベース（<http://projects.tcag.ca/variation/>）は、記載されたCNVの位置、型および大きさについての最新の情報を含む。当該データベースは、現在、21,000超のCNVのデータを含む。

30

40

【0110】

いくつかの例では、参照対立遺伝子を選択することなく、多形性部位の異なる対立遺伝子が参照される。あるいは、参照配列は、特定の多形性部位を意味し得る。参照対立遺伝子は、「野生型」の対立遺伝子と呼ばれることもあり、それは、第1に配列決定された対立遺伝子、または「非罹患の」個体（例えば、形質または疾病の表現型を示さない個体）からの対立遺伝子のいずれかとして通常は選択される。

【0111】

50

本願明細書で述べられたSNPマーカの対立遺伝子は、多形性部位に生じる場合、塩基A、C、GまたはTを意味する。本願明細書で使用されるSNPについての対立遺伝子コードは、下記の通りである；1 = A、2 = C、3 = G、4 = T。ヒトのDNAは二本鎖であるので、当業者は、逆のDNA鎖のアッセイまたは解読によって、相補的な対立遺伝子が各症例において判定され得ることを理解するだろう。従って、A/G多型性によって特徴付けられる多形性部位（多型マーカ）について、当該マーカを検出するために使用される方法論は、可能性のある2つの塩基（すなわちAおよびG）のうち1つまたは両方の存在を特異的に検出するように設計されてもよい。あるいは、鋳型DNA上の逆ストランドを検出するように設計されたアッセイを設計することで、相補的塩基TおよびCの存在が判定され得る。（例えば、相対リスクに関して）量的に、同一の結果が、いずれかのDNA鎖（+鎖または-鎖）の判定から得られうる。

10

【0112】

典型的には、参照配列は、特定の配列を意味する。当該参照と異なる対立遺伝子は、「変異」対立遺伝子と呼ばれることもある。変異配列は、本願明細書で使用されるように、参照配列と異なるが、その他の点では実質的に類似である配列を意味する。本願明細書記載の多型の遺伝子マーカの対立遺伝子は変異体である。変異体は、ポリペプチドに影響する変化を含み得る。配列の相違は、参照ヌクレオチド配列と比較した場合、以下を含み得る；フレームシフトをもたらす、1つのヌクレオチドの挿入もしくは欠損、または1より多いヌクレオチドの挿入もしくは欠損；コードされるアミノ酸中の変化をもたらす、少なくとも1つのヌクレオチドの変化；中途終止コドンの生成をもたらす、少なくとも1つのヌクレオチドの変化；ヌクレオチドによってコードされた1つまたは複数のアミノ酸の欠損をもたらす、いくつかのヌクレオチドの欠損；読み枠のコード配列の中断をもたらす、不等組み換えまたは遺伝子変換などによる1つまたはいくつかのヌクレオチドの挿入；全てまたは一部の配列の重複；転座；またはヌクレオチド配列の再編成。このような配列変化は、核酸によってコードされたポリペプチドを変化させ得る。例えば、核酸配列の変化がフレームシフトを引き起こすと、当該フレームシフトはコードされたアミノ酸の変化をもたらし得、および/または、切断されたポリペプチドの生成をもたらす、中途終止コドンの生成をもたらし得る。あるいは、多型性は、1つまたは複数のヌクレオチドにおける同義の変化（すなわち、アミノ酸配列の変化をもたらさない変化）であり得る。このような多型性は、例えば、スプライス部位を変化させ得、mRNAの安定性または輸送に影響し得、またはそうでなければコードされたポリペプチドの転写または翻訳に影響し得る。このような多型性はまた、構造の変化（増幅または欠損など）が体細胞レベルで生じる可能性を増加させるようにDNAを変化させ得る。参照ヌクレオチド配列にコードされたポリペプチドは、特定の参照アミノ酸配列を有する「参照」ポリペプチドであり、変異対立遺伝子によってコードされたポリペプチドは、変異アミノ酸配列を有する「変異」ポリペプチドと呼ばれる。

20

30

【0113】

ハプロタイプは、DNAの一本鎖の分節に沿って並んだ対立遺伝子の特定の組み合わせによって特徴付けられる、そのDNAの一本鎖の分節を意味する。2倍体の生物（ヒトなど）では、ハプロタイプは、各多型マーカまたは遺伝子座の対の対立遺伝子のうちの1のメンバーを含む。ある実施態様では、当該ハプロタイプは、2つ以上の対立遺伝子、3つ以上の対立遺伝子、4つ以上の対立遺伝子、または5つ以上の対立遺伝子（各対立遺伝子は分節に沿った特定の多型マーカに対応する）を含み得る。ハプロタイプは、多形性部位において特定の対立遺伝子を有する、様々な多型マーカの組み合わせ（例えば、SNPとマイクロサテライト）を含み得る。従って、ハプロタイプは、様々な遺伝子マーカの対立遺伝子の組み合わせを含む。

40

【0114】

特定の多型マーカおよび/またはハプロタイプの検出は、多形性部位の配列の検出のための、当該技術分野で公知の方法によって達成され得る。例えば、SNPおよび/またはマイクロサテライトマーカの存在についての遺伝子型の同定の標準的な技術（PCR

50

、LCR、ネステッドPCRおよび核酸増幅のための他の技術を用いた、蛍光に基づいた技術（Chen X.ら、Genome Res. 9(5):492-98(1999)）など）が使用され得る。SNP遺伝子型の同定について利用可能な特定の市販の方法論は、TaqMan遺伝子型の同定アッセイおよびSNPlexプラットフォーム（アプライドバイオシステムズ（Applied Biosystems））、ゲル電気泳動（アプライドバイオシステムズ）、質量分析（例えば、MassARRAYシステム、シーケノム（Sequenom））、ミニ配列分析法、リアルタイムPCR、Bio-Plexシステム（バイオラッド（BioRad））、CEQおよびSNPstreamシステム（ベックマン（Beckman））、アレイハイブリダイゼーション技術（例えば、Affymetrix遺伝子チップ、パーレジェン（Perlegen））、ビーズアレイ技術（例えば、イルミナ GoldenGateおよびInfiniumアッセイ）、アレイタグ技術（例えば、Parallelle）、およびエンドヌクレアーゼに基づいた蛍光ハイブリダイゼーション技術（インベーター法、サードウェーブ（Third Wave））を含むが、これらに限定されない。利用できるアレイプラットフォームのいくつか（Affymetrix（Affymetrix）SNPアレイ 6.0およびイルミナ CNV370-デュオ（Duo）および1M ビーズチップス（BeadChips）を含む）は、あるCNVをタグするSNPを含む。これにより、これらのプラットフォームに含まれた代用SNPによって、CNVの検出が可能となる。当業者に利用可能なこれらのまたは他の方法の使用によって、多型マーカー（マイクロサテライト、SNPまたは他のタイプの多型マーカーを含む）の1つまたは複数の対立遺伝子を同定し得る。

【0115】

ある実施態様では、多型マーカーは、配列決定技術によって検出される。個体についての配列情報を入手することで、当該配列の構成の中の特定のヌクレオチドが特定される。SNPについては、その特定のSNPにおける対立遺伝子を特定するためには、1つの特有の配列部位についての配列情報で十分である。複数のヌクレオチドを含むマーカーについては、当該多形性部位を含む個体のヌクレオチドについての配列情報は、特定の部位について当該個体の対立遺伝子を特定する。この配列情報は、当該個体由来の試料から入手され得る。ある実施態様では、当該試料は核酸試料である。ある他の実施態様では、当該試料はタンパク質試料である。

【0116】

核酸配列を得るための種々の方法が当業者に公知であり、すべてのこのような方法が本発明を実施するために有用である。Sanger配列決定は、核酸配列情報を生成するための周知の方法である。大量の配列データを入力するための最近の方法が開発され、このような方法もまた、配列情報を入手するために有用であると企図される。これらとしては、パイロシーケンス技術（Ronaghi, M.ら Anal Biochem 267:65-71(1999); Ronaghiら、Biotechniques 25:876-873(1998)、例えば 454 パイロシーケンス（Nyren, P.ら、Anal Biochem 208:171-175(1993)）、Illumina/Solexa配列決定技術（<http://www.illumina.com>; Strausberg, RLら、Drug Disc Today 13:569-577(2008)も参照）、およびSupported Oligonucleotide Ligation and Detection Platform (SOLiD)技術（アプライドバイオシステムズ、<http://www.appliedbiosystems.com>）; Strausberg, RLら、Drug Disc Today 13:569-577(2008)が挙げられる。

【0117】

遺伝子型が同定された個体の遺伝子型が同定されていない近親者についての遺伝子型を帰属させるまたは予測することが可能である。遺伝子型が同定されていない症例すべてについて、その4つの可能性のある段階的な遺伝子型が与えられれば、その近親者の遺伝子型の可能性を算出することが可能である。実際面では、その症例の親、子供、兄弟姉妹、

片親違いの兄弟姉妹 (half-siblings) (およびその片親違いの兄弟姉妹の親)、祖父母、孫 (およびその孫の親) ならびに配偶者の遺伝子型だけを含めることが好ましくあり得る。各症例の周りで作り上げられた小さい下位系統 (sub-pedigree) 中の個体は、その系統に含まれないいずれの経路を介しても関連しないと仮定されるであろう。当該症例に伝えられていない対立遺伝子は、同じ頻度 - 集団対立遺伝子頻度 - を有するという仮定される。対立遺伝子 A および G を有する SNP マーカーを考えよう。この症例の近親者の遺伝子型の確率は、

【数 1】

$$\Pr(\text{近親者の遺伝子型}; \theta) = \sum_{h \in \{AA, AG, GA, GG\}} \Pr(h; \theta) \Pr(\text{近親者の遺伝子型} | h),$$

10

によって計算され得る。式中、 θ は当該症例における A 対立遺伝子の頻度を表す。近親者の各組の遺伝子型が独立であると仮定すると、これより θ についての尤度関数を書き下すことができる；

【数 2】

$$L(\theta) = \prod_i \Pr(\text{症例 } i \text{ の近親者の遺伝子型}; \theta), \quad (*)$$

【0118】

この独立性の仮定は、通常は正しくない。個体間の依存性を考慮することは、困難で極めて高価になる可能性がある計算タスクである。(*) の尤度関数は、すべての依存性を適正に考慮する θ についての全尤度関数 (full likelihood function) の擬似尤度近似と考えられ得る。一般に、症例 - 対照関連研究における遺伝子型が同定された症例および対照は独立ではなく、症例および対照を関連付けるために症例 - 対照法を適用することは、同様の近似である。ゲノム制御の方法 (Devlin, B. ら、Nat Genet 36, 1129 - 30; 著者による回答 1131 (2004)) は、関連性についての症例対照検定統計量を調整するうえで成功していることが判明している。従って出願人は、本発明の擬似尤度における項目間の依存性を考慮するためにゲノム制御の方法を適用し、妥当な検定統計量を生成する。

20

【0119】

遺伝子型が同定されていない症例に起因する擬似尤度の一部分の有効サンプルサイズを見積もるために、Fisher 情報量が使用され得る。全 Fisher 情報量、 I 、を遺伝子型が同定された症例に起因する部分、 I_g 、および遺伝子型が同定されていない症例に起因する部分、 I_u に分け、 $I = I_g + I_u$ 、そして遺伝子型が同定された症例の数を N で表すと、遺伝子型が同定されていない症例に起因する有効サンプルサイズは、

30

【数 3】

$$\frac{I_u}{I_g} N$$

によって評価される。

40

【0120】

本発明の構成では、疾病に対して増加した感受性 (すなわち、増加したリスク) を有する個体は、当該疾病について増加した感受性 (増加したリスク) を与える 1 つもしくは複数の多型マーカーの少なくとも 1 つの特定の対立遺伝子またはハプロタイプが同定された (すなわち、リスクのあるマーカー対立遺伝子またはハプロタイプ) 個体である。当該リスクのあるマーカーまたはハプロタイプは、当該疾病の増加したリスク (または感受性) を与えるものである。1 つの実施態様では、マーカーまたはハプロタイプに関連する有意性は、相対リスク (RR) によって判定される。別の実施態様では、マーカーまたはハプロタイプに関連する有意性は、オッズ比 (OR) によって判定される。さらなる実施態様では、当該有意性は、割合 (%) によって判定される。1 つの実施態様では、有意に増加

50

したリスクは、少なくとも1.2のリスク（相対リスクおよび/またはオッズ比）（少なくとも1.2、少なくとも1.3、少なくとも1.4、少なくとも1.5、少なくとも1.6、少なくとも1.7、1.8、少なくとも1.9、少なくとも2.0、少なくとも2.5、少なくとも3.0、少なくとも4.0、および少なくとも5.0を含むがこれらに限定されない）として判定される。特定の実施態様では、少なくとも1.2のリスク（相対リスクおよび/またはオッズ比）が有意である。別の特定の実施態様では、少なくとも1.3のリスクが有意である。さらに別の実施態様では、少なくとも1.4のリスクが有意である。さらなる実施態様では、少なくとも約1.5の相対リスクが有意である。別のさらなる実施態様では、リスクの有意な増加は、少なくとも約1.7が有意である。しかし、他のカットオフもまた考えられ、例えば、少なくとも1.15、1.25、1.35など、およびこのようなカットオフもまた本発明の範囲内である。他の実施態様では、リスクの有意な増加は、少なくとも約20%（約25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、100%、150%、200%、300%、および500%を含むがこれらに限定されない）である。1つの特定の実施態様では、リスクの有意な増加は、少なくとも20%である。他の実施態様では、リスクの有意な増加は、少なくとも40%、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%、および少なくとも100%である。本発明を特徴付ける、当業者によって適しているとみなされる他のカットオフまたは範囲も、しかしまた考えられ、それらもまた本発明の範囲内である。ある実施態様では、リスクの有意な増加は、p値が0.05未満、0.01未満、0.001未満、0.0001未満、0.00001未満、0.000001未満、0.0000001未満、0.00000001未満、または0.0000000001未満などの、p値によって特徴付けられる。

【0121】

本発明のリスクのある多型マーカーまたはハプロタイプは、少なくとも1つのマーカーの少なくとも1つの対立遺伝子またはハプロタイプが、比較群（対照）におけるその存在の頻度と比較して、疾病もしくは形質のリスクのある（罹患した）、または疾病もしくは形質と診断された個体においてより頻繁に存在するものであり、当該マーカーまたはハプロタイプの存在が疾病または形質（例えば、乳癌）に対する感受性を示す。1つの実施態様では、対照群は、集団の試料、すなわち一般的な集団から得た確率標本であってもよい。対照群は、別の実施態様では、疾病フリーの個体群（すなわち、乳癌と診断されていない個体）によって表される。このような疾病フリーの対照は、1つの実施態様では、1つまたは複数の特定の疾病に関連する症候の不存在によって特徴付けられてもよい。別の実施態様では、当該疾病フリーの対照群は、1つまたは複数の疾病特異的な危険因子の不存在によって特徴付けられる。このような危険因子は、1つの実施態様では、少なくとも1つの環境的な危険因子である。代表的な環境的因子は、特定の疾病または形質について発生のリスクに影響することが知られる、または影響すると考えられる天然物、鉱物または他の化学薬品である。他の環境的な危険因子は、生活スタイルに関連する危険因子である（飲食の習慣、主な居住環境の地理的な位置、および職業性危険因子を含むが、これらに限定されない）。別の実施態様では、危険因子は、少なくとも1つの遺伝的危険因子である。

【0122】

相関についての簡易な試験の例は、2×2の表のフィッシャーの正確確率検定であり得る。所定の染色体コホートについて、マーカーまたはハプロタイプの両方を含む染色体数、マーカーまたはハプロタイプのうち片方のみを含むが、もう片方は含まない染色体数、およびマーカーまたはハプロタイプのどちらも含まない染色体数から、2×2の表が作成される。当業者に知られた他の関連の統計検定もまた考えられ、これらもまた本発明の範囲内である。

【0123】

当業者は、2つの対立遺伝子を有する多型マーカーが、検討された集団に存在し（SN

10

20

30

40

50

P など)、1つの対立遺伝子が、当該集団において形質または疾病を有する個体群中で対照と比較して増加した頻度が見出されると、マーカーのもう1つの対立遺伝子は、形質または疾病を有する個体群中で、対照と比較して、減少した頻度で見出されることが理解するだろう。このような場合は、マーカーの1つの対立遺伝子(形質または疾病を有する個体群中で増加した頻度で見出されたもの)は、リスクのある対立遺伝子であり、一方、もう1つの対立遺伝子は保護的対立遺伝子となるだろう。

【0124】

従って、本発明の他の実施態様では、疾病または形質について減少した感受性(すなわち、減少したリスク)を有する個体は、疾病または形質について減少した感受性を与える1つまたは複数の多型マーカーの少なくとも1つの特定の対立遺伝子またはハプロタイプが同定された個体である。減少したリスクを与えるマーカー対立遺伝子および/またはハプロタイプはまた、保護的であると言える。1つの態様では、保護マーカーまたはハプロタイプは、疾病または形質の有意に減少したリスク(または感受性)を与えるものである。1つの実施態様では、有意に減少したリスクは、0.90未満の相対リスク(0.85未満、0.80未満、0.75未満、0.7未満、0.6未満、0.5未満、0.4未満、0.3未満、0.2未満および0.1未満を含むが、これらに限定されない)として判定される。1つの特定の実施態様では、有意に減少したリスクは0.90未満である。別の実施態様では、有意に減少したリスクは、0.85未満である。さらに別の実施態様では、有意に減少したリスクは、0.80未満である。別の実施態様では、リスク(または感受性)の減少は、少なくとも10%(少なくとも15%、少なくとも20%、少なくとも25%、少なくとも30%、少なくとも35%、少なくとも40%、少なくとも45%、少なくとも50%、少なくとも55%、少なくとも60%、少なくとも65%、少なくとも70%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも95%および少なくとも98%を含むが、これらに限定されない)である。1つの特定の実施態様では、リスクの有意な減少は、少なくとも約15%である。別の実施態様では、リスクの有意な減少は、少なくとも約20%である。別の実施態様では、リスクの減少は、少なくとも約25%である。本発明を特徴付ける、当業者によって適しているとみなされる他のカットオフまたは範囲も、しかしまた考えられ、それらもまた本発明の範囲内である。

10

20

【0125】

疾病または形質(例えば、乳癌)に関連する遺伝的変異は、所定の遺伝子型についての疾病のリスクを予測するために、単独で使用され得る。2対立形質のマーカー(SNPなど)については、3つの可能性のある遺伝子型がある(リスク変異のホモ接合体、ヘテロ接合体、およびリスク変異の非保有者)。複数の座位の変異と関連するリスクは、全体のリスクの評価に使用され得る。複数のSNP変異については、 k 個の可能性のある遺伝子型($k = 3^n \times 2^p$)がある。 n は、常染色体の座位数であり、 p は、ゴノソーマルな(gonosomal)(性染色体の)座位数である。全体のリスクアセスメントの算出は、通常、異なる遺伝的変異の相対リスクが掛かる、すなわち、特定の遺伝子型の組み合わせと関連する全体のリスク(例えば、RRまたはOR)が、各遺伝子座の遺伝子型のリスク値の積であることを想定している。性別および民族性をマッチさせた参照集団と比較して、示されたリスクが、ヒトまたはヒトの特定の遺伝子型の相対リスクである場合は、複合リスクは遺伝子座特異的なリスク値の積であり、また、これは当該集団と比較した全体的リスク推定値と一致する。ヒトのリスクが、リスク対立遺伝子の非保有者との比較に基づく場合、複合リスクは、全ての座位の遺伝子型の所定の組み合わせを有する個人を、これらの座位のいずれでもリスク変異を有さない個体群と比較した推定値と一致する。いずれのリスク変異も有さない非保有者群は、最も低いリスク推定値を有し、自身(すなわち、非保有者)と比較した複合リスクが1.0であるが、集団と比較した全体のリスクは1.0未満である。非保有者群は潜在的に、特に多数の座位について、非常に小さいものであり得ることに注意すべきであり、この場合、その関連性は同様に小さい。

30

40

【0126】

50

複合モデルは、通常、複合形質のデータにまずまず適合する簡潔なモデルである。多重度からの偏差は、一般的な疾病の一般的な変異に関連してまれに記載され、報告されたとしても、座位間の統計的相互作用を示すことができるように、非常に大きなサンプルサイズが通常要求されることから、通常、示唆的であるのみである。

【0127】

例として、全部で7つの変異が乳癌に関連付けられている場合を考えることにする。1つのそのような例は、マーカー $rs13387042$ 、 $rs4415084$ 、 $rs1219648$ 、 $rs3803662$ 、 $rs13281615$ 、 $rs3817198$ および $rs889312$ によって与えられ、それらのうちのすべては、乳癌感受性についての市販されている `decODE BreastCancer` 試験で使用される (<http://www.decodediagnostics.com>)。理論上の遺伝子型の組み合わせの総数は、 $3^7 = 2187$ である。これらの遺伝子型のクラスのいくつかは非常にまれであるが、依然として可能性があり、全体のリスクアセスメントに考慮されるべきである。複数の遺伝的変異の場合に適用される複合モデルはまた、遺伝的変異が「環境的な」因子と明瞭には相関していないとすると、非遺伝的危険変異と合わせることが妥当であろうと思われる。換言すれば、遺伝的リスクおよび非遺伝的リスクのある変異は、非遺伝的危険因子および遺伝的危険因子が相互作用していないとすると、複合リスクを推定するために複合モデル下で評価し得る。

10

【0128】

同様の定量的アプローチを使用し、乳癌に関連する多数のこれらのおよび他の変異に関連する複合または全体のリスクを評価してもよい。これは、乳癌のリスクを予測すると本願明細書中で示され主張される変異を含む。

20

【0129】

(連鎖不平衡)

各減数分裂事象の間に各染色体対に平均1度生じる、組み換えの自然現象は、自然が配列変異(および、結果生じる生物学的機能)をもたらす1つの方法を表す。組み換えはゲノム中でランダムには生じないことが発見された。ある程度、組み換え率の頻度に大きな変異があり、高い組み換え頻度の小領域(組み換えホットスポットとも呼ばれる)、および、低い組み換え頻度のより大きな領域(通常、連鎖不平衡(LD)ブロックと呼ばれる)をもたらす(Myers, S.ら、*Biochem Soc Trans* 34:526-530(2006); Jeffrey, A. J.ら、*Nature Genet* 29:217-222(2001); May, C. A.ら、*Nature Genet* 31:272-275(2002))。

30

【0130】

連鎖不平衡(LD)は、2つの遺伝因子の非ランダム分類を意味する。例えば、特定の遺伝因子(例えば、多型マーカーの対立遺伝子、またはハプロタイプ)が0.50(50%)の頻度で集団に生じ、別の因子が0.50(50%)の頻度で生じると、当該因子のランダム分布を仮定すると、ヒトが両因子を有すると予測される発生率は0.25(25%)である。しかし、2つの因子が0.125より高い頻度でともに生じることが発見されると、それらの別個の発生頻度(例えば、対立遺伝子またはハプロタイプの頻度)における予測よりも高い割合で、ともに遺伝する傾向にあるため、当該因子は、連鎖不平衡であると言える。大まかに言うと、LDは、一般的に、2つの因子間の組み換え事象の頻度と相関する。対立遺伝子またはハプロタイプの頻度は、集団内の個体の遺伝子型の同定、および集団の各対立遺伝子またはハプロタイプの発生頻度の測定によって、集団において判定し得る。二倍体の集団(例えば、ヒト集団)については、個体は、典型的に各遺伝因子(例えば、マーカー、ハプロタイプまたは遺伝子)について2つの対立遺伝子を有するだろう。

40

【0131】

多くの異なる基準が、連鎖不平衡(LD; Devlin, B. および Risch, N., *Genomics* 29:311-22(1995))において概説されている)の強さ

50

の評価について提案されてきた。大部分は、対の2対立形質の部位間の関連の強さをとらえる。2つの重要な対のLDの基準は、 r^2 (r^2 と表示されることもある)および $|D'|$ である(Lewontin, R., *Genetics* 49:49-67 (1964); Hill, W. G. および Robertson, A., *Theor. Appl. Genet.* 22:226-231 (1968))。両基準とも、0(不平衡がない)から1(「完全な」不平衡)の範囲であるが、それらの解釈は少し異なる。 $|D'|$ は、2つのマーカーについて2つまたは3つのみの可能性のあるハプロタイプが存在する場合1に等しく、全ての4つの可能性のあるハプロタイプが存在する場合 < 1 となるように定義される。従って、 < 1 の $|D'|$ 値は、背景の組み換えが、2つの部位間(再発性変異もまた、 $|D'|$ を < 1 にし得るが、一塩基多型(SNP)については、これは通常、組み換えよりも可能性が少ないと考えられる)で起きたかも知れないことを示す。基準の r^2 は、2つの部位間の統計的相関を表し、2つのハプロタイプのみが存在する場合は、値は1である。

【0132】

r^2 基準は、 r^2 と、感受性座位とSNPとの間の関連を検出するのに必要な試料サイズとの間に単純な逆の関連性があるため、関連マッピングについてのほぼ間違いなく最も関連する基準である。これらの基準は、対の部位について定義されるが、いくつかの適用については、どの位強いLDが、多くの多形性部位を含む全体の領域をわたっているかについての判定が望ましいだろう(例えば、LDの強さが座位間もしくは集団にわたって有意に異なっているかについての試験、または特定のモデル下で、領域に、予測されるよりも多かれ少なかれLDがあるかについての試験)。大まかに言うと、 r は、データに見られるLDを生じるため、特定の集団モデル下で、どの位の組み換えが必要とされるかを判定する。この型の方法はまた、潜在的に、LDデータが組み換えホットスポットの存在についてのエビデンスを提供するかどうかについての判定の問題に対する、統計的に厳密なアプローチを提供し得る。本願明細書記載の方法において、連鎖不平衡であるマーカーを示すマーカーの間の有意な r^2 値は、少なくとも0.1(少なくとも0.15、0.2、0.25、0.3、0.35、0.4、0.45、0.5、0.55、0.6、0.65、0.7、0.75、0.8、0.85、0.9、0.91、0.92、0.93、0.94、0.95、0.96、0.97、0.98、または少なくとも0.99など)であり得る。1つの好ましい実施態様では、有意な r^2 値は、少なくとも0.2であり得る。あるいは、連鎖不平衡のマーカーは、本願明細書に記載される場合、少なくとも0.2の $|D'|$ 値(0.3、0.4、0.5、0.6、0.7、0.8、0.85、0.9、0.95、0.96、0.97、0.98、または少なくとも0.99など)によって特徴付けられる。従って、連鎖不平衡は、異なるマーカーの対立遺伝子の間の相関を表す。ある実施態様では、連鎖不平衡は、 r^2 および $|D'|$ 基準の両方についての値によって定義される。1つのこのような実施態様では、有意な連鎖不平衡は $r^2 > 0.1$ および $|D'| > 0.8$ として定義され、これらの基準を満たすマーカーは連鎖不平衡であると言われる。別の実施態様では、有意な連鎖不平衡は、 $r^2 > 0.2$ および $|D'| > 0.9$ として定義される。連鎖不平衡を判定するための r^2 および $|D'|$ の値の他の組み合わせおよび順列(permutation)もまた企図され、それらも本発明の範囲内である。連鎖不平衡は、本願明細書で定義されたように、単一のヒト集団において判定し得、または1人より多いヒト集団からの個体を含む試料の収集物において判定し得る。本発明の1つの実施態様では、LDは、(<http://www.hapmap.org>)で定義されたように、1つまたは複数のHapMap集団(コーカサス人、アフリカ人(ヨルバ族)、日本人、中国人)から得られた試料で判定される。このような実施態様の1つでは、LDは、HapMap試料のCEU集団(北ヨーロッパおよび西ヨーロッパ出身の祖先を有するユタ州の居住者)で判定される。別の実施態様では、LDは、HapMap試料(ナイジェリア、イバダン(Ibadan)のヨルバ族)のYRI集団において判定される。別の実施態様では、LDは、HapMap試料(中国、北京出身の漢民族)のCHB集団において判定される。別の実施態様では、LDは、HapMap試料(日本、東京出

10

20

30

40

50

身の日本人)のJPT集団において判定される。さらに別の実施態様では、LDは、アイ
スランド人の集団由来の試料において判定される。

【0133】

ゲノム中の全ての多型性が、集団レベルで同一の場合、それらのことごとくが、関連研
究において検討されることが必要だろう。しかし、多型間の連鎖不平衡により、密接に連
鎖した多型は強く関連し、有意な関連を観察するために関連研究において検討される必要
がある多型の数は減る。LDの別の結果は、多くの多型が、これらの多型が強く関連する
という事実によって関連シグナルを与える可能性がある、ということである。

【0134】

ゲノムLDマップは、ゲノムにわたって作成されてきており、このようなLDマップは
、疾病遺伝子のマッピングのためのフレームワークとして役立つと提案されてきている (10
Risch N. および Merkiangas K, Science 273:1516
-1517 (1996); Maniatis N. ら、Proc Natl Acad
Sci USA 99:2228-2233 (2002); Reich DE ら、Nat
ure 411:199-204 (2001))。

【0135】

今では、ヒトゲノムの多くの部分が、少数の共通のハプロタイプを含む一連の別々のハ
プロタイプブロックに切断され得ることが確立されている。これらのブロックについては
、連鎖不平衡データは、組み換えを示すエビデンスをわずかに提供するにすぎない (例え
ば、Wall J. D. および Pritchard, J. K., Nature Rev
20
iews Genetics 4:587-597 (2003); Daly, M. ら、N
ature Genet. 29:229-232 (2001); Gabriel, S. B
. ら、Science 296:2225-2229 (2002); Patil, N. ら
、Science 294:1719-1723 (2001); Dawson, E. ら、
Nature 418:544-548 (2002); Phillips, M. S. ら、
Nature Genet. 33:382-387 (2003) を参照)。

【0136】

これらのハプロタイプブロックを定義する2つの主な方法がある。ブロックは、限られ
たハプロタイプの多様性を有するDNAの領域 (例え
ば、Daly, M. ら、Natur
e Genet. 29:229-232 (2001); Patil, N. ら、Scien
30
ce 294:1719-1723 (2001); Dawson, E. ら、Natur
e 418:544-548 (2002); Zhang, K. ら、Proc. Natl. A
cad. Sci. USA 99:7335-7339 (2002) 参照)、または、連鎖
不平衡を使用して同定された、広範な背景の組み換えを有する転位ゾーン間の領域 (例え
ば、Gabriel, S. B. ら、Science 296:2225-2229 (20
02); Phillips, M. S. ら、Nature Genet. 33:382-3
87 (2003); Wang, N. ら、Am. J. Hum. Genet. 71:1227
-1234 (2002); Stumpf, M. P. および Goldstein, D. B.
、Curr. Biol. 13:1-8 (2003) 参照) として定義され得る。つい最近
40
では、ヒトゲノムにわたる組み換え率および対応するホットスポットの、詳細なスケール
のマップが、作成されてきている (Myers, S. ら、Science 310:32
1-32324 (2005); Myers, S. ら、Biochem Soc Tran
s 34:526530 (2006))。当該マップは、組み換え率が、ホットスポット
中で10~60cM/Mbほど高く、一方、介在性領域では0に近く、従って、限られた
ハプロタイプの多様性および高いLDの領域を表す、ゲノムにわたる組み換えにおける莫
大な変異を明らかにした。従って、当該マップは、組み換えホットスポットに隣接した領
域として、ハプロタイプブロック/LDブロックを定義することに使用され得る。本願明
細書で使用されるように、用語「ハプロタイプブロック」または「LDブロック」は、上
述の特性、またはこのような領域を定義するために当業者によって使用される他の別の方
法のいずれかによって定義されるブロックを含む。

10

20

30

40

50

【 0 1 3 7 】

例えば本願明細書に記載される用語「LDブロックC09」は、NCBI（国立バイオテクノロジー情報センター）ビルド36の位置36,806,001～36,859,001の間の染色体9上の連鎖不平衡（LD）ブロックを指す。以下の表は、本願中で言及される、NCBIビルド36中の9つのLDブロックの染色体上の位置を定める。

【 0 1 3 8 】

【表2】

LD-ブロックの表 (NCBI ビルド 36 からの位置)

染色体	LD ブロック	ブロック開始 B36	ブロック末端 B36
C09	LDブロックC09	36,806,001	36,859,001
C10	LDブロックC10A	8,643,001	8,817,001
C10	LDブロックC10B	9,077,001	9,264,001
C11	LDブロックC11A	8,053,268	8,191,268
C11	LDブロックC11B	74,886,341	74,971,341
C14	LDブロックC14	68,035,712	68,130,712
C18	LDブロックC18	32,110,012	32,145,012
C22	LDブロックC22A	38,704,907	38,859,907
C22	LDブロックC22B	38,859,907	39,411,907

【 0 1 3 9 】

ハプロタイプブロック（LDブロック）は、1つのマーカーまたは複数のマーカーを含むハプロタイプを使用して、表現型とハプロタイプ状態との間の関連をマッピングするために使用され得る。これらの主なハプロタイプは各ハプロタイプブロックにおいて特定され得、次いで一連の「タグした（tagging）」SNPまたはマーカー（当該ハプロタイプを区別するために必要とされるSNPまたはマーカーの最小の組）が同定され得る。これらのタグしたSNPまたはマーカーは、表現型とハプロタイプとの間の関連を特定するために、個体群由来の試料の評価で使用され得る。乳癌に関連すると本願明細書に示されるマーカーは、このようなタグしたマーカーである。当該ハプロタイプブロック間で連鎖不平衡が存在し得るため、所望の場合、隣接するハプロタイプブロックは、同時に評価され得る。

【 0 1 4 0 】

従って、ゲノム中の多型マーカーへの、所定の認められた関連のいずれについても、ゲノム中のさらなるマーカーもまた関連を示すことが明らかになって来た。このことは、組み換え率の大きな変動によって認められたように、ゲノムにわたるLDの不均等な分布の当然の結果である。従って、関連の検出のために使用されたマーカーは、ある意味では、所定の疾病または形質と関連する、ゲノム領域（すなわち、ハプロタイプブロックまたはLDブロック）の「タグ」を表す。1つまたは複数の原因である（機能的）変異体または変異は、疾病または形質と関連することが見出された領域内に属していてもよい。機能的な変異体は、別のSNP、タンデムリピート多型性（ミニサテライトまたはマイクロサテライトなど）、転位因子、またはコピー数の変異（逆位、欠損または挿入など）であってもよい。本願明細書に記載された変異体とLDにある、このような変異は、関連の検出のために使用されたタグしたマーカーについて認められたものより、高い相対リスク（RR）またはオッズ比（OR）を与えてもよい。従って、本発明は、本願明細書記載の疾病の関連の検出のために使用されたマーカー、および当該マーカーとの連鎖不平衡のマーカーを意味する。従って、本発明のある実施態様では、関連を検出するために最初に使用され

たマーカーとLDのマーカーは、代用マーカーとして使用されてもよい。代用マーカーは、1つの実施態様では、相対リスク(RR)および/またはオッズ比(OR)の値が、最初に検出されたものより小さい。他の実施態様では、代用マーカーは、RRまたはOR値が、当該疾病に関連することが最初に見出されたマーカーで最初に判定されたものより大きい。このような実施態様の例は、当該疾病と関連することが最初に見出された、より一般的な変異体(>10%集団における頻度)とLDにある、まれ、または相対的にまれ(対立遺伝子の集団における頻度が<10%)な変異体であろう。関連の検出のためにこのようなマーカーを同定および使用することは、当業者に周知のルーチンの方法によって行われ得、従ってこれらは本発明の範囲である。

【0141】

10

(ハプロタイプ頻度の測定)

患者および対照群におけるハプロタイプの頻度は、期待値最大化アルゴリズムを使用して評価され得る(Dempster A.ら、J. R. Stat. Soc. B、39:1-38(1977))。フェーズとともに、欠損した遺伝子型および不確定度を扱うことができる当該アルゴリズムの実行が使用され得る。帰無仮説下では、患者および対照は、同一の頻度を有すると想定される。尤度法を使用し、別の仮説が検討され、候補のリスクのあるハプロタイプ(本願明細書記載のマーカーを含み得る)は、対照よりも患者において高い頻度を有していてもよく、一方、他のハプロタイプの頻度の割合は、両群で同一であると想定される。尤度は、両仮説下で別々に最大化され、対応する1-dfの可能性の割合の統計量は、統計的有意性を評価するために使用される。

20

【0142】

感受性領域内(例えば、LDブロック領域内)のリスクのあるマーカーおよび保護マーカーならびにハプロタイプを探索するため、当該領域内の遺伝子型を同定されたマーカーの全ての可能性のある組み合わせの関連が検討される。合わされた患者および対照群は、本来の患者群および対照群とサイズが等しい、ランダムに2つのセットに分けられ得る。次いで、マーカーおよびハプロタイプ分析は、繰り返され、最も有意な記載されたp値が判定される。当該ランダム化スキームは、例えば、実験に基づいたp値の分布を作成するために、100回を超えて繰り返され得る。好ましい実施態様では、p値が<0.05であることは、有意なマーカーおよび/またはハプロタイプの関連性を示す。

【0143】

30

(ハプロタイプ分析)

ハプロタイプ分析の1つの一般的なアプローチは、ネステッド(Nested)モデルに適用した尤度に基づいた推論の使用に関する(Gretarsdottir S.ら、Nat. Genet. 35:131-38(2003))。当該方法は、多くの多型マーカー、SNPおよびマイクロサテライトに適用してもよい、プログラムNEMOで行われる。当該方法およびソフトウェアは、異なるリスクを与えるハプロタイプ群を同定することを目的とする、症例と対照の研究のために特に設計される。これはまた、LD構造を研究するためのツールである。NEMOでは、EMアルゴリズムを活用して、観察されたデータについて最尤推定値、尤度比およびp値が直接的に算出され、観察されたデータについては欠損データの問題として扱う。

40

【0144】

フェーズにおける不確定度および欠損遺伝子型による情報の損失をとらえる、観察されたデータについて直接的に算出される尤度に基づく尤度比検定は、所定の妥当なp値を与えるために依存され得るが、情報が不完全であることでどの位の量の情報が失われているのかについて知ることは依然として興味深い。ハプロタイプ分析のための情報測定は、NicolaieおよびKong(Technical Report 537、シカゴ大学、統計大学(University of Statistics)、統計学科; Biometrics、60(2):368-75(2004))に、連鎖分析のために定義された情報測定の自然な伸長として記載され、NEMOにおいて実行される。

【0145】

50

(関連分析)

疾病に関連する単一のマーカーについては、各個体の対立遺伝子の両側 p 値を算出するために、フィッシャーの直接確率検定が使用され得る。同胞群のためのこれまでに記載された分散調整の手順 (Risch, N. および Teng, J. *Genome Res.*, 8: 1273 - 1288 (1998)) を拡張することで、患者間の関連性の補正がなされ得、それは一般的な家族性の関連に適用され得る。ゲノム制御の方法 (Devlin, B. および Roeder, K. *Biometrics* 55: 997 (1999)) はまた、個体および可能な層別化の関連性を調整するために使用され得る。

【0146】

単一のマーカーおよびハプロタイプ分析の両方について、相対リスク (RR) および集団に起因するリスク (PAR) は、複合モデル (ハプロタイプの相対リスクモデル) を仮定して算出され得る (Terwilliger J. D. および Ott, J., *Hum. Hered.* 42: 337 - 46 (1992)、ならびに Falk, C. T. および Rubinstein, P., *Ann. Hum. Genet.* 51 (Pt 3): 227 - 33 (1987))。すなわち、ヒトが保有する2つの対立遺伝子/ハプロタイプのリスクは乗算で表される。例えば、RR が、a に関連する A のリスクである場合、ヒトホモ接合体 AA のリスクは、ヘテロ接合体 Aa のリスクの RR 倍となり、ホモ接合体 aa のリスクの RR^2 倍となる。複合モデルは、分析および計算を単純化する良好な性質を有する。ハプロタイプは罹患した集団内で、対照集団内同様に別個であり、すなわち、ハーディ・ワインベルグ平衡にある。結果として、罹患者および対照のハプロタイプ数は、それぞれ多項分布を有するが、別の仮説下で異なるハプロタイプの頻度を有する。特に、2つのハプロタイプ (h_i および h_j) については、 $\text{リスク}(h_i) / \text{リスク}(h_j) = (f_i / p_i) / (f_j / p_j)$ であり、f および p は、それぞれ、罹患した集団における頻度および対照集団における頻度を表す。真のモデルが増殖性ではない場合、いくらかの検出力の損失がある一方で、当該損失は、極端な症例以外では、軽度である傾向がある。最も重要なのは、帰無仮説に関して算出されるため、p 値が常に妥当であることである。

【0147】

1つの関連研究で検出される関連シグナルは、理想的には、同一のまたは異なる民族性の、異なる集団 (例えば、同一の国の異なる領域、または異なる国) からの、第2のコホートで再現されてもよい。追試の利点は、追試で実施される試験の数が通常は非常に小さく、従って、適用される必要がある統計的基準はより厳密ではないことである。例えば、300,000 SNP を使用した特定の疾病または形質の感受性変異体のゲノム全体での検索に対し、実施される300,000試験 (各SNPについて1つ) について補正を行うことができる。典型的に使用されるアレイ上の多くのSNPは、関連しているため (すなわち、LDにあるため)、これらは独立ではない。従って、この補正は、保守的である。それにもかかわらず、この補正因子の適用は、この保守的検定 (conservative test) を単一の研究コホートから得た結果に適用するとき、そのシグナルが有意であるとみなされるためには、観察されたP値が、 $0.05 / 300,000 = 1.7 \times 10^{-7}$ 未満であることを必要とする。P値がこの保守的な閾値未満である、ゲノム全体での関連研究で見出されたシグナルは、真の遺伝効果の基準であることは明らかであり、さらなるコホートでの再現は、必ずしも統計的な観点からのものではない。しかしながら重要なことには、この閾値よりも大きいP値を有するシグナルは、真の遺伝効果にも起因し得る。この第1の研究におけるサンプルサイズは、ゲノム全体での有意性についての保存的な閾値を満たす観察されたP値を与えるのに十分に大きくはなかったかも知れないし、または当該第1の研究は、サンプリングに起因する内在的なゆらぎに起因してゲノム全体での有意性に到達しなかったかも知れない。しかし、当該補正因子は、実施された統計検定の数に依存するため、最初の研究から得た1つのシグナル (1つのSNP) が、第2の症例 - 対照コホートで再現された場合、有意性についての適切な統計検定は、1回の統計検定についてのものであり、すなわち、P値が0.05未満である。1つまたはさらにいくつかの追加の症例 - 対照コホートでの追試は、さらなる集団における関連シグナル

の評価を与えるという付加された利点を有し、従って、最初の発見を確認し、ヒト集団一般で試験される遺伝的変異（1つまたは複数）の全体の有意性の評価を与える。

【0148】

いくつかの症例 - 対照コホートから得た結果は、根底にある効果の全体の評価を与えるために合わされてもよい。複数の遺伝関連研究から得た結果を合わせるために一般的に使用される方法論は、マンテル・ヘンツェルモデルである（MantelおよびHaenszel、J Natl Cancer Inst 22:719-48（1959））。当該モデルは、それぞれがおそらく遺伝的変異の異なる集団頻度を有する、異なる集団から得た関連結果が合わされる状況を扱うために設計されている。当該モデルは、ORまたはRRによって判定される当該疾病のリスクに対する変異体の効果は、全ての集団で同一であるが、当該変異体の頻度は集団間で異なっているかも知れないと仮定して結果を合わせる。いくつかの集団から得た結果を合わせることは、合わせられたコホートによって与えられる、増加した統計的検出力による、現実の根底にある関連シグナルの全体の検出力が増加するという、付加された利点を有する。さらに、個体研究のいずれの欠損（例えば、症例および対照の不均等な組み合わせまたは集団の層別化による）は、複数のコホートから得た結果が合わせられると、相殺される傾向があるだろうから、ここでも再度、真の根底にある遺伝効果のより良い評価を与えるだろう。

10

【0149】

（リスクアセスメントおよび診断）

いずれの所定の集団内でも、疾病または形質が発生する絶対的なリスクがあり、特定の期間にわたってヒトが特定の疾病または形質を生じる可能性として定義される。例えば、乳癌についての女性の生涯の絶対的なリスクは9分の1である。つまり、9人のうち1人の女性は、その人生のある時点で乳癌を生じる。典型的には、リスクは、特定の個体を見るよりも、非常に多くの人々を見ることによって判定される。リスクはしばしば絶対的なリスク（AR）および相対リスク（RR）の観点から示される。相対リスクは、2つの変異に関連するリスクまたは人々の2つの異なる群のリスクを比較するために使用される。例えば、ある遺伝子型を有する人々の群を、異なる遺伝子型を有する別の群と比較するために使用され得る。疾病について、2の相対リスクは、1つの群が他の群に対して疾病を生じる2倍の可能性を有することを意味する。示されたリスクは、通常、性別および民族性をマッチさせた集団と比較した、ヒト、またはヒトの特定の遺伝子型についての相対リスクである。同一の性別および民族性の2個体のリスクは、簡潔な様式で比較され得る。例えば、集団と比較して、第1の個体の相対リスクが1.5であり、第2の個体の相対リスクが0.5である場合、第2の個体と比較した第1の個体のリスクは、 $1.5 / 0.5 = 3$ である。

20

30

【0150】

（リスクの算出）

全体の遺伝的危険率を算出するためのモデルの創出には、2つの工程を伴う：i) 1つの遺伝的変異についてのオッズ比を相対リスクに変換すること、およびii) 異なる遺伝子座における複数の変異に由来するリスクを1つの相対リスク値へと組み合わせること。

【0151】

（オッズ比からのリスクの誘導）

現在までに権威ある雑誌で刊行された、複合疾患についてのほとんどの遺伝子発見の研究は、それらのレトロスペクティブな設定のため、症例対照設計を用いてきた。これらの研究は、選択された組の症例（特定された病状を有する人々）および対照個体をサンプリングし、その遺伝子型を同定する。興味があるのは、症例および対照における頻度が有意に異なる遺伝的変異（対立遺伝子）である。

40

【0152】

結果は、典型的には、オッズ比で報告される。これは、罹患者の群 対 対照におけるリスク変異体（保有者）対非リスク変異体（非保有者）の割合（確率）の間の比であり、すなわち当該罹患状態を条件とする確率によって表される：

50

$$OR = (Pr(c|A) / Pr(nc|A)) / (Pr(c|C) / Pr(nc|C))$$

【0153】

しかしながら、時には、出願人らが興味を持っているのは、当該疾病についての絶対的なリスク、すなわち当該疾病にかかる当該リスク変異体を保有する個体の割合、または換言すると当該疾病にかかる確率であるということもある。この数は、症例対照研究では直接判定され得ない。なぜなら、一部には、症例対対照の比は、典型的には、一般的な集団におけるものとは同じではないからである。しかしながら、ある仮定の下で、このオッズ比からリスクを評価し得る。

【0154】

稀な疾病の仮定の下では、疾病の相対リスクはオッズ比によって近似し得るということは周知である。しかしながらこの仮定は、多くの一般的な疾病に対しては適用できないかもしれない。それでもなお、別の遺伝子型変異体に対する1つの遺伝子型変異体のリスクは、上で表されたオッズ比から評価され得るということは明らかである。対照が症例と同じ集団に由来する確率標本であり、厳密に罹患していない個体であるというよりは罹患した人々を含むランダム集団の対照という仮定の下では、算出は特に簡単である。サンプルサイズおよび検出力を高めるために、大きいゲノム全体での関連および再現研究の多くは、症例と年齢をマッチさせていない対照を使用し、かつその対照は、対照がその研究時に当該疾病を有していなかったということを確認するために慎重に精査されてもいない。

【0155】

従って、正確ではないが、それらは、一般的集団に由来する確率標本を近似することが多い。この仮定が正確に満足されるということは稀であると予想されるが、当該リスク評価は、この仮定からの少しのずれに対しては通常はロバストであるということに留意されたい。

【0156】

リスク変異体保有者「c」、および非保有者「nc」を有する優性および劣性モデル(the dominant and the recessive model)について、個体のオッズ比はこれらの変異体間のリスク比と同じであるということが、算出によって示される：

$$OR = Pr(A|c) / Pr(A|nc) = r$$

【0157】

また同様に、リスクが2つの対立遺伝子のコピーに関連するリスクの積である複合モデルについて、対立遺伝子のオッズ比は危険因子に等しい：

$$OR = Pr(A|aa) / Pr(A|ab) = Pr(A|ab) / Pr(A|bb) = r$$

【0158】

ここで「a」はリスク対立遺伝子を表し、「b」は非リスク対立遺伝子を表す。従って因子「r」は、対立遺伝子のタイプの間の相対リスクである。

【0159】

複合疾患に関連する一般的な変異を報告する過去数年に刊行された研究の多くについて、当該複合モデルは、当該効果を十分に要約し、かつ優性および劣性モデルなどの代替のモデルよりも優れた当該データに対する一致をもっとも頻繁に与えるということが見出された。

【0160】

(平均的な集団のリスクに対するリスク)

平均的な集団に対する遺伝的変異のリスクを表すことがもっとも簡便である。なぜなら、それは、ベースラインの集団のリスクと比べて、当該疾病の発生についての生涯リスクを伝えることをより容易にするからである。例えば、複合モデルでは、変異体「aa」についての相対的な集団リスクは、

$$RR(aa) = Pr(A|aa) / Pr(A) = (Pr(A|aa) / Pr(A|bb)) / (Pr(A) / Pr(A|bb)) = r^2 / (Pr(aa)r^2 + Pr(ab)r + Pr(bb)) = r^2 / (p^2r^2 + 2p$$

10

20

30

40

50

$$q(r + q^2) = r^2 / R$$

として算出され得る。

【0161】

ここで「p」および「q」は、それぞれ「a」および「b」の対立遺伝子の頻度である。同様に、 $RR(ab) = r/R$ および $RR(bb) = 1/R$ が得られる。対立遺伝子の頻度の評価は、オッズ比を報告する刊行物から、および HapMap データベースから、入手され得る。個体の遺伝子型が不明である症例においては、その試験またはマーカーについての相対的な遺伝的危険率は単に 1 に等しいということに留意されたい。

【0162】

一例として、染色体 14 上のマーカー rs999737 について、対立遺伝子 C は、コカサス人の集団では 1.15 の乳癌についての対立遺伝子の OR、および約 0.76 の頻度 (p) を有する。遺伝子型 TT と比較した当該遺伝子型の相対リスクは、上記複合モデルに基づき評価される。

10

【0163】

CC については、それは $1.15 \times 1.15 = 1.32$ である；CT については、それは単に OR 1.15 であり、TT については定義によりそれは 1.0 である。

【0164】

対立遺伝子 T の頻度は、 $q = 1 - p = 1 - 0.76 = 0.24$ である。このマーカーにおけるこの 3 つの可能性のある遺伝子型の各々の集団頻度は、 $Pr(CC) = p^2 = 0.58$ 、 $Pr(CT) = 2pq = 0.36$ 、および $Pr(TT) = q^2 = 0.06$

20

である。

【0165】

遺伝子型 TT (これは 1 のリスクを有すると定義されている) に対する平均的な集団リスクは、

$$R = 0.50 \times 1.32 + 0.36 \times 1.15 + 0.06 \times 1 = 1.13$$

である。

【0166】

従って、このマーカーにおける以下の遺伝子型のうちの 1 つを有する個体についての一般的集団に対するリスク (RR) は、

30

$$RR(CC) = 1.32 / 1.13 = 1.17, \quad RR(CT) = 1.15 / 1.13 = 1.02, \quad RR(TT) = 1 / 1.13 = 0.88$$

である。

【0167】

複数のマーカーからのリスクの組み合わせ：

多くの SNP 変異の遺伝子型が個体についてのリスクを評価するために使用される場合、リスクについての複合モデルが一般的に仮定され得る。これは、当該集団に対する合わせた遺伝的危険率が、個々のマーカーについて、例えば 2 つのマーカー g1 および g2 についての対応する評価の積として算出される：

$$RR(g1, g2) = RR(g1) RR(g2)$$

40

ということの意味する。

【0168】

根底にある仮定は、危険因子は独立に発生し振舞うということ、すなわち結合条件付き確率は積として表され得る：

$$Pr(A | g1, g2) = Pr(A | g1) Pr(A | g2) / Pr(A) \text{ および } Pr(g1, g2) = Pr(g1) Pr(g2)$$

ということである。

【0169】

この仮定に対する明白な違反は、2 以上のリスク対立遺伝子の同時発生に相関性があるほどに、当該ゲノム上で密に間隔をあけられた、すなわち連鎖不平衡にあるマーカーであ

50

る。このような場合では、相関性があるSNPのすべての対立遺伝子の組み合わせに対してオッズ比が定義されるいわゆるハプロタイプモデリングが使用され得る。

【0170】

統計モデルが利用されるほとんどの場合のように、適用されるモデルは、まさに真実であるとは予想されない。なぜなら、それは、根底にある生物物理学的モデルに基づくからである。しかしながら、当該複合モデルは、今までのところ十分にデータに一致することが見出されている。すなわち多くのリスク変異体が発見されてきた多くの一般的な疾病について、著しいずれは検出されていない

【0171】

一例として、各マーカーにおけるこの集団に対するリスクとともに、特定の疾病に関連する4つの仮定上のマーカーにおいて以下の遺伝子型を有する個体について：

【0172】

【表3】

マーカー	遺伝子型	算出されたリスク
M1	CC	1.03
M2	GG	1.30
M3	AG	0.88
M4	TT	1.54

10

20

【0173】

合わせると、この個体についてのこの集団に対する全体のリスクは、 $1.03 \times 1.30 \times 0.88 \times 1.54 = 1.81$ である。同様にして、いずれかの複数のマーカー（またはハプロタイプ）についての全体のリスクが評価され得る。

【0174】

（調整された生涯リスク）

個体の生涯リスクは、集団に対する全体の遺伝的危険率に、同じ民族性および性別の一般的集団におけるおよび当該個体の地理的起源の領域における当該疾病の平均の生涯リスクを乗じることにより誘導される。一般的集団リスクを定める際に選択するためのいくつかの疫学的研究が通常は存在するので、遺伝的変異について使用されてきた疾病の定義のために十分に検討された（well-powered）研究を取り上げることにする。

30

【0175】

例えば、特定の疾病について、集団に対する全体の遺伝的危険率が個体について1.8である場合、およびその個体の層（demographic）の個体についての当該疾病の平均的な生涯リスクが20%である場合、その個体についての調整された生涯リスクは $20\% \times 1.8 = 36\%$ である。

【0176】

集団についての平均RRは1であるため、この複合モデルは当該疾病の同じ平均的な調整された生涯リスクを与えるということに留意されたい。さらに、現実の生涯リスクは100%を超え得ないため、遺伝的RRに対して上限があるべきである。

40

【0177】

（乳癌についてのリスクアセスメント）

本願明細書記載のように、このようなマーカーを含むある多型マーカーおよびハプロタイプは、乳癌のリスクアセスメントに有用であることが見出された。リスクアセスメントは、乳癌に対する感受性の診断のマーカーの使用に関連し得る。ある多型マーカーの特定の対立遺伝子は、乳癌と診断されていない個体と比較して、乳癌を有する個体において、より頻繁であることが見出された。従って、これらのマーカー対立遺伝子は、個体において、乳癌または乳癌に対する感受性の検出の予測値を有する。本願明細書記載のリスクのある変異体（または保護的変異体）と連鎖不平衡のタグしたマーカーは、これらのマーカー（および/またはハプロタイプ）の代替物として使用され得る。このような代用マーカー

50

ーは、特定のハプロタイプブロックまたはLDブロック内に位置し得る。このような代用マーカーはまた、時には、当該LDブロック/ハプロタイプブロックのごく近傍のいずれかに、このようなハプロタイプブロックまたはLDブロックの物理的な境界の外側に位置し得るが、より遠く離れたゲノム位置に位置する可能性もある。

【0178】

長距離のLDは、例えば、特定のゲノムの領域（例えば、遺伝子）が機能的に関連している場合に生じ得る。例えば、2つの遺伝子が共通の代謝経路において役割を果たすタンパク質をコードする場合、1つの遺伝子における特定の変異は、他の遺伝子についての観察された変異に対して直接的な影響を及ぼし得る。1つの遺伝子の中の変異が遺伝子産物の増加した発現につながるという場合を考えよう。この効果を弱めて特定の経路の全体のフラックスを保存するために、この変異体は、その遺伝子の減少した発現レベルを与える第2の遺伝子での1つ（以上）の変異の選択につながったかも知れない。これらの2つの遺伝子は、恐らく異なる染色体上の異なるゲノムの部位に位置し得るが、当該遺伝子内の変異は、高LDの領域内のそれらの共通の物理的な位置のためではなく、進化力により見かけ上LDである。このようなLDも本発明の範囲内で企図される。当業者なら、機能的な遺伝子-遺伝子相互作用の多くの他のシナリオが可能性を有し、本願明細書で論じられる特定の例は、1つのこのような可能性のあるシナリオを表しているに過ぎないということとはわかるであろう。

【0179】

r^2 値が1に等しいマーカーは、リスクのある変異体（アンカー変異体）の完全な代替物であり、すなわち、1つのマーカーの遺伝子型は、別の遺伝子型を完全に予測する。1より小さい r^2 値を有するマーカーはまた、リスクのある変異体の代替物となり得、あるいは、リスクのある変異体よりも高いかまたはさらに高い相対リスク値を有する変異体を表す。ある好ましい実施態様では、リスクのあるアンカー変異体に対して r^2 の値をもつマーカーは、有用な代用マーカーである。同定されたリスクのある変異体は、それ自体機能的変異体ではないかも知れないが、当該例では、真の機能的変異体と連鎖不平衡にある。機能的な変異体はSNPであってもよいが、例えば、タンデムリピート（ミニサテライトもしくはマイクロサテライトなど）、転位因子（例えば、Alu因子）、または構造上の変化（欠損、挿入または逆位（コピー数変異、またはCNVとも呼ばれることがある）など）であってもよい。本発明は、本願明細書で開示されたマーカーのこのような代用マーカーの評価を包含する。このようなマーカーは、当業者に周知のように、注解され、マッピングされ、および公的なデータベースに記載され、または、個体群における本発明のマーカーによって同定された領域または領域の一部の配列決定によって、代わりにすぐに同定され得、そして得られた配列群の多型性を同定する。結果として、当業者は、すぐに、かつ過度の実験をすることなく、本願明細書記載の、マーカーおよび/またはハプロタイプと連鎖不平衡の代用マーカーの遺伝子型を同定できる。検出されたリスクのある変異体とLDにあるこれらのタグしたマーカーまたは代用マーカーはまた、予測値を有する。

【0180】

ある実施態様では、本発明は、乳癌に関連すると本願明細書に記載の変異体の存在についての、個体から入手したゲノムDNAを含む試料の評価によって行われ得る。このような評価は、当業者に周知の方法およびさらに本願明細書記載の方法を使用し、少なくとも1つの多型マーカーの少なくとも1つの対立遺伝子の存在または不存在を検出する工程を含み、このような評価の成果に基づき、試料の由来である個体が、乳癌の増加したリスクまたは減少したリスク（増加した感受性または減少した感受性）にあるかを判定する。あるいは、本発明は、乳癌に関連すると本願明細書に記載の少なくとも1つの多型マーカー（または乳癌に関連すると本願明細書で示された少なくとも1つのマーカーと連鎖不平衡のマーカー）の遺伝子型の状態についての情報を含むデータセットを利用して行われ得る。換言すれば、例えば、ある多型マーカーまたは多数のマーカーにおける遺伝子型の数の形態にあるこのような遺伝的状態についての情報（例えば、あるリスクのある対立遺伝子の存在または不存在の指標）を含むデータセット、または1つ以上のマーカーの実際の遺

10

20

30

40

50

伝子型は、乳癌に関連すると本願発明者によって示された、ある多型マーカーのあるリスクのある対立遺伝子の存在または不存在についてクエリーされ得る。乳癌の増加したリスクに関連する変異体（例えば、マーカー対立遺伝子）についての陽性の結果は、本願明細書で示されたように、データセットの由来である個体が乳癌の増加した感受性（増加したリスク）にあることを示す。

【0181】

本発明のある実施態様では、多型マーカーは、多型マーカーの遺伝子型のデータを、多型性の少なくとも1つの対立遺伝子と乳癌との間の相関データを含む検査表などのデータベースへ照会することによって、乳癌に関連付けられる。いくつかの実施態様では、その表は1つの多型性についての相関を含む。他の実施態様では、表は多数の多型性についての相関を含む。両方のシナリオで、マーカーと乳癌との間の相関の指標を与える検査表へ照会することによって、乳癌のリスク、または乳癌に対する感受性が試料の由来である個体において同定され得る。いくつかの実施態様では、相関は統計的測定として報告される。統計的測定はリスク判定基準（相対リスク（RR）、絶対的なりスク（AR）またはオッズ比（OR）など）として報告されてもよい。

10

【0182】

リスクマーカーは、単独でまたは組合せて、リスクアセスメントおよび診断目的のために有用であり得る。本願明細書に記載されたマーカーに基づく疾病のリスクアセスメントの結果はまた、全体のリスクを確立するために、当該疾病についての他の遺伝子マーカーまたは危険因子についてのデータと合わされ得る。従って、個々のマーカーによるリスクの増加が相対的に少ない、例えば10～30%の程度である場合でさえ、その関連は、他のリスクマーカーと合わされた場合には、重要な影響をもたらし得る。従って、相対的に共通の変異は、全体のリスクに対して重要な寄与をし得る（集団に起因するリスクが高い）か、またはマーカーの組み合わせは、当該マーカーの合わせたリスクに基づいて、当該疾病の発生の有意な合わせたリスクにある個体群を定義するために使用され得る。例えば、rs2005154、rs2184380、rs2224696、rs2242503、rs12291026、rs999737、rs9956546、rs1191292、およびrs6001954のうちのいずれか1つ、またはこれらの組み合わせについての遺伝子型の結果に基づいて、合わせたリスクが評価され得る。このような組み合わせは、乳癌についての他の感受性マーカー、例えば染色体5p12上のマーカーおよび染色体10q26上のマーカー、例えば、マーカーrs10941679およびマーカーrs1219648をも含み得る（Stacey, S.N.ら、Nat Genet 40:703-6(2008)）。あるいは、これらのマーカーのうちのいずれか1つとLDのマーカーが評価され得る。乳癌のリスクを与えることが公知の他のマーカーもまた、染色体2q14上のマーカー（例えば、マーカーrs4848543またはそれらと連鎖不平衡のマーカー）、2q35上のマーカー（例えば、マーカーrs13387042、またはそれらと連鎖不平衡のマーカー）、および染色体16上のマーカー（例えば、マーカーrs3803662、またはそれらと連鎖不平衡のマーカー）などの本願明細書に記載されたマーカーと一緒に、評価され得る（Stacey, S.N.ら Nat Genet 39:865-9(2007)）。

20

30

40

【0183】

従って、本発明のある実施態様では、複数の変異体（マーカーおよび/またはハプロタイプ）は、全体のリスクアセスメントのために用いられる。これらの変異体は、1つの実施態様では、本願明細書に開示された変異体から選択される。他の実施態様は、乳癌に対する感受性の診断に有用であることが知られた他の変異体と組み合わせた本発明の変異体の使用を含む。このような実施態様では、多数のマーカーおよび/またはハプロタイプの遺伝子型の状態は、関連した変異体の集団における頻度、または臨床的に健康な被験者（年齢をマッチさせた被験者および性別をマッチさせた被験者など）の変異体の頻度と比較した、個体、および個体の状態について判定される。当該技術分野で公知の方法（本願明細書に記載されたものなどの多変量解析もしくは共同リスク解析、または当業者に知られ

50

た他の方法など)は、その後、複数の部位の遺伝子型の状態に基づいて与えられた全体のリスクの判定のために使用されてもよい。このような分析に基づくリスクの評価は、本願明細書記載のように、その後、本発明の方法、使用およびキットで使用されてもよい。

【0184】

一般的な意味では、本願明細書に記載される方法およびキットは、いずれの供給源およびいずれの個体)由来の核酸物質(DNAもしくはRNA)を含む試料から、またはこのような試料由来の遺伝子型または配列データから利用され得る。好ましい実施態様では、個体はヒト個体である。当該個体は、成人、子供、または胎児であり得る。当該核酸供給源は、核酸物質を含むいずれの試料であってもよく、例としては生物学的試料、またはそれらに由来する核酸物質を含む試料が挙げられる。本発明はまた、標的集団のメンバーである個体のマーカーおよび/またはハプロタイプの評価について提供する。このような標的集団は、1つの実施態様では、他の遺伝要因、バイオマーカー、生物物理的なパラメータ(例えば、体重、BMD、血圧)、または一般的な健康および/または生活習慣のパラメータ(例えば、癌の履歴、乳癌の履歴、疾病の以前の診断、癌の家族歴、乳癌の家族歴)に基づく、疾病を発生させるリスクがある個体の集団または群である。

【0185】

本発明は、特定の年齢亜群(40歳超、45歳超、または50、55、60、65、70、75、80、または85歳超など)からの個体を含む実施態様を提供する。本発明の他の実施態様は、他の年齢群(85歳未満の個体など、80歳未満、75歳未満、または70、65、60、55、50、45、40、35、または30歳未満など)に関する。他の実施態様は、乳癌の発生時の年齢が上述の年齢幅のいずれかである個体に関する。年齢幅は、ある実施態様では、発生時の年齢が45歳超だが60歳未満などであることに関連していてもよいことも考慮される。しかし、他の年齢幅もまた考慮され、上記の年齢の値によってひとまとめにされる全ての年齢幅を含む。本発明は、さらに、いずれかの性別(男性または女性)の個体に関する。

【0186】

アイスランド人の集団は、北欧の系統のコーカサス人の集団である。アイスランド人の集団の遺伝的連鎖および関連の結果を報告する数多くの研究が、直近の数年間に発表されてきている。これらの研究の多くは、特定の疾病に関連しているとしてアイスランド人の集団で最初に同定された変異の再現を、他の集団でも示した(Sulem, P.ら Nat Genet 2009年5月17日(印刷物に先駆けた電子出版); Rafnar, T.ら Nat Genet 41:221-7(2009); Gretarsdottir, S.ら Ann Neurol 64:402-9(2008); Stacey, S.N.ら Nat Genet 40:1313-18(2008); Gudbjartsson, D.F.ら Nat Genet 40:886-91(2008); Styrkarsdottir, U.ら N Engl J Med 358:2355-65(2008); Thorsteinsdottir, T.ら Nature 452:638-42(2008); Gudmundsson, J.ら Nat Genet 40:281-3(2008); Stacey, S.N.ら Nat Genet 39:865-69(2007); Helgadóttir, A.ら Science 316:1491-93(2007); Steinthorsdottir, V.ら Nat Genet 39:770-75(2007); Gudmundsson, J.ら Nat Genet 39:631-37(2007); Frayling, T.M. Nature Reviews Genet 8:657-662(2007); Amundadóttir, L.T.ら Nat Genet 38:652-58(2006); Grant, S.F.ら Nat Genet 38:320-23(2006)。従って、アイスランド人の集団における遺伝的発見は、一般的に、他の集団(アフリカ人の集団およびアジア人の集団を含む)において再現されている。

【0187】

乳癌のリスクに関連すると本願明細書に記載されるマーカーは、他のヒト集団において同様の関連性を示すと考えられている。従って、ヒト個体集団を含む特定の実施態様もまた考慮され、本発明の範囲内である。このような実施態様は、コーカサス人の集団、ヨーロッパ人の集団、アメリカ人の集団、ユーラシア人の集団、アジア人の集団、中央/南アジア人の集団、東アジア人の集団、中東人の集団、アフリカ人の集団、スペイン系アメリカ人の集団、およびオセアニア人の集団を含む（しかしこれらに限定されない）1つ以上のヒト集団の出身であるヒト被験者に関する。ヨーロッパ人の集団は、スウェーデン人、ノルウェー人、フィンランド人、ロシア人、デンマーク人、アイスランド人、アイルランド人、ケルト人、英国人、スコットランド人、オランダ人、ベルギー人、フランス人、ドイツ人、スペイン人、ポルトガル人、イタリア人、ポーランド人、ブルガリア人、スラブ人、セルビア人、ボスニア人、チェコ人、ギリシャ人およびトルコ人集団を含むが、これらに限定されない。1つの実施態様では、本発明は、コーカサス人の血統の個体に関する。

10

20

30

40

50

【0188】

個々の被験者における人種の寄与はまた、遺伝的分析によって判定されてもよい。系統の遺伝的分析は、非連鎖のマイクロサテライトマーカー（Smithら、Am J Hum Genet 74、1001-13（2004）に提示されるものなど）を使用して行ってもよい。

【0189】

ある実施態様では、本発明は、上述のように、特定の集団で同定されたマーカーおよび/またはハプロタイプに関する。当業者は、連鎖不平衡（LD）の判定は、異なる集団に適用した場合は異なる結果を与えてもよいことを理解するだろう。このことは、特定のゲノム領域におけるLDの相違を引き起こしたかも知れない、異なるヒト集団の異なる集団歴および異なる選択圧による。あるマーカー（例えば、SNPマーカー）は、1つの集団では多型だが、別の集団ではそうではないこともまた、当業者に周知である。しかし、当業者は、いずれかの所定のヒト集団で本発明を実行するために、利用可能であり、本願明細書で教示された方法を適用するだろう。このことは、特定の集団内で最も強い関連を与えるこれらのマーカーを同定するための、本発明のLD領域における多型マーカーの評価を含んでいてもよい。従って、本発明のリスクのある変異体は、異なるハプロタイプの背景に、および様々なヒト集団において異なる頻度で属していてもよい。しかし、当該技術分野で公知の方法および本発明のマーカーを利用し、本発明は、所定のヒト集団いずれにおいても実行され得る。

【0190】

（乳癌の遺伝するリスクを予測するモデル）

乳癌のリスクアセスメントの目的は、全ての女性に対して、高いリスクの女性においては生存期間および生活の質の増加のため、一方、低いリスクの女性においては費用、不要な介入および不安の最小化のための、個別化された医学的管理戦略の開発の合理的なフレームワークを提供することである。リスク予測モデルは、先天性リスク特性の所定のセット（例えば、家族歴、術前良性乳房病変、以前の乳腺腫瘍）を有する個体の乳癌のリスクを推定しようと試みる。診療において最も一般的に使用される乳癌のリスクアセスメントモデルは、家族歴を考慮して、遺伝する危険因子を評価する。リスク評価は、以前に乳癌と診断された1つ以上の近親者を有する個体の増加したリスクの観察に基づく。これらは、複雑な系統構造を考慮しない。これらのモデルは、乳癌になりやすい変異を有する遺伝子の保有者と非保有者とを区別できないというさらなる不利益を有する。

【0191】

さらに洗練されたリスクモデルは、特定の家族歴を扱うためのより良い機構を有し、BRCA1変異およびBRCA2変異の保有者状態を考慮する能力を有する。例えば、疾病発生率についての乳房分析および卵巣分析ならびに保有者推定アルゴリズム（BOADICEA）（Antoniouら、2004）は、系統分析プログラムMENDELによる個体系統構造に基づく家族歴を考慮している。公知のBRCA1およびBRCA2の状態

に関する情報もまた、考慮される。現在使用されている B O A D I C E A および全ての他の乳癌のリスクモデルの主な制限は、これらは他の素因遺伝子からの遺伝子型の情報を、取り込まないことである。現在のモデルは、リスクの非 B R C A 遺伝的判定要因についての知識の欠如を補うための代替物として作用する、家族歴に強く依存している。従って、利用可能なモデルは、疾病の家族歴が分かっている状況に制限される。より低い浸透率の乳癌素因遺伝子は、集団において相対的に一般的であるかも知れず、B R C A 1 および B R C A 2 遺伝子のように、家族性クラスター形成を促進するこのような強い傾向を示さないかも知れない。素因対立遺伝子について相対的に高い遺伝的負荷を有する患者は、ほとんど、または全く疾病の家族歴を示さないかも知れない。従って、遺伝子に基づいた試験を通じて、直接的に得られた遺伝する感受性データを取り込むモデルの構築についてニーズがある。モデルをより正確にすることに加えて、これは、家族歴パラメータへの依存性を減少させ、リスク特性を、家族歴がそれほど鍵因子とならないより幅広いリスクのある集団へ拡張することを支援するだろう。

10

20

30

40

50

【0192】

(改善された遺伝的危険率モデルの乳癌一次予防の臨床的管理への統合)

臨床的一次予防の選択肢は、現在、化学予防(またはホルモン)治療および予防的手術に分類され得る。高いリスクと同定される患者は、長期コースの化学予防的治療を処方され得る。この構想は、循環器学の分野で広く受け入れられているが、臨床的腫瘍学では、今ようやく、影響を与え始めている。最も広く使用されている腫瘍学の化学予防は、タモキシフェン(選択的エストロゲン受容体調節物質(SERM))である。乳癌再発に対して向けられたアジュバント治療方法として最初に使用され、タモキシフェンは今、乳癌の予防的薬剤としての有効性が証明されている[Cuzickら、(2003)、Lancet、361、296-300][Martinoら、(2004)、Oncologist、9、116-25]。FDAは、ある高いリスクの女性の化学予防薬剤として、タモキシフェンの使用を承認した。

【0193】

残念ながら、長期間のタモキシフェンの使用は、子宮内膜癌のリスクを約2.5倍増加させ、静脈血栓症のリスクを約2.0倍増加させる。肺塞栓症、脳卒中、および白内障のリスクもまた、増加する[Cuzickら、(2003)、Lancet、361、296-300]。従って、乳癌発生率を減少させるタモキシフェン使用の利益は、全体的な死亡率における対応する減少に簡単に解釈されないかも知れない。ラロキシフェンと呼ばれる別のSERMは、予防的方法においてより効果的であるかも知れず、子宮内膜の癌についての同様のリスクはもたらさない。しかし、血栓症のリスクは、ラロキシフェンで長期間治療された患者において、依然として上昇する[Cuzickら、(2003)、Lancet、361、296-300; Martinoら、(2004)、Oncologist、9、116-25]。さらに、タモキシフェンおよびラロキシフェンの両方が、それらに関連する生活の質の問題点を有する。化学予防的方法におけるSERM治療方法の合理的なリスク:利益分析を行うために、最も乳癌のリスクのある個体の同定について臨床的なニーズがある。乳癌のリスクの実質的な部分が遺伝的であるとすると、当該状況での個体のリスクを定量化するための遺伝的試験について明確な臨床的ニーズがある。利用可能になるかも知れない将来の癌の化学予防的治療(アロマターゼ阻害剤など)のいずれかから生じる同様の問題点について予測し得る。さらに、化学予防的治療が安全になるにつれ、遺伝的な傾向があるが、B R C A 1 および 2 変異の保有者に関連する、大量の上昇したリスクは有さない患者を同定するニーズが増加する。

【0194】

乳癌について高いリスクを有すると同定された患者は、予防的手術(両側乳房切除術もしくは卵巣摘出術のいずれか、または両方)について考慮される。明らかにこのように強烈な治療は、極度に高いリスクが認知された患者に対してのみ推奨される。実際に、このようなリスクは、現在、B R C A 1、B R C A 2 または、まれな乳癌素因の症候群に関連することが知られる遺伝子(リー・フラウメニ症候群におけるp53、カウデン症候群に

おける P T E N など) の変異を有する個体のみ同定され得る。

【 0 1 9 5 】

B R C A 1 変異および B R C A 2 変異の浸透率の評価は、集団に基づいた推定に由来している場合に比べて、複数の症例の家族に由来している場合に高い傾向にある。これは、異なる変異を有する家族は、乳癌に対して異なる浸透率を呈するからである (例えば、[Thorlaciusら、(1997)、Am J Hum Genet、60、1079-84]参照)。当該変異に寄与する主な因子の1つは、B R C A 1 変異および B R C A 2 変異の浸透率を修飾する効果を有する、今までのところ未知の素因遺伝子の作用である。従って、B R C A 1 または B R C A 2 遺伝子の変異を有する個体の絶対的なリスクは、修飾遺伝子の存在および作用についての知識がない場合では、正確に数量化されない。B R C A 1 および B R C A 2 保有者の治療上の選択肢は、苛酷であり得るため、当該状況では、それぞれの B R C A 保有者のリスクについて、できるだけ最大の正確性で定量化することが重要である。従って、B R C A 1 および B R C A 2 保有者の乳癌の浸透率を修飾する効果を有する素因遺伝子を同定し、およびこれらの遺伝子に基づく改善されたりリスクアセスメントモデルの開発についてニーズがある。

10

【 0 1 9 6 】

さらに、恐らく強い乳癌の家族歴が原因である、乳癌について非常に高いリスクが認知されたが、公知の素因遺伝子の変異は同定され得ないという個体がいる。高い浸透率の素因遺伝子を受け継いでいるかどうかを発見するために個体を試験することはできないため、このような症例において予防的手術の考慮は困難である。従って、個体のリスクは、正確に評価され得ない。従って、発見されていないままのいずれかの高い浸透率の素因遺伝子の同定、および一次予防戦略で使用するための関連する遺伝的試験の開発について、明確な臨床的ニーズがある。このような遺伝子は、例えば、乳癌のリスクと関連があると本願明細書で開示された遺伝子であってもよい。乳癌のリスクと関連があると本願明細書で示された変異体は極めて一般的であり、相対的に低い乳癌のリスクを与えるが、より高いリスク変異体が1つ以上のこれらの遺伝子の中に存在する可能性が十分にある。従って、P A X 5、T U B、S E R P I N H 1、R A D 5 1 L 1、F H O D 3 および / または T N R C 6 B 遺伝子のうちの1つ以上の中の、またはこれらと関連する高いリスクの遺伝的変異は、個体が乳癌について高いリスクの (かつ高い浸透率の) 遺伝的要因の保有者であるかどうかを判定するのに有用であり得ると考えられる。

20

30

【 0 1 9 7 】

(早期診断)

大部分の西洋諸国における乳癌の臨床的スクリーニングは、定期的な臨床的な乳房の診察 (C B E) および X 線マンモグラフィからなる。C B E が、良好な乳房 X 線撮影によるスクリーニングプログラムに関連して使用されても、ほとんど利益が付加されないことを示す優れたエビデンスがある。イギリスでは、50~70歳の女性は、3年ごとにスクリーニングマンモグラフィを受けるように勧められる。米国における状況は医療提供者によって異なるが、しかし、アメリカ癌協会は、40歳からは年1回の乳房 X 線撮影によるスクリーニングを推奨している。乳房 X 線撮影によるスクリーニングは、50歳を超えたスクリーニングされた女性の間の死亡率の減少について証明された有効性を有する。

40

【 0 1 9 8 】

遺伝子検査が、既存の乳房 X 線撮影によるスクリーニングプログラムへのアクセスを減少させる手段として採用されることは、絶対にありそうにない。しかし、乳房 X 線撮影によるスクリーニングは、欠点がないわけではなく、遺伝子検査は、増強されたスクリーニングプログラムのための人々を選別するのに使用されるべきであると考えられる。乳房 X 線撮影によるスクリーニングの弱点の1つは、これまでは、50歳未満のスクリーニングされた女性の改善された生存について、著しい効果を示す可能性がないことである。

【 0 1 9 9 】

マンモグラフィが50歳未満の女性において効果が少ない1つの理由は、乳房組織の密度は、より若い女性において高く、腫瘍の乳房 X 線撮影の検出をより困難にしているこ

50

とかも知れない。しかし、遺伝的な傾向がある個体の乳癌は、若年齢群で生じる傾向があり、高い乳房密度と乳癌のリスクとの間の明確な関連がある。従って、最も高いリスク群において最適でなく行われる技術によって管理され得るため、高い素因を有する個体の乳房X線撮影によるスクリーニングの単純な増加に問題がある。近年の研究では、乳房X線撮影によるスクリーニングよりも、造影磁気共鳴画像（CE-MRI）が、当該高いリスク群において、より感度が高く、より早期の段階で腫瘍を検出することが示された [Warnerら、(2004)、Jama、292、1317-25；Leachら、(2005)、Lancet、365、1769-78]。CE-MRI戦略は、特に、ルーチンのX線マンモグラフィと組み合わせて使用される場合に、よく役立つ [Leachら、(2005)、Lancet、365、1769-78]。CE-MRIは、高い費用がかかる専門家のセンターを必要とするため、50歳未満のスクリーニングは、最も高いリスクを有する個体に制限されなければならない。現在のCE-MRI試験は、エントリーをBRCA1、BRCA2もしくはp53変異ならびに疾病の非常に強い家族歴を有する個体に制限している。当該スクリーニング様式をより幅広い範囲の高いリスクの患者に拡張することは、遺伝子に基づいたリスクプロファイリングの手段の供給によって、大いに補助される。

10

【0200】

早期発生の乳癌および遺伝的に傾向がある女性に生じる癌は、より高齢の、遺伝的な傾向が強くない女性の癌よりも増殖が早いという概念を支持する優れたエビデンスがある。これは、より若い女性における高率の中間期癌の観察に由来し、つまり、よくスクリーニングされた集団においてスクリーニング来診の間隔で生じる癌は、より若い女性の間で高い。従って、スクリーニング間隔は、どんな方法によるものであれ、より若い女性については減少されるべきであることの示唆がある。ここに、より高価な方法論を用いたより頻繁なスクリーニングは、乳癌の全体的な割合が比較的低い年齢群に対して必要とされると思われるという逆説がある。ここに、早期の疾病を生じる最も強い遺伝的な傾向がある若い個体を同定し、彼らをより高価かつ広範なスクリーニング管理に導く、明確な臨床的ニーズがある。乳癌のリスクを与えると本願明細書で開示された変異体は、乳癌の発生について特に高いリスクにある個体の同定に有用であり得る。このような個体は、癌の早期の同定の尤度を最大にするための、初期の、かつ活発なスクリーニング計画からの恩恵を最も得られる可能性がある。

20

30

【0201】

(治療)

現在、原発性乳癌は、手術、アジュバント化学療法、放射線療法、その後の長期間のホルモン療法によって治療される。しばしば、3つまたは4つの治療の組み合わせが使用される。

【0202】

疾病の同様の段階の乳癌患者は、全体的な治療成果における広い変異をもたらす、アジュバント化学療法に対する非常に異なる反応を有し得る。総合指針（ザンクトガレン（St Gallen）およびNIHの判断基準）が、アジュバント化学療法の治療に対する乳癌患者の適格性を判定するために開発された。しかし、転移の最も強い臨床的前兆および組織学的前兆さえも、正確に乳腺腫瘍の臨床的反応を予測するには至っていない [Goldhirschら、(1998)、J Natl Cancer Inst、90、1601-8；Eifelら、(2001)、J Natl Cancer Inst、93、979-89]。化学療法またはホルモン療法は、転移のリスクを約1/3しか減少させないが、しかし、70～80%の当該治療を受けている患者は、それなくして生存している。従って、大部分の乳癌患者は、現在、無効または不要な治療のいずれかを提供されている。臨床医が、最善の利益を受ける者に対して治療をより適切に調整することができる予後診断の基準の開発の改善について、明確な臨床的ニーズがある。遺伝的素因についての個体の鑑定は、彼らの治療成果に関連する情報を明らかにし、それによって合理的な治療計画を助けるかも知れないと予期するのは合理的である。乳癌のリスクを与える本発

40

50

明のマーカーは、この状況において有用であると考えられる。

【0203】

いくつかの以前の研究は、当該概念を例証する。BRCA変異保有者の乳癌患者は、アジュバント化学治療で治療された場合、より良い臨床的反応の割合および生存を示すことが明らかとなった [Chappuisら、(2002)、J Med Genet、39、608-10; Goffinら、(2003)、Cancer、97、527-36]。BRCA変異保有者は、卵巣癌の白金化学療法に対して、非保有者よりも、改善された反応を示した [Cassら、(2003)、Cancer、97、2187-95]。同様の考察は、関連する遺伝子が未知である遺伝的な傾向がある患者に対しても当てはまるかも知れない。例えば、浸潤性小葉乳癌 (ILBC) は、強い家族性要素を有することが知られるが、関連する遺伝的変異は未だ同定されていない。ILBCを有する患者は、一般的な化学療法管理に対して乏しい反応を示す [Mathieura、(2004)、Eur J Cancer、40、342-51]。

10

【0204】

遺伝的素因モデルは、治療戦略の個別化を補助するだけでなく、これらの戦略の設計において不可欠な役割を果たすかも知れない。例えば、BRCA1およびBRCA2変異の腫瘍細胞は、ポリ (ADPリボース) ポリメラーゼ (PARP) 阻害剤に対し、その欠損したDNA修復経路の結果、深刻な感受性があることが見出された [Farmerら、(2005)、Nature、434、917-21]。これは、BRCA保有者である患者特異的な使用を目的とした、PARPを標的にした小分子薬剤の開発を刺激した。この例から、遺伝的素因の知識が、遺伝的危険率特性と組み合わせて使用される個別化された化学療法管理の開発を導く薬剤標的を同定するかも知れないことは明らかである。同様に、本発明のマーカーは、例えば、PAX5、TUB、SERPINH1、RAD51L1、FHOD3および/またはTNRC6B遺伝子のうち1つ以上を標的にする新規な薬物の同定の助けになるかも知れない。

20

【0205】

癌化学療法は、周知の、通常の組織 (特に、高増殖性ヘモポエティック (hemopoetic) 細胞区画および腸上皮細胞区画) に対する用量制限副作用を有する。遺伝に基づいた個体差が、細胞障害性薬剤に対する通常の組織の感受性に存在することが予測され得る。これらの因子についての理解は、合理的な治療計画および通常の組織を化学療法の副作用から保護するように設計された薬剤開発を助けるかも知れない。

30

【0206】

遺伝的特性はまた、改善された放射線療法のアプローチに寄与するかも知れない。標準的な放射線療法管理を受けている乳癌患者群内で、患者の一部は、通常は許容的な放射線量に対する有害反応を経験するだろう。急性反応は、紅斑、湿性落屑、浮腫および放射線気腫 (radiation pneumatitis) を含む。毛細血管拡張症、浮腫、肺線維症および乳腺線維症を含む、長期間の反応は、放射線療法後何年も生じるかも知れない。急性反応および長期間の反応は両方とも罹患率の大きな原因であり、致死的であり得る。1つの研究では、11%は深刻な有害反応を有する一方で、87%の患者は、放射線療法に対するいくつかの有害な副作用を有することが見出された (LENT/SOMA 段階3~4) [Hoellerら、(2003)、Int J Radiat Oncol Biol Phys、55、1013-8]。放射線療法に対する有害反応を経験する可能性は、主に通常の組織反応における恒常的な個体差によるものであり、これらは強い遺伝的要素を有するという疑いがある。いくつかの公知の乳癌素因遺伝子 (例えば、BRCA1、BRCA2、ATM) は、DNA二本鎖の切断修復経路に影響する。DNA二本鎖の切断は、放射線療法によって引き起こされた原発性細胞障害性の病変である。これは、これらの経路に属する遺伝子の変異を保有することを通じて乳癌に遺伝的な傾向がある個体はまた、放射線療法から過剰な通常組織の障害を受ける高いリスクを有するかも知れないという懸念を導く。例えば、PAX5、TUB、SERPINH1、RAD51L1、FHOD3および/またはTNRC6B遺伝子のうち1つ以上を通じて乳癌のリス

40

50

クを与えると本願明細書に記載された遺伝的変異体は、放射線療法に対する有害反応という特定のリスクを有する個体の同定に有用であるかも知れないと考えられる。

【0207】

集団における恒常的に放射線感受性な個体の存在は、有害反応の頻度を容認できるレベルに維持するため、大部分の患者集団に対する放射線療法の線量率は制限されなければならないことを意味する。従って、放射線療法に対する有害反応に対する上昇したリスクを有する個体を同定できる信頼できる試験について臨床的ニーズがある。このような試験は、相対的に放射線抵抗性な大部分の患者に対する放射線治療量の増大を許容する一方で、放射線感受性の個体に対する保守的なまたは別の治療を示すだろう。用量の増大は、乳癌患者を単純に放射線感受性分類、中間分類および放射線抵抗性分類ヘトリアージ方式に分類するための試験によって可能となり、これは、局所的な腫瘍管理について約35%の増加、および結果的な生存率の改善をもたらし得ることが推定される [Burnetら、(1996)、Clin Oncol (R Coll Radiol)、8、25-34]。

10

【0208】

電離放射線への曝露は、乳房における発癌に寄与する証明された因子である。 [DumitrescuおよびCotarla、(2005)、J Cell Mol Med、9、208-21]。公知の乳癌素因遺伝子は、放射線誘発のDNA損傷に対する細胞応答の経路の要素をコードする [NarodおよびFoulkes、(2004)、Nat Rev Cancer、4、665-76]。従って、続発性原発性乳腺腫瘍のリスクが、放射線療法の領域内で通常の組織に対する照射によって増加するかも知れないという懸念がある。放射線療法によるBRCA保有者についてのいずれの判定可能な増加したリスクは明らかではないが、しかし、続発性原発腫瘍についてのそのリスクは、すでに例外的に高い。続発性原発腫瘍のリスクは、放射線療法で治療された、ATMおよびCHEK2遺伝子における乳癌になりやすい対立遺伝子の保有者で増加することを示唆するエビデンスがある [Bernsteinら、(2004)、Breast Cancer Res、6、R199-214; Broeksら、(2004)、Breast Cancer Res Treat、83、91-3]。放射線療法による(および、恐らく、集中的な乳房X線撮影によるスクリーニングによる)続発性原発腫瘍のリスクは、治療計画段階に間の患者の正確な遺伝的危険率特性を得ることによって、より良く定められるだろうことが予期される。

20

30

【0209】

(再発予防)

ステージ1または2の乳癌と診断された患者の約30%は、それらの原発腫瘍の局所領域転移性再発または遠隔転移性再発のいずれかを経験するだろう。原発性乳癌を有していた患者はまた、乳房保存手術が行われた場合、対側乳房または同側乳房において、続発性原発腫瘍と診断される非常に増加したリスクを有する。再発予防は、再発予防または続発性原発腫瘍の発生の予防に使用される方法を意味する。現在使用される方法は、以下を含む。タモキシフェンもしくは別のSERM(単独またはアロマターゼ阻害剤と交互)による長期間の治療、リスクを減少する対側乳房の乳房切除術、ならびにリスクを減少する卵巣摘出術(家族性乳房癌および卵巣癌のリスクを有する患者)。タモキシフェンの使用に関する考察は上記で議論されている。リスクを減少する外科的選択肢について、費用:利益分析を知らせるために、できるだけリスクが数量化される必要があることは明らかである。

40

【0210】

乳癌についての公知の遺伝的素因を有する患者が、大部分の患者よりも悪化したことについていくつかの指標がある。CHEK2遺伝子1100delC変異を有する患者は、非保有者と比較して、推定2.8倍の遠隔転移についての増加したリスク、3.9倍の疾病再発の増加したリスクを有する [de Bockら、(2004)、J Med Genet、41、731-5]。BRCA1リンパ節転移陰性の腫瘍を有する患者は、BR

50

CA1変異を有さない同様の患者よりも、転移についてより大きなリスクを有する [Goffinら、(2003)、Cancer、97、527-36; Mollerら、(2002)、Int J Cancer、101、555-9; Eerolaら、(2001)、Int J Cancer、93、368-72]。従って、遺伝的特性は、局所的な再発および転移のリスクを評価するのに助けるために使用され得、それによって再発予防的な治療の選択を導く。本願明細書に記載された変異体に基づく遺伝的プロファイリングは、この状況において有用であるかも知れない。ある実施態様では、このようなプロファイリングは、本願明細書に記載された1つ以上の変異体に基づいていてもよい。他の実施態様では、このようなプロファイリングは、1つまたはいくつかの他の既知の、乳癌の遺伝的危険因子を含んでいてもよい。このような危険因子は、確立した高い浸透性の危険因子であってもよく、または、これらは前述の、1つ以上の、一般的な、より低い浸透率危険因子(例えば、染色体2q14上のマーカー(例えば、マーカーrs2005154、rs2184380、rs2224696、rs2242503、rs12291026、rs999737、rs9956546、rs11912922およびrs6001954、またはそれらと連鎖不平衡のマーカー、例えば表4のマーカー)であってもよい。

10

【0211】

通常は、原発性腫瘍と診断された患者は、一定の年間発生率(0.7%)の続発性原発腫瘍のリスクを有する [PetoおよびMack、(2000)、Nat Genet、26、411-4]。BRCA変異を有する患者は、大部分の乳癌患者よりも、40~60%の範囲の絶対的なリスクで、続発性原発腫瘍について著しく大きいリスクを有する [Easton、1999、Breast Cancer Res、1、14-7]。BRCA変異の保有者は、続発性原発腫瘍の非常に増加したリスクを有する [Staceyら、(2006)、PLoS Med、3、e217; Metcalfeら、(2004)、J Clin Oncol、22、2328-35]。CHEK2遺伝子変異の患者は、対側乳癌について、推定5.7倍の増加したリスクを有する [de Bockら、(2004)、J Med Genet、41、731-5]。BARD1 Cys557Ser変異の保有者は、続発性原発腫瘍と診断される傾向が2.7倍である [Staceyら、(2006)、PLoS Med、3、e217]。遺伝的危険率の特性は、患者における続発性原発腫瘍のリスクを評価するのに使用され得、予防措置がどの位激しいものであるべきかについての判定を知らせるだろう。

20

30

【0212】

(方法)

疾病のリスクアセスメントおよびリスクマネジメントについての方法が本願明細書に記載され、本発明により包含される。本発明はまた、治療薬に対する反応の可能性について個体を評価する方法、治療薬、核酸、ポリペプチドおよび抗体およびコンピューターで実施される機能の有効性を予測する方法を包含する。本願明細書中に与えられる種々の方法で使用するためのキットも、本発明によって包含される。

【0213】

(診断方法およびスクリーニング方法)

ある実施態様では、本発明は、乳癌の被験者もしくは乳癌に感受性がある被験者においてより頻繁に表れる遺伝子マーカーの特定の対立遺伝子を検出することによる、乳癌もしくは乳癌に対する感受性の診断方法、または診断の補助方法に関する。特定の実施態様では、本発明は、少なくとも1つの多型マーカー(例えば、本願明細書記載のマーカー)の少なくとも1つの対立遺伝子を検出することによる、乳癌に対する感受性の判定方法である。他の実施態様では、本発明は、少なくとも1つの多型マーカーの少なくとも1つの対立遺伝子を検出することによる、乳癌に対する感受性の診断方法に関する。本発明は、それによって特定のマーカーの特定の対立遺伝子またはハプロタイプの検出が乳癌に対する感受性を示す方法について記載する。このような予後診断のアッセイまたは予測的なアッセイはまた、乳癌に関連する症状の発生前の被験者の予防的な治療の判定に使用され得る。

40

50

【0214】

本発明は、いくつかの実施態様では、診断の臨床的適用の方法、例えば、医療専門家によって行われる診断に関する。他の実施態様では、本発明は、素人によって行われる感受性の診断方法または判定方法に関する。当該素人は、遺伝子型の同定サービスの顧客であり得る。当該素人はまた、個体（すなわち、当該顧客）の遺伝子型の状態に基づく、特定の形質または疾病についての遺伝的危険因子に関するサービスを提供するための、当該個体から得たDNA試料上の遺伝子型分析を行う遺伝子型サービスの提供者であってもよい。遺伝子型の同定技術についての近年の技術的進歩（SNPマーカーの高処理の遺伝子型の同定（分子反転プローブアレイ技術（例えば、Affymetrix遺伝子チップ）など）、およびビーズアレイ技術（例えば、イルミナ（イルミナ）GoldenGateアッセイおよびInfiniumアッセイ）を含む）は、個体が、彼ら自身のゲノムが最大100万のSNPについて同時に、相対的に低い費用で、評価されることを可能にする。得られた遺伝子型の情報は、当該個体に利用可能となり、様々なSNPに関連する疾病または形質リスクについての情報（公的な文献および科学出版物からの情報）と比較され得る。従って、本願明細書記載の疾病に関連する対立遺伝子の診断の適用は、例えば、当該個体の遺伝子型のデータの分析を通じて当該個体、臨床試験の結果に基づいて医療従事者、または遺伝子型サービス提供者を含む第三者によって行われ得る。当該第三者は、特定の遺伝的危険因子（本願明細書に記載された遺伝子マーカーを含む）に関するサービスを提供するために当該顧客から得た遺伝子型情報を解釈するサービス提供者であってもよい。換言すれば、遺伝的危険率に基づく感受性の診断もしくは判定は、個体の遺伝子型の状態についての情報および特定の遺伝的危険因子（例えば、特定のSNP）によって与えられるリスクについての知識に基づいて、医療従事者、遺伝的カウンセラー、遺伝子型の同定サービスを提供する第三者、リスクアセスメントサービスを提供する第三者、または素人（例えば、当該個体）によってされ得る。本願明細書の文脈では、用語「診断」、「感受性を診断する」および「感受性を判定する」とは、利用可能な診断方法（上記記載のものを含む）のいずれをも指すことを意味する。

10

20

【0215】

ある実施態様では、個体から得たゲノムDNAを含む試料が集められる。このような試料は、例えば、さらに本願明細書に記載されるように、頬スワブ、唾液試料、血液試料、またはゲノムDNAを含む他の適した試料であり得る。次いで、当該ゲノムDNAは、当業者が利用できるいずれかの一般的な技術（高処理アレイ技術など）を使用して分析される。このような遺伝子型の同定から得た結果は、簡便なデータ保存ユニット（コンピューターデータベースを含むデータ保有物、データ保存ディスク、または他の簡便なデータ保存手段など）に保存される。ある実施態様では、当該コンピューターデータベースは、オブジェクトデータベース、リレーショナルデータベースまたはポストリレーショナルデータベースである。当該遺伝子型データは、引き続いて、特定のヒトの状態の感受性変異体であることが知られるある変異体（本願明細書に記載された遺伝的変異体など）の存在について分析される。遺伝子型データは、いずれの簡便なデータクエリ方法を使用して、データ保存ユニットから検索され得る。当該個体の特定の遺伝子型によって与えられるリスクの算出は、当該個体の遺伝子型を、当該遺伝子型について以前に測定されたリスク（例えば、相対リスク（RR）またはおよびオッズ比（OR）として表わされる）（例えば、特定の疾病または形質についてのリスクのある変異体のヘテロ接合性の保有者）と比較することに基づくことができる。当該個体について算出されたリスクは、性別および民族性がマッチされた平均的な集団と比較した、ヒトについての相対リスクまたはヒトの特定の遺伝子型の相対リスクであってもよい。当該平均的な集団のリスクは、参照集団から得た結果を使用して、異なる遺伝子型のリスクの加重平均として表わしてもよく、当該集団に対する遺伝子型群のリスクを算出するための適切な算出を行ってもよい。あるいは、個体のリスクは、特定の遺伝子型の比較、例えば、リスクのある対立遺伝子の非保有者と比較した、マーカーのリスクのある対立遺伝子のヘテロ接合性の保有者に基づく。集団平均の使用は、ある実施態様では、それが当該使用者にとって解釈しやすい基準、すなわち、当

30

40

50

のDNA配列を含むがまた、マーカーまたはハプロタイプと強いLD（連鎖不平衡）のDNA分節を含んでいてもよい。1つの実施態様では、このような分節は、（0.1超の r^2 値および/または $|D'| > 0.8$ によって判定されるような）当該マーカーまたはハプロタイプとLDにある分節を含む。

【0221】

1つの実施態様では、乳癌に対する感受性の診断は、ハイブリダイゼーション法を使用して達成され得る（Current Protocols in Molecular Biology、Ausubel、F.ら、編集、John Wiley & Sons参照、全ての補遺を含む）。ゲノムDNA、RNA、もしくはcDNAの、試験の被験者もしくは個体から入手した生物学的試料（「試験試料」）は、乳癌を有する疑いのある被験者、乳癌に対する感受性がある疑いがある被験者、もしくは乳癌に遺伝的な傾向がある疑いがある被験者（「試験被験者」）から得られる。被験者は、成人、子供、または胎児であり得る。試験試料は、ゲノムDNAを含むいずれの供給源（血液試料、羊水試料、脳脊髄液試料、または皮膚、筋肉、頬側粘膜もしくは結膜粘膜、胎盤、消化管または他の器官などから得た組織試料）からでもあり得る。胎児の細胞または組織から得たDNAの試験試料は、適切な方法（羊水穿刺または絨毛採取などによる）によって得られうる。次いで、DNA、RNA、またはcDNA試料を試験した。特異的なマーカー対立遺伝子の存在は、特定の対立遺伝子に特異的な核酸プローブの配列特異的なハイブリダイゼーションによって示され得る。より多くの特定のマーカー対立遺伝子または特定のハプロタイプの存在は、それぞれ特定の対立遺伝子に特異的ないくつかの配列特異的な核酸プローブを使用することによって示され得る。1つの実施態様では、ハプロタイプは、特異的なハプロタイプに特異的な単一の核酸プローブ（すなわち、当該ハプロタイプの特徴を示す特定のマーカー対立遺伝子を含むDNA鎖に特異的にハイブリダイズする）によって示され得る。配列特異的なプローブは、ゲノムDNA、RNA、またはcDNAにハイブリダイズするように向けられ得る。「核酸プローブ」は、本願明細書で使用される場合、相補的配列にハイブリダイズするDNAプローブまたはRNAプローブであり得る。当業者であれば、配列特異的なハイブリダイゼーションが、特定の対立遺伝子が試験試料のゲノム配列に存在する場合のみに生じるように、このようなプローブを設計する方法を知っているだろう。本発明はまた、いずれの簡便な遺伝子型の同定方法（特定の多型マーカーの遺伝子型の同定のための市販の技術および方法を含む）を使用して、実行に移すことができる。

【0222】

乳癌に対する感受性を判定するため、ハイブリダイゼーション試料は、試験試料（ゲノムDNA試料など）を、少なくとも1つの核酸プローブに接触させることで形成され得る。mRNAまたはゲノムDNAを検出するプローブの非限定的な例は、本願明細書記載のmRNAまたはゲノムDNA配列にハイブリダイズできる標識された核酸プローブである。核酸プローブは、例えば、全長の核酸分子、またはその部分（長さが少なくとも15、30、50、100、250または500ヌクレオチドの、適切なmRNAまたはゲノムDNAにストリンジェントな条件下で特異的にハイブリダイズするのに十分なオリゴヌクレオチドなど）であり得る。ある実施態様では、当該オリゴヌクレオチドは、長さが約15～約100ヌクレオチドである。ある他の実施態様では、当該オリゴヌクレオチドは、長さが約20～約50ヌクレオチドである。例えば、核酸プローブは、本願明細書に明記されるLDブロックC09、LDブロックC10A、LDブロック10B、LDブロックC11A、LDブロックC11B、LDブロックC14、LDブロックC18、LDブロックC22A、およびLDブロックC22Bのうちのいずれか1つの全てもしくは一部のヌクレオチド配列を含み得；あるいは、当該核酸プローブは、本願明細書中の表1および4（配列番号：1～562）に記載されるマーカーを含むヌクレオチド配列の全てもしくは一部、または本願明細書に記載されたマーカーの少なくとも1つの対立遺伝子、もしくは本願明細書に記載された少なくとも1つのハプロタイプを任意に含む本願明細書記載のPAX5、TUB、SERPINH1、RAD51L1、FHOD3およびTNRC6B遺伝子のいずれか1つを含むヌクレオチド配列を含むことができ、または、当該プローブ

はこのような配列の相補的配列であり得る。特定の実施態様では、当該核酸プローブは、表1、2、3および4（配列番号：1～562）のいずれか1つに記載されたマーカを含むヌクレオチド配列一部、あるいは、本願明細書に記載されたマーカの少なくとも1つの対立遺伝子、または1つの多型マーカの少なくとも1つの対立遺伝子または本願明細書に記載された少なくとも1つの多型マーカを含むハプロタイプを任意に含む本願明細書記載のPAX5、TUB、SERPINH1、RAD51L1、FHOD3およびTNRC6B遺伝子のいずれか1つもしくはその断片を含むヌクレオチド配列であるか、またはプローブは、このような配列の相補的配列であり得る。本発明の診断のアッセイの使用のための他の適したプローブは、本願明細書に記載される。ハイブリダイゼーションは、当業者に周知の方法（例えば、Current Protocols in Molecular Biology、Ausubel, F.ら、編集、John Wiley & Sonsを参照、全ての補遺を含む）によって行われ得る。1つの実施態様では、ハイブリダイゼーションは、特異的なハイブリダイゼーション、すなわち、ミスマッチのないハイブリダイゼーション（正確なハイブリダイゼーション）を意味する。1つの実施態様では、特異的なハイブリダイゼーションのハイブリダイゼーション条件は、高度にストリンジェンシーである。

10

【0223】

特異的なハイブリダイゼーションは、存在する場合は、次いで、標準的な方法を使用して検出される。特異的なハイブリダイゼーションが、試験試料中で核酸プローブと核酸との間で生じた場合は、当該試料は当該核酸プローブに存在する当該ヌクレオチドに相補的な対立遺伝子を含む。この工程は、いずれの本発明のマーカ、または本発明のハプロタイプを構成するマーカについて繰り返すことができ、または複数のプローブは1より多くのマーカ対立遺伝子を同時に検出するために、同時に使用され得る。特定のハプロタイプの1より多くのマーカ対立遺伝子を含む単一のプローブ（例えば、特定のハプロタイプを構成する2、3、4、5または全てのマーカに相補的な対立遺伝子を含むプローブ）を設計することもまた、可能である。試料中の当該ハプロタイプの特定のマーカの検出は、当該試料の供給源が特定のハプロタイプ（例えば、ハプロタイプ）を有し、従って、乳癌に感受性があることを示す。

20

【0224】

1つの好ましい実施態様では、Kutyavinら（Nucleic Acid Res. 34:e128(2006)）によって記載されているように、その3'末端に蛍光部分または基、およびその5'末端に消光剤、およびエンハンサーオリゴヌクレオチドを含む、検出オリゴヌクレオチドプローブを利用する方法が使用される。蛍光部分は、Gig Harbor GreenまたはYakima Yellow、または他の適した蛍光部分であり得る。検出プローブは、検出されるSNP多型を含む短いヌクレオチド配列にハイブリダイズするように設計される。好ましくは、SNPは、末端残基から、当該検出プローブの3'末端から-6残基のどこでもよい。エンハンサーは、当該検出プローブに対して3'の鋳型DNAにハイブリダイズする短いオリゴヌクレオチドプローブである。当該検出プローブと当該エンハンサーヌクレオチドプローブとが当該鋳型に結合すると、両方の間に1つのヌクレオチドの間隙が存在するように、当該プローブが設計される。間隙は、エンドヌクレアーゼ（エンドヌクレアーゼIVなど）によって認識される合成の脱塩基の部位を作る。酵素は、色素を完全に相補的な検出プローブから切断するが、ミスマッチを含む検出プローブを切断することはできない。従って、放出された蛍光部分の蛍光を測定することで、検出プローブのヌクレオチド配列によって定義される特定の対立遺伝子の存在の評価を行い得る。

30

40

【0225】

検出プローブは、いずれの適したサイズでもあり得るが、好ましくは、当該プローブは相対的に短い。1つの実施態様では、当該プローブは長さが5～100ヌクレオチドである。別の実施態様では、当該プローブは、長さが10～50ヌクレオチドであり、別の実施態様では、当該プローブは、長さが12～30ヌクレオチドである。プローブの他の長

50

さも可能であり、当業者の技能の範囲内である。

【0226】

好ましい実施態様では、SNP多型を含む鋳型DNAは、検出の前にポリメラーゼ連鎖反応(PCR)によって増幅される。このような実施態様では、増幅されたDNAは、当該検出プローブおよび当該エンハンサープローブの鋳型として役立つ。

【0227】

検出プローブのある実施態様は、PCRによる鋳型の増幅に使用される当該エンハンサープローブおよび/または当該プライマーは、修飾塩基(修飾されたAおよび修飾されたGを含む)の使用を含む。修飾塩基の使用は、鋳型DNAへのヌクレオチド分子(プローブおよび/またはプライマー)の融解温度を調整することに対して有用であり得、例えば、低い割合のGもしくはC塩基を含む領域の融解温度を高めるため、その相補的Tに3つの水素結合を形成できる修飾されたAが使用され得、または、高い割合のGまたはC塩基を含む領域の融解温度を下げるため、例えば、二本鎖DNA分子中のその相補的C塩基に2つの水素結合のみを形成する修飾されたG塩基を使用することによって、有用であり得る。好ましい実施態様では、修飾塩基は、検出ヌクレオチドプローブの設計において使用される。当業者に知られたいずれの修飾塩基もこれらの方法において選択され得、適した塩基の選択は、本願明細書の教示および当業者に知られた市販の供給源から利用可能な公知の塩基に基づき、当業者の範囲内である。

【0228】

さらに、または、あるいは、ペプチド核酸(PNA)プローブは、本願明細書記載のハイブリダイゼーション法の核酸プローブに加えて、またはその代わりに使用され得る。PNAは、ペプチド様の無機の主鎖(有機の塩基(A、G、C、TまたはU)が、メチレンカルボニルリンカーを介してグリシンの窒素に結合している、N-(2-アミノエチル)グリシンユニットなどを有するDNA模倣体である(例えば、Nielsen, P.ら、Bioconjug. Chem. 5:3-7(1994)参照)。PNAプローブは、乳癌に関連する1つ以上のマーカー対立遺伝子またはハプロタイプを含む疑いのある試料中の分子に特異的にハイブリダイズするように設計され得る。本発明の1つの実施態様では、被験者から入手されたゲノムDNAを含む試験試料は収集され、ポリメラーゼ連鎖反応(PCR)が、本発明の1つ以上のマーカーまたはハプロタイプを含む断片を増幅するために使用される。本願明細書記載のように、乳癌に関連する特定のマーカー対立遺伝子またはハプロタイプの同定は、様々な方法(例えば、配列分析、制限消化による分析、特異的なハイブリダイゼーション、一本鎖高次構造多型分析(SSCP)、電気泳動分析など)を使用して達成され得る。別の実施態様では、診断は、定量的PCR(動力学的熱サイクル)を使用した発現分析によって達成される。当該技術は、例えば、市販の技術(TaqMan(登録商標)(アプライドバイオシステムズ社(Applied Biosystems)、カリフォルニア州、フォスターシティーなど)を利用し得る。当該技術は、乳癌に関連する核酸によってコードされるポリペプチドもしくはスプライシング変異体(単数または複数)の発現もしくは構成における変化の存在を評価できる。さらに、変異体(単数または複数)の発現は、物理的に数量化され得るか、または機能的に異なり得る。

【0229】

別の本発明の方法では、制限消化による分析は、対立遺伝子が参照配列に関連する制限部位の創造または除去をもたらす場合、特定の対立遺伝子の検出に使用され得る。制限断片長多型(RFLP)分析は、例えば、Current Protocols in Molecular Biology(上記参照)に記載のように行うことができる。関連するDNA断片の消化様式は、試料中の特定の対立遺伝子の存在または不存在を示す。

【0230】

配列分析はまた、特定の対立遺伝子またはハプロタイプの検出に使用され得る。従って、1つの実施態様では、特定のマーカー対立遺伝子またはハプロタイプの存在または不在の判定は、被験者または個体から入手したDNAまたはRNAの試験試料の配列分析を

10

20

30

40

50

含む。PCRまたは他の適切な方法は、乳癌に関連する核酸の一部の増幅に使用され得、次いで、特定の対立遺伝子の存在は、試料中のゲノムDNAの多形性部位（またはハプロタイプの複数の多形性部位）の配列決定によって直接的に検出され得る。

【0231】

別の実施態様では、被験者から得た標的の核酸配列の分節に相補的なオリゴヌクレオチドプローブのアレイは、多形性部位での特定の対立遺伝子を同定するために使用され得る。例えば、オリゴヌクレオチドアレイが使用され得る。オリゴヌクレオチドアレイは、典型的には、異なる公知の部位の基質の表面に結合される複数の異なるオリゴヌクレオチドプローブを含む。これらのアレイは、一般的に、フォトリソグラフィ法および固相オリゴヌクレオチド合成法の組み合わせを取り込んだ機械的合成法もしくは光指向合成法、または当業者に知られた他の方法（例えば、Bier, F. F. *ら* *Adv Biochem Eng Biotechnol* 109: 433-53 (2008); Hoheisel, J. D., *Nat Rev Genet* 7: 200-10 (2006); Fan, J. B., *ら* *Methods Enzymol* 410: 57-73 (2006); Raquousis, J. および Elvidge, G., *Expert Rev Mol Diagn* 6: 145-52 (2006); Mockler, T. C. *ら* *Genomics* 85: 1-15 (2005)、およびそれらの参考文献を参照（それぞれの全体の教示を参照によって本願明細書に援用する））を使用して作製され得る。多型性の検出のためのオリゴヌクレオチドアレイの調製および使用についての多くの付加的な記載は、例えば、米国特許第6,858,394号明細書、米国特許第6,429,027号明細書、米国特許第5,445,934号明細書、米国特許第5,700,637号明細書、米国特許第5,744,305号明細書、米国特許第5,945,334号明細書、米国特許第6,054,270号明細書、米国特許第6,300,063号明細書、米国特許第6,733,977号明細書、米国特許第7,364,858号明細書、欧州特許第619321号明細書、および欧州特許第373203号明細書に見出すことができ、これらの全体の教示は、参照によって本願明細書に組み込まれる。

【0232】

当業者に利用可能な核酸分析の他の方法は、乳癌に関連する多形性部位の特定の対立遺伝子を検出するために使用され得る。代表的な方法は、例えば、直接のマニュアル配列決定（ChurchおよびGilbert, *Proc. Natl. Acad. Sci. USA*, 81: 1991-1995 (1988); Sanger F. *ら*, *Proc. Natl. Acad. Sci. USA*, 74: 5463-5467 (1977); Beavis *ら*, 米国特許第5,288,644号明細書); 自動化した蛍光配列決定; 一本鎖高次構造多型性アッセイ (SSCP); クランプ変性ゲル電気泳動 (clamped denaturing gel electrophoresis) (CDGE); 変性剤濃度勾配ゲル電気泳動 (DGGE) (Sheffield V. *ら*, *Proc. Natl. Acad. Sci. USA*, 86: 232-236 (1989))、移動度シフトアッセイ (Orita M. *ら*, *Proc. Natl. Acad. Sci. USA*, 86: 2766-2770 (1989))、制限酵素解析 (Flavell R. *ら*, *Cell*, 15: 25-41 (1978); Geever R. *ら*, *Proc. Natl. Acad. Sci. USA*, 78: 5081-5085 (1981)); ヘテロ二本鎖分析; 化学的ミスマッチ切断 (chemical mismatch cleavage) (CMC) (Cotton R. *ら*, *Proc. Natl. Acad. Sci. USA*, 85: 4397-4401 (1985)); リボヌクレアーゼプロテクションアッセイ (Myers R. *ら*, *Science*, 230: 1242-1246 (1985)); ヌクレオチドのミスマッチを認識するポリペプチド (E. coli mutSタンパク質など)の使用; および対立遺伝子特異的なPCRを含む。

【0233】

本発明の別の実施態様では、乳癌に対する感受性の判定は、本発明の遺伝子マーカー（単数または複数）もしくはハプロタイプ（単数または複数）がポリペプチドの構成もしくは

10

20

30

40

50

は発現の変化をもたらすそれらの例における乳癌に関連した核酸によってコードされたポリペプチドの発現および/または構成を試験することでなされ得る。従って、乳癌に対する感受性の診断は、これらのポリペプチドの1つ、または、本発明の遺伝子マーカーもしくはハプロタイプが当該ポリペプチドの構成または発現の変化をもたらす例における、乳癌に関連した核酸（例えば、PAX5、TUB、SERPINH1、RAD51L1、FHOD3およびTNRC6B遺伝子のうち1つ以上）にコードされた別のポリペプチドの、発現および/または構成を試験することによってなされ得る。乳癌への関連を示す本願明細書記載のマーカーはまた、これらの近くの遺伝子の発現に影響を及ぼしているのかも知れない。遺伝子発現に影響する調節エレメントは、遺伝子のプロモーター領域から数十キロベースまたは数百キロベースでさえ離れて位置しているかも知れないことは周知である。本発明の少なくとも1つの多型マーカーの少なくとも1つの対立遺伝子の存在または不存在をアッセイすることによって、このような近くの遺伝子の発現レベルを評価できる。これらの遺伝子に影響する可能性のある機構は、例えば、転写への効果、RNAスプライシングへの効果、mRNAの別のスプライシングの形態の相対的な量の変化、RNAの安定性への効果、核から細胞質への輸送への効果、ならびに翻訳の効率および正確性への効果を含む。

10

【0234】

酵素結合免疫吸着検定法（ELISA）、ウエスタンブロット、免疫沈降および免疫蛍光を含む、様々な方法がタンパク質発現レベルの検出に使用され得る。被験者から得た試験試料は、発現の変化および/または乳癌に関連した核酸にコードされたポリペプチドの構成の変化の存在について評価される。乳癌に関連した核酸にコードされたポリペプチドの発現の変化は、例えば、定量的ポリペプチドの発現（すなわち、産生されたポリペプチドの量）の変化であり得る。乳癌に関連した核酸によってコードされたポリペプチドの構成の変化は、定性的なポリペプチドの発現（例えば、変異ポリペプチドの発現または異なるスプライシング変異体の発現）の変化である。1つの実施態様では、乳癌に対する感受性の診断は、乳癌に関連する核酸によってコードされた特定のスプライシング変異体、またはスプライシング変異体（例えば、PAX5、TUB、SERPINH1、RAD51L1、FHOD3およびTNRC6B遺伝子をコードする核酸）の特定の様式の検出によってなされる。

20

【0235】

このような変化の両方（定量的および定性的）もまた、存在し得る。ポリペプチドの発現または構成の「変化」は、本願明細書で使用される場合、対照試料のポリペプチドの発現または構成と比較した、試験試料の発現または構成の変化を意味する。対照試料は、試験試料に対応する試料（例えば、同様のタイプの細胞からもの）であり、乳癌に罹患していない被験者および/または乳癌に対する感受性を有さない被験者から得たものである。1つの実施態様では、対照試料は、本願明細書記載のマーカー対立遺伝子またはハプロタイプを有さない被験者から得たものである。同様に、対照試料と比較した、試験試料における1つ以上の異なるスプライシング変異体の存在、または試験試料における著しく異なる量の異なるスプライシング変異体の存在は、乳癌に対する感受性を示し得る。対照試料と比較した、試験試料におけるポリペプチドの発現または構成の変化は、対立遺伝子が対照試料における参照に関連するスプライシングの部位を変化させる例における、特定の対立遺伝子を示し得る。核酸によってコードされたポリペプチドの発現または構成を試験する様々な手段が当業者に知られ、かつ、使用され得、分光法、比色分析、電気泳動、等電点電気泳動、および、イムノブロッティング（例えば、Current Protocols in Molecular Biology、特に10章、上記参照）などの、イムノアッセイ（例えば、Davidら、米国特許第4,376,110号明細書）を含む。

30

40

【0236】

例えば、1つの実施態様では、乳癌に関連した核酸にコードされたポリペプチドに結合できる抗体（例えば、検出可能な標識を有する抗体）が使用され得る。抗体は、ポリクロ

50

ーナル抗体またはモノクローナル抗体であり得る。インタクトな抗体、またはその断片（例えば、Fv、Fab、Fab'、F(ab')₂）が使用され得る。用語「標識された」は、プローブまたは抗体に関し、検出可能な物質をプローブまたは抗体に結合することによって（すなわち、物理的に連結することによって）プローブまたは抗体を直接的に標識すること、および直接的に標識された別の試薬との反応性によってプローブまたは抗体を間接的に標識することを含むことが意図される。間接的な標識の例は、標識された2次抗体（例えば、蛍光標識された2次抗体）を使用した1次抗体の検出、および蛍光標識されたストレプトアビジンによって検出され得るようにビオチンによるDNAプローブの末端標識を含む。

【0237】

当該方法の1つの実施態様では、試験試料中の乳癌に関連する核酸（例えば、PAX5、TUB、SERPINH1、RAD51L1、FHOD3および/またはTNRC6B）にコードされたポリペプチドのレベルまたは量は、対照試料中のポリペプチドのレベルまたは量と比較される。対照試料中のポリペプチドのレベルまたは量よりもより高く、またはより低く、そのため、相違が統計的に有意な、試験試料中のポリペプチドのレベルまたは量は、核酸によってコードされたポリペプチドの発現の変化を示し、これは発現の相違を引き起こす原因となる特定の対立遺伝子またはハプロタイプであると診断される。あるいは、試験試料中のポリペプチドの構成は、対照試料中のポリペプチドの構成と比較される。別の実施態様では、ポリペプチドのレベルまたは量ならびに構成の両方は、試験試料および対照試料において評価され得る。

【0238】

別の実施態様では、乳癌に対する感受性の判定は、本発明の少なくとも1つのマーカーまたはハプロタイプを、さらにタンパク質に基づいたアッセイ、RNAに基づいたアッセイまたはDNAに基づいたアッセイと組み合わせて検出することによってなされる。

【0239】

（キット）

本発明の方法において有用なキットは、本願明細書記載の方法のいずれにおいても有用な要素を含み、例えば、核酸増幅のためのプライマー、ハイブリダイゼーションプローブ、制限酵素（例えば、RFLP分析のため）、対立遺伝子特異的なオリゴヌクレオチド、本願明細書記載の本発明の核酸（例えば、本発明の少なくとも1つの多型マーカーおよび/またはハプロタイプを含むゲノム分節）によってコードされた変化したポリペプチドに結合する抗体、または、本願明細書記載の本発明の核酸によってコードされた変化していない（ネイティブな）ポリペプチドに結合する抗体、乳癌に関連する核酸を増幅する手段、乳癌に関連する核酸の核酸配列を分析する手段、乳癌に関連する核酸によってコードされたポリペプチドのアミノ酸配列を分析する手段などを含む。当該キットは、例えば、必要なバッファー、本発明の核酸（例えば、本願明細書記載の1つ以上の多型マーカー）を増幅する核酸プライマー、ならびにこのようなプライマーおよび必要な酵素（例えば、DNAポリメラーゼ）を使用して増幅した断片の対立遺伝子特異的な検出のための試薬を含み得る。さらに、キットは、本発明の方法と組み合わせて使用されるアッセイのための試薬、例えば、乳癌の診断のアッセイと使用するための試薬を提供し得る。

【0240】

1つの実施態様では、本発明は、被験者における乳癌または乳癌の感受性の存在を検出するために、被験者から入手した試料をアッセイするためのキットであり、当該キットは、当該個体のゲノムにおける本発明の少なくとも1つの多型の少なくとも1つの対立遺伝子を選択的に検出するために必要な試薬を含む。特定の実施態様では、当該試薬は、本発明の少なくとも1つの多型を含む当該個体のゲノムの断片にハイブリダイズする少なくとも1つの近接するオリゴヌクレオチドを含む。別の実施態様では、当該試薬は、被験者から入手したゲノム分節の逆ストランドにハイブリダイズする少なくとも1対のオリゴヌクレオチドを含み、各オリゴヌクレオチドプライマー対は、少なくとも1つの多型を含む当該個体のゲノムの断片を選択的に増幅するように設計され、当該多型は、表1および4（

10

20

30

40

50

配列番号：1～562)に記載された多型、ならびにそれらと連鎖不平衡の多型マーカーからなる群から選択される。さらに別の実施態様では、当該断片は、サイズが少なくとも20塩基対である。このようなオリゴヌクレオチドまたは核酸(例えば、オリゴヌクレオチドプライマー)は、乳癌を示す多型に隣接する核酸配列(例えば、SNPまたはマイクロサテライト)の部分を使用して設計され得る。別の実施態様では、当該キットは、乳癌に関連する1つ以上の特定の多型マーカーまたはハプロタイプを対立遺伝子特異的に検出できる1つ以上の標識された核酸、および当該標識の検出のための試薬を含む。適した標識は、例えば、放射性同位元素、蛍光標識、酵素標識、酵素共因子標識、磁性標識、スピン標識、エピトープ標識を含む。

【0241】

特定の実施態様では、当該キットの試薬によって検出される当該多型マーカーまたはハプロタイプは、表1および4のマーカーからなる群から選択される1つ以上のマーカー、2つ以上のマーカー、3つ以上のマーカー、4つ以上のマーカーまたは5つ以上のマーカーを含む。別の実施態様では、検出されるマーカーまたはハプロタイプは、 0.2 より大きい r^2 値によって定義される、表1、2、3および4に記載されたマーカーからなるマーカー群の少なくとも1つと強い連鎖不平衡のマーカー群のうち少なくとも1つのマーカーを含む。別の実施態様では、検出されるマーカーまたはハプロタイプは、マーカーrs2005154、rs2184380、rs2224696、rs2242503、rs12291026、rs999737、rs9956546、rs11912922およびrs6001954からなるマーカー群から選択される少なくとも1つのマーカーを含む。

【0242】

好ましい実施態様では、SNP多型性を含むDNA鋳型は、検出の前にポリメラーゼ連鎖反応(PCR)によって増幅され、このような増幅のためのプライマーは、試薬キットに含まれる。このような実施態様では、増幅されたDNAは、当該検出プローブの鋳型および当該エンハンサープローブの鋳型として役立つ。

【0243】

1つの実施態様では、当該DNA鋳型は、本願明細書に記載される特定の多型マーカーの存在についての評価に先立って、全ゲノム増幅(WGA)法によって増幅される。WGAを実施するための当業者に周知の標準的な方法が利用され得、それらは本発明の範囲内にある。1つのこのような実施態様では、WGAを実施するための試薬は試薬キットの中に含まれる。

【0244】

ある実施態様では、特定のマーカー対立遺伝子またはハプロタイプの存在の特定は、乳癌に対する感受性(増加した感受性または減少した感受性)を示す。別の実施態様では、当該マーカー対立遺伝子またはハプロタイプの存在の特定は、乳癌治療薬に対する反応を示す。別の実施態様では、マーカー対立遺伝子またはハプロタイプの存在は、乳癌の予後診断を示す。さらに別の実施態様では、対立遺伝子マーカーまたはハプロタイプの存在は、乳癌治療の進行を示す。このような治療は、手術、薬物療法、または他の手段(例えば、生活スタイルの変化)による介入を含んでいてもよい。

【0245】

本発明のさらなる態様では、医薬品パック(キット)が提供され、当該パックは、本願明細書に開示される、治療薬および本発明の1以上の変異について診断上試験された当該治療薬のヒトへの投与のための一連の指示を含む。当該治療薬は、小分子薬物、抗体、ペプチド、アンチセンスもしくはrna i分子、または他の治療用分子であり得る。1つの実施態様では、本発明の少なくとも1つの変異体の保有者として同定された個体は、当該治療薬の処方される用量を摂取するように指示される。1つのこのような実施態様では、本発明の少なくとも1つの変異体のホモ接合保有者として同定された個体は、当該治療薬の処方される用量を摂取するように指示される。別の実施態様では、本発明の少なくとも1つの変異体の非保有者として同定された個体は、当該治療薬の処方される用量を摂取す

10

20

30

40

50

るように指示される。

【0246】

ある実施態様では、当該キットはさらに、当該キットを含む試薬を使用するための一連の指示を含む。ある実施態様では、当該キットはさらに、当該キットにより評価される多型マーカーと乳癌に対する感受性との間の相関データを含むデータの収集物を含む。

【0247】

(治療薬)

本願明細書に提示された乳癌についてのリスク変異体は、乳癌の新規な治療上の標的を同定するために有用であり得る。例えば、乳癌に関連する変異体(マーカーおよび/もしくはハプロタイプ)を含む遺伝子、もしくはこれらと連鎖不平衡の遺伝子(例えば、1つ以上のPAX5、TUB、SERPINH1、RAD51L1、FHOD3およびTNR C6B遺伝子)、またはこれらの生成物、ならびにこれらの変異遺伝子もしくはその産物によって直接的もしくは間接的に制御され、または相互作用する遺伝子またはこれらの産物は、乳癌の治療、または乳癌に関連する症候の発生の予防もしくは遅延のための治療薬の開発の標的にされ得る。治療薬は、1つ以上の、例えば、小さい非タンパク質分子および非核酸分子、タンパク質、ペプチド、タンパク質断片、核酸(DNA、RNA)、PNA(ペプチド核酸)、または標的遺伝子もしくはこれらの遺伝子産物の機能および/もしくはレベルを調節できるこれらの誘導体もしくは模倣薬を含んでいてもよい。

【0248】

本願明細書に記載された核酸および/もしくは変異体、またはこれらの相補的な配列を含む核酸は、細胞、組織もしくは器官における遺伝子発現を制御するためのアンチセンス構築物として使用されてもよい。アンチセンス技術に関連する方法論は当業者に周知であり、例えばAntisense Drug Technology: Principles, Strategies, and Applications、Crooke、編集、Marcel Dekker Inc.、New York(2001)において記載され、概説されている。通常は、アンチセンス薬剤(アンチセンスオリゴヌクレオチド)は、相補的なヌクレオチド分節に結合することができる一本鎖オリゴヌクレオチド(RNAまたはDNA)から構成される。この適切な標的配列を結合することにより、RNA-RNA、DNA-DNAまたはRNA-DNA二本鎖が形成される。当該アンチセンスオリゴヌクレオチドは、遺伝子のセンスまたはコード鎖に相補的である。当該アンチセンスオリゴヌクレオチドが二本鎖DNAに結合する三重らせんを形成することも可能である。

【0249】

アンチセンスオリゴヌクレオチドのいくつかのクラスは当業者に公知であり、切断剤(cleaver)および遮断剤(blocker)を含む。前者は標的RNA部位に結合し、当該標的RNAを切断する細胞内ヌクレアーゼ(例えば、リボヌクレアーゼHまたはリボヌクレアーゼL)を活性化する。遮断剤は、標的RNAに結合し、リボソームの立体的な障害によってタンパク質の翻訳を阻害する。遮断剤の例は、核酸、モルホリノ化合物、ロックド核酸およびメチルホスホン酸を含む(Thompson、Drug Discovery Today、7:912-917(2002))。アンチセンスオリゴヌクレオチドは、治療薬として直接的に有用であり、また、例えば、遺伝子ノックアウトまたは遺伝子ノックダウン実験によって、遺伝子機能を判定し、検証するのにも有用である。アンチセンス技術は、さらに、Laveryら、Curr. Opin. Drug Discov. Devel. 6:561-569(2003)、Stephensら、Curr. Opin. Mol. Ther. 5:118-122(2003)、Kurreck、Eur. J. Biochem. 270:1628-44(2003)、Diasら、Mol. Cancer Ther. 1:347-55(2002)、Chen、Methods Mol. Med. 75:621-636(2003)、Wangら、Curr. Cancer Drug Targets 1:177-96(2001)、およびBennett、Antisense Nucleic Acid Drug. Dev. 12:215-24(2002)に記載される。

10

20

30

40

50

【0250】

ある実施態様では、当該アンチセンス薬剤は、特定のヌクレオチド分節に結合することができるオリゴヌクレオチドである。ある実施態様では、当該ヌクレオチド分節は、PAX5、TUB、SERPINH1、RAD51L1、FHOD3およびTNRC6B遺伝子のいずれかを含む。ある他の実施態様では、当該アンチセンスヌクレオチドは配列番号：1～562に記載されるヌクレオチド分節に結合することができる。アンチセンスヌクレオチドは、長さが5～500ヌクレオチド（5～200ヌクレオチド、5～100ヌクレオチド、10～50ヌクレオチド、および10～30ヌクレオチドを含む）であり得る。ある好ましい実施態様では、当該アンチセンスヌクレオチドは、長さが14～50ヌクレオチド（14～40ヌクレオチドおよび14～30ヌクレオチドを含む）である。本願明細書に記載された変異はまた、特定の変異に特異的なアンチセンス試薬の選択および設計のために使用され得る。本願明細書に記載された変異についての情報を使用して、本発明の1以上の変異を含むmRNA分子を特異的に標的とするアンチセンスオリゴヌクレオチドまたは他のアンチセンス分子が設計され得る。このようにして、本発明の1以上の変異体（すなわちあるマーカー対立遺伝子および/またはハプロタイプ）を含むmRNA分子の発現は、阻害され得またはブロックされ得る。1つの実施態様では、当該アンチセンス分子は、標的核酸の特定の対立遺伝子の形態（すなわち、1つまたはいくつかの変異（対立遺伝子および/またはハプロタイプ））に特異的に結合して、これによりこの特定の対立遺伝子またはハプロタイプに由来する産物の翻訳を阻害するが、当該アンチセンス分子は当該標的核酸分子の特定の多形性部位での他のまたは別の変異には結合しないように設計される。アンチセンス分子は、遺伝子発現、従ってタンパク質発現を阻害するようにmRNAを不活化するために使用され得るので、当該分子は疾病治療のために使用され得る。この方法論は、mRNAが翻訳される能力を減弱させるmRNAの中の1以上の領域に相補的なヌクレオチド配列を含むリボザイムによる切断を含み得る。このようなmRNA領域としては、例えば、タンパク質コード領域、特にタンパク質の触媒活性、基質および/もしくはリガンド結合部位、または他の機能ドメインに対応するタンパク質コード領域が挙げられる。

10

20

【0251】

RNA干渉(RNAi)の現象は、C. elegans [Fireら、Nature 391:806-11(1998)]でのその最初の発見から、過去10年間で活発に研究されており、近年、ヒト疾病の治療におけるその有望な使用が、活発に探究されている(KimおよびRossi、Nature Rev. Genet. 8:173-204(2007)における概説)。RNA干渉(RNAi)はまた、遺伝子サイレンシングとも呼ばれ、特定の遺伝子を遮断するための二本鎖RNA分子(dsRNA)の使用に基づく。細胞では、細胞質二本鎖RNA分子(dsRNA)は、細胞の複合体によって、低分子干渉RNA(siRNA)に処理される。siRNAは、標的mRNA上の特定の部位へとタンパク質とRNAの複合体のターゲティングを導き、mRNAの切断を引き起こす[Thompson、Drug Discovery Today、7:912-917(2002)]。siRNA分子は、典型的には、長さが約20、21、22または23ヌクレオチドである。従って、本発明の1つの態様は、単離した核酸分子、およびRNA干渉のためのこれらの分子の使用、すなわち低分子干渉RNA分子(siRNA)としての使用に関する。1つの実施態様では、単離された核酸分子は長さが18～26ヌクレオチド、好ましくは長さが19～25ヌクレオチド、より好ましくは長さが20～24ヌクレオチド、およびより好ましくは長さが21、22または23ヌクレオチドである。

30

40

【0252】

RNAiを介した遺伝子サイレンシングの別の経路は、miRNA前駆体(pre-miRNA)を生じるために細胞内で処理される、内因的にコードされたプライマリーマイクロRNA(pri-miRNA)転写物に始まる。これらのmiRNA分子は、核から細胞質に輸送され、そこでこれらは成熟したmiRNA分子(miRNA)を生じるための処理を経て、mRNAの3'の非翻訳領域の標的部位を認識することによって翻訳の阻

50

害に方向付けられ、引き続いて、P体(P - bodies)の処理によってmRNAが分解する(KimおよびRossi、Nature Rev. Genet. 8:173-204(2007)における概説)。

【0253】

RNAiの臨床的適用は、好ましくはサイズが約20~23ヌクレオチド、および好ましくは2ヌクレオチドの3'重複を有する合成二本鎖siRNAを組み入れることを含む。遺伝子発現のノックダウンは、標的mRNAの配列特異的な設計によって確立される。このような分子の至適な設計および合成のためのいくつかの市販の部位は、当業者に知られている。

【0254】

他の適用は、より長いsiRNA分子(典型的には長さが25~30ヌクレオチド、好ましくは約27ヌクレオチド)、および低分子ヘアピン型RNA(shRNA;典型的には長さが約29ヌクレオチド)を提供する。後者は、Amarguiraら(FEB S Lett. 579:5974-81(2005))が記載しているように、自然に発現される。化学的に合成したsiRNAおよびshRNAは、in vivoでの処理のための基質であり、いくつかの場合では、より短い設計よりも、より強力な遺伝子サイレンシングを提供する[Kimら、Nature Biotechnol. 23:222-226(2005); Siolasら、Nature Biotechnol. 23:227-231(2005)]。一般的に、siRNAは、その細胞内濃度が、引き続く細胞分裂によって希釈されるため、遺伝子発現の一過性のサイレンシングを提供する。対照的に、発現されたshRNAは、shRNAの転写が起こっている限り、標的転写物の、長期間の安定したノックダウンを媒介する(Marquesら、Nature Biotechnol. 23:559-565(2006); Brummelkampら、Science 296:550-553(2002))。

【0255】

siRNA、miRNAおよびshRNAを含めたRNAi分子は、配列依存性様式で作用するため、本願明細書で提示された変異体は、特定の対立遺伝子および/またはハプロタイプ(例えば、本発明の対立遺伝子および/またはハプロタイプ)を含む特定の核酸分子を認識する一方、他の対立遺伝子またはハプロタイプを含む核酸分子を認識しないRNAi試薬を設計するために使用され得る。従って、これらのRNAi試薬は、標的核酸分子を認識し、破壊できる。アンチセンス試薬と同様に、RNAi試薬は治療薬として(すなわち、疾病に関連する遺伝子または疾病に関連する遺伝子変異を遮断するために)有用であり得るが、遺伝子機能の特徴付け、検証すること(例えば、遺伝子ノックアウト実験または遺伝子ノックダウン実験による)にもまた、有用であるかも知れない。

【0256】

RNAiの送達は、当業者に知られた方法論の範囲によって行ってもよい。非ウイルスデリバリーを利用する方法は、コレステロール、安定した核酸脂質粒子(SNALP)、重鎖抗体の断片(Fab)、アプタマーおよびナノ粒子を含む。ウイルスデリバリー方法は、レンチウイルス、アデノウイルスおよびアデノ随伴ウイルスの使用を含む。siRNA分子は、いくつかの実施態様では、その安定性を増加するために化学的に修飾される。これは、2'-O-メチルプリンおよび2'-フルオロピリミジンを含む、リボースの2'位での修飾を含むことができ、リボヌクレアーゼ活性に対する抵抗性を提供する。他の化学的修飾も可能であり、当業者に知られている。

【0257】

下記の参考文献は、さらなるRNAiの概略、およびRNAiを用いた特定の遺伝子のターゲティングの可能性を提供する: KimおよびRossi、Nat. Rev. Genet. 8:173-184(2007)、ChenおよびRajewsky、Nat. Rev. Genet. 8:93-103(2007)、Reynoldsら、Nat. Biotechnol. 22:326-330(2004)、Chiら、Proc. Natl. Acad. Sci. USA 100:6343-6346(2003)、Vicker

10

20

30

40

50

sら、J. Biol. Chem. 278:7108-7118 (2003)、Agami、Curr. Opin. Chem. Biol. 6:829-834 (2002)、Laveryら、Curr. Opin. Drug Discov. Devel. 6:561-569 (2003)、Shi、Trends Genet. 19:9-12 (2003)、Shueyら、Drug Discov. Today 7:1040-46 (2002)、McManusら、Nat. Rev. Genet. 3:737-747 (2002)、Xiaら、Nat. Biotechnol. 20:1006-10 (2002)、Plasterkerら、curr. Opin. Genet. Dev. 10:562-7 (2000)、Bosherら、Nat. Cell Biol. 2:E31-6 (2000)、およびHunter、Curr. Biol. 9:R440-442 (1999)。

10

【0258】

乳癌の発生について増加した素因またはリスクを引き起こす遺伝的欠陥、または乳癌の原因となる欠陥は、欠陥を有する被験者に、通常のノ野生型のヌクレオチド（単数または複数）を遺伝的欠陥部位で提供する修復配列を取り込む核酸断片を投与することによって、持続的に矯正されるかも知れない。このような部位特異的な修復配列は、被験者のゲノムDNAの内因性修復を促進するように機能するRNA/DNAオリゴヌクレオチドを含んで（concompas）もよい。修復配列の投与は、アニオン性リボソームに被包されたポリエチレンジミンとの複合体、ウイルスベクター（アデノウイルスベクターなど）、または投与した核酸の細胞内吸収を促進するのに適した他の医薬組成物などの、適切な媒体によって行われてもよい。次いで、遺伝的欠陥は、キメラのオリゴヌクレオチドが被験者のゲノムへの、通常の配列の組み込みを引き起こし、通常のノ野生型の遺伝子産物の発現を導くため、克服されるかも知れない。置換が増加し、従って、持続性の修復および疾病もしくは疾患に関連する症候の軽減を与える。

20

【0259】

本発明は、乳癌の治療に使用され得る化合物または薬剤を同定するための方法を提供する。従って、本発明の変異体は、治療薬の同定およびノまたは開発の標的として有用である。ある実施態様では、このような方法は、本発明の少なくとも1つの変異体（マーカーおよびノまたはハプロタイプ）、または核酸のコードされた生成物を含む核酸の活性およびノまたは発現を調節するための薬剤または化合物の能力のアッセイを含む。これは、例えば、1つ以上のFGF10、MRPS30、HCN1およびFGFR2遺伝子、ならび

30

【0260】

患者における変異遺伝子の発現は、変異を含む核酸配列（例えば、少なくとも1つの変異を含むRNAに転写されることができ、次にタンパク質に翻訳されることができ、少なくとも1つの本発明の変異を含む遺伝子）の発現、または通常の転写物の発現のレベルもしくは様式に影響する変異（例えば、遺伝子の制御領域または調節領域における変異）によって、通常のノ野生型の核酸配列が変化した発現によって評価され得る。遺伝子発現のアッセイは、直接の核酸アッセイ（mRNA）、発現したタンパク質レベルのアッセイ、または経路（例えば、シグナル経路）に関連する付随化合物のアッセイを含む。さらに、シグナル経路に反応して上方制御または下方制御される遺伝子の発現もまた、アッセイされ得る。1つの実施態様は、興味ある遺伝子（単数または複数）の制御領域に、レポーター遺伝子（ルシフェラーゼなど）を操作可能に結合することを含む。

40

【0261】

遺伝子発現の調節因子は、1つの実施態様では、細胞が候補化合物または候補薬と接触された場合に同定されることができ、mRNAの発現が判定される。候補化合物または候補薬の存在下でのmRNAの発現レベルは、当該化合物または薬剤の不存在下での発現レ

50

ベルと比較される。当該比較に基づき、乳癌の治療のための候補化合物または候補薬剤は、変異遺伝子の遺伝子発現を調節するものとして同定され得る。候補化合物または候補薬の存在下で、その不存在下よりも、mRNAまたはコード化タンパク質の発現が統計的に有意に大きい場合、当該候補化合物または候補薬は、核酸の発現の刺激剤または核酸の上方制御因子として同定される。候補化合物または候補薬の存在下で、その不存在下よりも、核酸の発現またはタンパク質レベルが統計的に有意に低い場合、その候補化合物は、核酸の発現の阻害剤または下方制御因子として同定される。

【0262】

本発明は、さらに、遺伝子の修飾因子（すなわち、遺伝子発現の刺激剤および/または阻害剤）としての薬剤（drug）（化合物および/または薬剤（agent））スクリーニングを通じて同定される化合物を使用した治療方法を提供する。

10

【0263】

（治療薬に対する反応の可能性の評価方法、治療の進展をモニターする方法および乳癌を治療する方法）

当該技術分野で公知のように、個体は、特定の治療（例えば、治療薬または治療方法、）に対する反応差を有しうる。反応差の基盤は、部分的に遺伝的に判定されてもよい。薬理ゲノミクスは、変化した薬剤処分および/または薬剤の異常な作用もしくは変化した作用によって、遺伝的変異（例えば、本発明の変異体（マーカーおよび/またはハプロタイプ））がどのように薬剤反応に影響するか、という問題点について取り組んでいる。従って、反応差の基盤は部分的に遺伝的に判定されてもよい。薬剤反応に影響する遺伝的変異に起因する臨床的成果は、ある個体（例えば、本発明の遺伝的変異の保有者または非保有者）における薬剤毒性、または当該薬剤の治療不全をもたらすかも知れない。従って、本発明の変異は、治療薬および/または方法が、身体に作用する様式、または身体が治療薬を代謝する方法を判定してもよい。

20

【0264】

従って、1つの実施態様では、多形性部位またはハプロタイプの特定の対立遺伝子の存在は、特定の治療様式に対する異なる反応の速度を示す。これは、乳癌と診断され、本発明の、多型における特定の対立遺伝子またはハプロタイプ（例えば、本発明の、リスクがありかつ保護的な対立遺伝子および/またはハプロタイプ）を有する患者は、疾病の治療に使用される、特定の治療上の薬剤および/または他の治療に対して、より良好に、または悪化して反応することを意味する。従って、マーカー対立遺伝子またはハプロタイプの存在または不存在は、患者に対して使用されるべき治療の決定の助けとなり得る。例えば、新規に診断された患者について、本発明のマーカーまたはハプロタイプの存在が（例えば、本願明細書記載のように、血液試料由来のDNAの試験を通じて）評価されてもよい。患者が、マーカー対立遺伝子またはハプロタイプに対して陽性である場合（つまり、マーカーの少なくとも1つの特定の対立遺伝子、またはハプロタイプが存在している）、医師は、1つの特定の治療方法を推奨し、一方、患者がマーカーの少なくとも1つの対立遺伝子、またはハプロタイプに陰性である場合、異なるコースの治療方法が推奨されてもよい（疾病の進行の連続的なモニター以外は、緊急ではない治療方法が行われることを推奨することを含んでいてもよい）。従って、患者の保有者状態は、特定の治療様式が施されるべきかどうかを判定するのに補助するために使用され得る。その価値は、最も適切な治療を適用できるようにするため、早期の段階で疾病を診断でき、最も適切な治療を選択でき、疾病の予後診断/悪性度について臨床医に情報を提供できる可能性の中にある。

30

40

【0265】

さらに本願明細書に記載のように、乳癌の現在の臨床予防的選択肢は、主に化学予防（化学療法、またはホルモン治療方法）および予防的手術である。最も一般的な化学予防は、タモキシフェンおよびラロキシフェンである。他の選択肢は、他の選択的エストロゲン受容体修飾因子（SERM）およびアロマターゼ阻害剤を含む。治療上の選択肢はまた、一部の患者が有害な症候を経験する放射線療法を含む。本発明のマーカーは、本願明細書記載のように、これらの治療上の選択肢に対する反応を評価するため、または、これらの

50

治療上の選択肢のいずれか1つを使用した治療方法の進行を予測するために使用されてもよい。従って、遺伝的鑑定は、個体の遺伝的状态に基づいた適切な治療戦略の選択に使用され得、または特定の治療の選択肢の成果の予測に使用されてもよく、従って、治療上の選択肢または利用可能な治療上の選択肢の組み合わせの戦略的選択について有用であり得る。

【0266】

本発明はまた、乳癌の治療の進行または有効性をモニターする方法に関する。これは、本発明のマーカーおよびハプロタイプの遺伝子型および/またはハプロタイプの状態に基づいて行われ得、すなわち、本願明細書で開示された少なくとも1つの多型マーカーの少なくとも1つの対立遺伝子の不存在または存在の評価、または本発明の変異体(マーカーおよびハプロタイプ)に関連する遺伝子の発現のモニターによってなされ得る。リスク遺伝子のmRNAまたはコードされたポリペプチドは、組織試料(例えば、抹消血試料、または生検試料)において測定され得る。従って、発現レベルおよび/またはmRNAレベルは、その有効性をモニターするために、治療の前および治療の間に判定され得る。代わりに、または同時に、本願明細書に示された乳癌の少なくとも1つのリスク変異の遺伝子型および/またはハプロタイプの状態は、その有効性をモニターするために、治療の前および治療の間に判定される。

10

【0267】

あるいは、本発明のマーカーおよびハプロタイプに関連する生物学的ネットワークまたは代謝経路は、mRNAおよび/またはポリペプチドレベルの測定によってモニターされ得る。例えば、これは、治療前および治療後に得られた試料中の、ネットワークおよび/または経路に属するいくつかの遺伝子について、発現レベルまたはポリペプチドをモニターすることによってなされ得る。あるいは、生物学的ネットワークまたは代謝経路に属する代謝物を、治療前および治療後に測定してもよい。治療の有効性は、治療の間に、発現レベル/代謝物レベルで認められた変化を、健康な被験者から得た対応するデータと比較することによって判定される。

20

【0268】

さらなる態様では、本発明のマーカーは、臨床試験の効力および有効性を増加するために使用され得る。従って、本発明のリスクのある変異の保有者である個体、すなわち、乳癌の発生について増加したリスクを与える少なくとも1つの多型マーカーの少なくとも1つの対立遺伝子の保有者である個体は、特定の治療様式に対し、より反応する可能性があってもよい。1つの実施態様では、特定の治療(例えば、小分子薬剤)が標的にしている経路および/または代謝ネットワークの遺伝子(単数または複数)についてリスクのある変異を保有する個体は、当該治療に対してより反応する可能性がある。別の実施態様では、発現および/または機能がリスクのある変異によって変化された遺伝子についてリスクのある変異を保有する個体は、その遺伝子、その発現またはその遺伝子産物を標的とした治療様式に対してより反応する可能性がある。

30

【0269】

さらなる態様では、本発明のマーカーおよびハプロタイプは、特定の個体の医薬品の選択のターゲティングに使用され得る。治療様式の個別化された選択、生活習慣の変化、またはこれら2つの組み合わせは、本発明のリスクのある変異体の利用によって実現され得る。従って、本発明の特定のマーカーについての個体の状態の知識は、本発明のリスクのある変異によって影響される遺伝子または遺伝子産物を標的にする、治療上の選択肢の選択について有用であり得る。変異の特定の組み合わせは、治療上の選択肢の1つの選択に適してもよく、一方で他の遺伝子変異の組み合わせは他の治療上の選択肢を標的としてよい。治療要素の選択を臨床的に信頼できる正確性をもって判定するために必要に応じて、このような変異の組み合わせは、1つの変異、2つの変異、3つの変異、または4つ以上の変異を含んでもよい。

40

【0270】

(コンピューターで実施される態様)

50

当業者が理解するとおり、本願明細書に記載される方法および情報は、全体がまたは一部が、公知のコンピューター可読媒体上のコンピューターが実行可能な命令として実施され得る。例えば、本願明細書に記載される方法は、ハードウェアで実施され得る。あるいは、当該方法は、例えば、1以上のメモリまたは他のコンピューター可読媒体に保存されたソフトウェアで実施され得、そして1以上のプロセッサ上で実施され得る。公知のように、当該プロセッサは、1以上の制御装置、計算ユニットおよび/またはコンピューターシステムの他のユニットに関連付けられ得るか、または所望に応じてファームウェアに埋め込まれ得る。ソフトウェアで実施される場合、ルーチンは、公知のようにRAM、ROM、フラッシュメモリ、磁気ディスク、レーザーディスク、または他の記憶媒体などのいずれかのコンピューター可読メモリの中に保存され得る。同様に、このソフトウェアは、

10

【0271】

より一般的には、そして当業者が理解するように、上記の種々の工程は、種々のブロック、演算、ツール、モジュールおよび技術として実施され得、これらは、ひいてはハードウェア、ファームウェア、ソフトウェア、またはハードウェア、ファームウェア、および/もしくはソフトウェアのいずれかの組み合わせで実施され得る。ハードウェアで実施される場合、当該ブロック、演算、技術などのうちのいくつかまたはすべては、例えば、カスタム集積回路(IC)、特定用途向け集積回路(ASIC)、フィールドプログラマブル論理アレイ(FPGA)、プログラマブル論理アレイ(PLA)などで実施され得る。

20

【0272】

ソフトウェアで実施される場合、当該ソフトウェアは、いずれかの公知のコンピューター可読媒体に、例えば磁気ディスク、光ディスク、もしくは他の記憶媒体上に、コンピューターのRAMまたはROMまたはフラッシュメモリに、プロセッサ、ハードディスクドライブ、光ディスクドライブ、テープドライブなどに保存され得る。同様に、当該ソフトウェアは、いずれかの公知の配信方法を介して(例えば、コンピューター可読ディスクまたは他の持ち運びできるコンピューター記憶装置上でなど)ユーザーまたはコンピューターシステムへと配信され得る。

【0273】

図1は、請求項に係る方法および装置の工程のためのシステムが実施され得る、適したコンピューターシステム環境100の例を図示する。コンピューターシステム環境100は、適したコンピューター環境の1つの例に過ぎず、請求項の方法または装置の使用または機能性の範囲に関して何らかの限定を示唆することを意図するものではない。コンピューター環境100は、代表的な動作環境100に図示された構成要素のいずれか1つまたは組み合わせに関して何らかの依存性または必要条件を有すると解釈されるべきでもない。

30

【0274】

請求項に係る方法およびシステムの工程は、多数の他の一般的な目的または特別の目的のコンピューターシステム環境または機器構成を用いて動作できる。請求項の方法またはシステムとの使用に適切であり得る周知のコンピューターシステム、環境、および/または機器構成の例としては、パーソナルコンピューター、サーバーコンピューター、携帯端末またはラップトップ型装置、マルチプロセッサシステム、マイクロプロセッサを用いるシステム、セットトップボックス、プログラムできる家庭用電化製品、ネットワークPC、ミニコンピューター、メインフレームコンピューター、上記のシステムまたは装置のいずれかを含む分散コンピューティング環境などが挙げられるが、これらに限定されない。

40

【0275】

請求項に係る方法およびシステムの工程は、コンピューターによって実行される、プログラムモジュールなどのコンピューターが実行可能な命令の一般的なコンテキストで記述され得る。一般的に、プログラムモジュールは、特定のタスクを実行し、または特定の抽

50

象データ型を実施するルーチン、プログラム、オブジェクト、構成要素、データ構造などを含む。当該方法および装置は、通信ネットワークを通してリンクされている遠隔の処理装置によってタスクが実行される、分散コンピューティング環境においても実行され得る。統合コンピューティング環境および分散コンピューティング環境の両方において、プログラムモジュールは、メモリ記憶装置を含むローカルのおよび遠隔のコンピューター記憶媒体の両方に位置し得る。

【0276】

図1を参照して、請求項に係る方法およびシステムの工程を実施するための代表的なシステムは、コンピューター110の形態の一般的な目的の計算装置を含む。コンピューター110の構成要素としては、処理装置120、システムメモリ130、およびシステムメモリを含む種々のシステム構成要素を処理装置120に結合するシステムバス121が挙げられ得るが、これらに限定されない。システムバス121は、メモリバスまたはメモリコントローラ、周辺機器用バス、および様々なバスアーキテクチャのうちのいずれかを使用するローカルバスを含めたいくつかの種類のバス構造のいずれかであり得る。例として、このようなアーキテクチャとしては、業界標準アーキテクチャ（ISA）バス、マイクロ・チャンネル・アーキテクチャ（MCA）バス、拡張ISA（EISA）バス、ビデオ・エレクトロニクス・スタンダーズ・アソシエーション（Video Electronics Standards Association）（VESA）ローカルバス、およびMezzanine（中二階）バスとしても知られるペリフェラル・コンポーネント・インターコネクト（Peripheral Component Interconnect）（PCI）バスが挙げられるが、これらに限定されない。

10

20

【0277】

コンピューター110は、典型的には様々なコンピューター可読媒体を含む。コンピューター可読媒体は、コンピューター110によってアクセスされ得るいずれの利用可能な媒体であり得、例としては揮発性媒体および不揮発性媒体の両方、リムーバブル媒体および非リムーバブル媒体が挙げられる。例として、コンピューター可読媒体は、コンピューター記憶媒体および通信媒体を含み得るが、これらに限定されない。コンピューター記憶媒体としては、コンピューター可読命令、データ構造、プログラムモジュールまたは他のデータなどの情報の保存のためのいずれかの方法または技術で実施される、揮発性および不揮発性の両方の、リムーバブル媒体および非リムーバブル媒体が挙げられる。コンピューター記憶媒体としては、RAM、ROM、EEPROM、フラッシュメモリもしくは他のメモリ技術、CD-ROM、デジタルバーサタイルディスク（DVD）もしくは他の光ディスク記憶装置、磁気カセット、磁気テープ、磁気ディスク記憶装置または他の磁気記憶装置、あるいは所望の情報をするために使用され得かつコンピューター110によってアクセスされ得るいずれかの他の媒体が挙げられるが、これらに限定されない。通信媒体は、典型的には、搬送波または他の輸送機構などの変調されたデータ信号の中にコンピューター可読命令、データ構造、プログラムモジュールまたは他のデータを具現化し、いずれかの情報配信媒体を含む。用語「変調されたデータ信号」は、ある信号の特性集合のうち1以上が、情報を当該信号の中でコード化するような様式で変更されている、その信号を意味する。例として、通信媒体としては、有線ネットワークまたは直接配線接続などの有線媒体、ならびに音響、RF、赤外線および他の無線媒体などの無線媒体が挙げられるが、これらに限定されない。上記のものうちのいずれかの組み合わせも、コンピューター可読媒体の範囲内に包含されるはずである。

30

40

【0278】

システムメモリ130は、読み出し専用メモリ（ROM）131およびランダムアクセスメモリ（RAM）132などの揮発性メモリおよび/または不揮発性メモリの形態のコンピューター記憶媒体を含む。起動中などにコンピューター110内の要素間で情報を転送することを助ける基本ルーチンを含む基本入出力システム133（BIOS）は、典型的にはROM 131に保存される。RAM 132は、典型的には、処理装置120が直ちにアクセス可能であり、かつ/または現在処理装置120が演算しているデータおよ

50

び/またはプログラムモジュールを含む。例として、図1は、オペレーティングシステム134、アプリケーションプログラム135、他のプログラムモジュール136、およびプログラムデータ137を図示するが、これらに限定されない。

【0279】

コンピューター110はまた、他のリムーバブル/非リムーバブル、揮発性/不揮発性のコンピューター記憶媒体を含み得る。例に過ぎないが、図1は、非リムーバブルの、不揮発性の磁気媒体から読み取りまたは非リムーバブルの、不揮発性の磁気媒体に書き込むハードディスクドライブ140、リムーバブルの、不揮発性の磁気ディスク152から読み取りまたはリムーバブルの、不揮発性の磁気ディスク152に書き込む磁気ディスクドライブ151、およびCD-ROMまたは他の光媒体などのリムーバブルの、不揮発性の光ディスク156から読み取りまたはリムーバブルの、不揮発性の光ディスク156に書き込む光ディスクドライブ155を図示する。代表的な動作環境で使用され得る他のリムーバブル/非リムーバブルの、揮発性/不揮発性のコンピューター記憶媒体としては、磁気テープカセット、フラッシュメモリカード、デジタルバーサタイルディスク、デジタルビデオテープ、ソリッドステートRAM、ソリッドステートROMなどが挙げられるが、これらに限定されない。ハードディスクドライブ141は、典型的には、非リムーバブルメモリインターフェース(インターフェース140など)を介してシステムバス121に接続され、磁気ディスクドライブ151および光ディスクドライブ155は、典型的にはリムーバブルメモリインターフェース(インターフェース150など)によってシステムバス121に接続される。

10

20

【0280】

上で論じ図1に図示したドライブおよびそれらの関連するコンピューター記憶媒体は、コンピューター110のためのコンピューター可読命令、データ構造、プログラムモジュールおよび他のデータの保存をもたらす。図1では、例えば、ハードディスクドライブ141は、保存用(storing)オペレーティングシステム144、アプリケーションプログラム145、他のプログラムモジュール146、およびプログラムデータ147として図示されている。これらの構成要素は、オペレーティングシステム134、アプリケーションプログラム135、他のプログラムモジュール136、およびプログラムデータ137と同じであってもよいし、異なってもよいことに留意されたい。オペレーティングシステム144、アプリケーションプログラム145、他のプログラムモジュール146、およびプログラムデータ147は、最小限それらが異なるコピーであることを図示するために、ここでは異なる数が与えられている。ユーザーは、キーボード162およびポインティングデバイス161(一般にマウスと呼ばれる)、トラックボールまたはタッチパッドなどの入力装置を通して、コマンドおよび情報をコンピューター20へと入力し得る。他の入力装置(図示せず)は、マイクロホン、ジョイスティック、ゲームパッド、パラボラアンテナ、スキャナーなどを含み得る。これらおよび他の入力装置は、システムバスに結合されているユーザー入力インターフェース160を通して処理装置120に接続されることが多いが、パラレルポート、ゲームポートまたはユニバーサル・シリアル・バス(USB)などの他のインターフェースおよびバス構造によって接続されてもよい。モニター191または他の種類のディスプレイ装置はまた、ビデオインターフェース190などのインターフェースを介してシステムバス121に接続されている。モニターに加えて、コンピューターはまた、スピーカー197およびプリンタ196などの他の周辺出力装置を含み得、これらは出力用周辺インターフェース190を通して接続され得る。

30

40

【0281】

コンピューター110は、リモートコンピューター180などの1以上のリモートコンピューターへの論理結合を使用して、ネットワーク化された環境の中で動作し得る。リモートコンピューター180は、パーソナルコンピューター、サーバー、ルーター、ネットワークPC、ピア装置(peer device)または他の共通のネットワークノードであり得、そして典型的には、コンピューター110に対する上記の要素のうち多くまたはすべてを含むが、メモリ記憶装置181だけが図1に図示されている。図1に図示さ

50

れる論理結合は、ローカルエリアネットワーク（LAN）171および広域ネットワーク（WAN）173を含むが、他のネットワークも含み得る。このようなネットワーク環境は、オフィス、企業規模のコンピューターネットワーク、イントラネットおよびインターネット内では普通のものである。

【0282】

LANネットワーク環境で使用される場合、コンピューター110は、ネットワークインターフェースまたはアダプター170を通してLAN171に接続される。WANネットワーク環境で使用される場合、コンピューター110は、典型的には、モデム172またはインターネットなどのWAN173にわたって通信を確立するための他の手段を含む。モデム172（内蔵モデムまたは外付けモデムであり得る）は、ユーザー入力インターフェース160、または他の適切な機構を介してシステムバス121に接続され得る。ネットワーク化された環境では、コンピューター110に対して示されたプログラムモジュール、またはその一部分は、遠隔のメモリ記憶装置に保存され得る。例として、図1は、遠隔のアプリケーションプログラム185をメモリ装置181上に存在するとして図示しているが、これに限定されない。示されるネットワーク接続は代表的なものであり、コンピューター間の通信リンクを確立するための他の手段が使用され得ることは分かるであろう。

10

【0283】

上記の本文は本発明の多数の異なる実施態様の詳細な説明を記載するが、本発明の範囲は、添付の特許請求の範囲の文言によって定められるということを理解されたい。この詳細な説明は、代表的なものに過ぎないと解釈されるべきであり、あらゆる可能性のある本発明の実施態様を記載しているわけではない。なぜなら、あらゆる可能性のある実施態様を記載することは、不可能ではないかも知れないが、実益がないと思われるからである。多数の代替の実施態様が、現在の技術または本特許の出願日以後に開発される技術のいずれかを使用して実施され得、それらは、依然として本発明を画定する特許請求の範囲の範囲内に含まれることになる。

20

【0284】

リスク評価システムおよび方法、ならびに他の要素が、好ましくはソフトウェアで実施されるとして記載されてきたが、それらはハードウェア、ファームウェアなどで実施され得、いずれかの他のプロセッサによって実施され得る。従って、本願明細書に記載された要素は、標準的な多目的CPUでまたは所望に応じて特定用途向け集積回路（ASIC）または他の配線接続された装置などの特定の設計されたハードウェアまたはファームウェア（図1のコンピューター110が挙げられるが、これに限定されない）上で実施され得る。ソフトウェアで実施される場合、当該ソフトウェアルーチンは、磁気ディスク、レーザーディスク、または他の記憶媒体上などのいずれかのコンピューター可読メモリに、コンピューターまたはプロセッサのRAMまたはROMに、いずれかのデータベースなどに保存され得る。同様に、このソフトウェアは、例えばコンピューター可読ディスクもしくは他の持ち運びできるコンピューター記憶装置上、または電話線、インターネット、無線通信など通信チャネルにわたって（これらは、持ち運びできる記憶媒体を介してこのようなソフトウェアを提供することと同じまたは置き換え可能であると考えられる）、を含めた、いずれかの公知のまたは所望の配信方法を介してユーザーまたは診断システムに配信され得る。

30

40

【0285】

従って、本発明の趣旨および範囲から逸脱せずに、本願明細書に記載および図示された技術および構造の中で、多くの改変態様および変更態様がなされ得る。従って、本願明細書に記載された方法および装置は例示的なものに過ぎず、本発明の範囲に対する限定にはならないということを知りたい。

【0286】

従って、本発明は、乳癌に関連すると本願明細書に記載された多型マーカーおよびハプロタイプのコンピューターで実施される応用例に関する。このような応用例は、本発明の

50

方法で有用な遺伝子型データの保存、操作、または別の態様での分析のために有用であり得る。1つの例は、遺伝子型情報を第三者（例えば、個体、その個体の保護者、医療提供者または遺伝学的解析サービスの提供者）に提供することができるように、または例えば当該遺伝子型データを癌に対する増加した感受性に寄与する遺伝的危険因子についての情報と比較することにより、当該遺伝子型データから情報を誘導し、そしてこのような比較に基づいて結果を報告するために、個体由来の遺伝子型情報を可読媒体上に保存することに関する。

【0287】

一般的に言えば、コンピューター可読媒体は、(i)本願明細書に記載された少なくとも1つの多型マーカーまたはハプロタイプの識別子の情報、(ii)当該疾病の個体における、当該少なくとも1つのマーカーの少なくとも1つの対立遺伝子の頻度の指標もしくはハプロタイプの頻度の指標、および参照集団における、当該少なくとも1つのマーカーの少なくとも1つの対立遺伝子の頻度の指標もしくはハプロタイプの頻度の指標、を保存する性能を有する。当該参照集団は、個体の疾病フリーの集団であり得る。あるいは、当該参照集団は、一般的な集団から得た確率標本であり、従って、集団全体を代表する。頻度の指標は、算出された頻度、対立遺伝子および/またはハプロタイプのコピーの数、または、特定の媒体に適した実際の頻度の規準化された値もしくは別の態様で操作された値であってもよい。

【0288】

本願明細書で、乳癌の増加した感受性（増加したリスク）に関連すると記載されたマーカーおよびハプロタイプは、ある実施態様では、遺伝子型のデータの解釈および/または分析に有用である。従って、ある実施態様では、本願明細書で示された、乳癌のリスクのある対立遺伝子の存在の判定、またはいずれかのこのようなリスク対立遺伝子とLDである多型マーカーの対立遺伝子の存在の判定は、遺伝子型のデータの由来である個体が乳癌について増加したリスクを有することを示す。そのような1つの実施態様では、遺伝子型のデータは、本願明細書で乳癌に関連すると示された少なくとも1つの多型マーカー、またはそれらと連鎖不平衡のマーカーについて生じる。当該遺伝子型のデータは、例えば、インターネットに接続可能なユーザーインターフェースを介して、当該遺伝子型のデータの解釈とともに、例えば、疾病についてのリスク判定基準（絶対的なリスク（AR）、リスク比（RR）またはオッズ比（OR）など）の形態で、その後、当該データの由来である個体、その個体の保護者または代表者、医師または医療従事者、遺伝子カウンセラー、または保険代理業者などの第三者に利用可能となる。別の実施態様では、個体由来の遺伝子型のデータセットで同定されたリスクのあるマーカーが評価され、当該データセットにおけるこのようなリスクのある変異体の存在によって与えられるリスク評価から得た結果は、例えば、安全なウェブインターフェースを介して、または他の伝達手段によって当該第三者に利用可能となる。このようなリスクアセスメントの結果は、数値的な形態（例えば、絶対的なリスク、相対リスク、および/もしくはオッズ比などのリスク値、または参照と比較したリスクの増加した割合による）、図面による方法、または当該遺伝子型のデータの由来である個体のリスクを図解するのに適した他の手段で報告され得る。

【0289】

（核酸およびポリペプチド）

本願明細書記載の核酸およびポリペプチドは、上記に記載のように、本発明の方法およびキットで使用され得る。「単離された」核酸分子は、本願明細書で使用される場合、（ゲノム配列における）遺伝子またはヌクレオチド配列に通常に隣接した核酸から分離されたもの、および/または、（例えば、RNAライブラリーにおける）他の転写された配列から完全にもしくは部分的に精製したものである。例えば、本発明の単離された核酸は、自然に生じる複雑な細胞環境、または組み換え技術によって産生された場合の培養培地、または化学的に合成された場合の化学的前駆体または他の化学物質に対して実質的に単離され得る。いくつかの例では、単離された物質は、組成物（例えば、他の物質を含む粗抽出物）、バッファーシステムまたは試薬混合物の一部を形成するだろう。他の環境では、

10

20

30

40

50

当該物質は、例えば、ポリアクリルアミドゲル電気泳動 (PAGE) またはカラムクロマトグラフィー (例えば、HPLC) によって判定されるように、基本的に均一に精製され得る。本発明の単離された核酸分子は、存在する全ての高分子の種のうち、少なくとも約 50%、少なくとも約 80% または少なくとも約 90% (モルベースによる) を構成し得る。ゲノム DNA に関しては、用語「単離された」はまた、当該ゲノム DNA が本来関連する染色体から分離される核酸分子も意味し得る。例えば、単離した核酸分子は、約 250 kb、200 kb、150 kb、100 kb、75 kb、50 kb、25 kb、10 kb、5 kb、4 kb、3 kb、2 kb、1 kb、0.5 kb または 0.1 kb 未満の、当該核酸分子が由来する細胞のゲノム DNA の核酸分子に隣接するヌクレオチドを含み得る。

10

【0290】

当該核酸分子は、他のコード配列または制御配列に融合され得、依然として単離されたものとみなされ得る。従って、ベクターに包含された組み換え DNA は、本願明細書で使用される「単離された」の定義に含まれる。また、単離された核酸分子は、異種の宿主細胞もしくは異種生命体中の組み換え DNA 分子、および部分的にもしくは実質的に精製された溶液中の DNA 分子を含む。「単離された」核酸分子はまた、*in vivo* および *in vitro* での本発明の DNA 分子の RNA 転写物を含む。単離した核酸分子またはヌクレオチド配列は、化学的に合成された、または組み換え手段によって合成された核酸分子またはヌクレオチド配列を含み得る。このような単離されたヌクレオチド配列は、例えば、コードされたポリペプチドの製造において、(例えば、他の哺乳類の種からの) 相同配列の単離のため、(例えば、染色体とのインサイツハイブリダイゼーションによる) 遺伝子マッピングのため、または組織 (例えば、ヒト組織) 中の遺伝子発現の検出 (ノーザンプロット分析または他のハイブリダイゼーション技術などによる) のためのプローブとして有用である。

20

【0291】

本発明はまた、高いストリンジェンシーハイブリダイゼーション条件下 (選択的ハイブリダイゼーションなど) で、本願明細書記載のヌクレオチド配列 (例えば、本願明細書記載のマーカーまたはハプロタイプに関連する多形性部位を含むヌクレオチド配列に特異的にハイブリダイズする核酸分子) にハイブリダイズする核酸分子に関する。このような核酸分子は、(例えば、高いストリンジェンシー条件下での) 対立遺伝子特異的なハイブリダイゼーションまたは配列特異的なハイブリダイゼーションによって検出および/または単離され得る。核酸ハイブリダイゼーションのストリンジェンシー条件および方法は、当業者に周知 (例えば、*Current Protocols in Molecular Biology*, Ausubel, F. ら、John Wiley & Sons, (1998)、および Kraus, M. および Aaronson, S., *Methods Enzymol.*, 200: 546-556 (1991) を参照、その全体の教示は参照によって本願明細書に援用する) である。

30

【0292】

2つのヌクレオチドまたはアミノ酸配列の同一性 (%) は、至適な比較目的で配列を整理することによって判定され得る (例えば、間隙は、第1の配列の配列に導入され得る)。対応する位置のヌクレオチドまたはアミノ酸は、次いで、比較され、2つの配列間の同一性 (%) は、当該配列によって共有される同一の位置の数の関数である (すなわち、同一性 (%) = 同一の位置の数 / 位置の総数 × 100)。ある実施態様では、比較目的で整理させた配列の長さは、参照配列の長さのうち、少なくとも 30%、少なくとも 40%、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%、または少なくとも 95% である。2つの配列の実際と比較は、周知の方法、例えば、数学的アルゴリズムを使用することによって達成され得る。このような数学的アルゴリズムの非限定的な例は、Karlin, S. および Altschul, S., *Proc. Natl. Acad. Sci. USA*, 90: 5873-5877 (1993) に記載される。このようなアルゴリズムは、Altschul S. ら、*Nucleic A*

40

50

cids Res.、25:3389-3402(1997)に記載されるように、NB
LASTプログラムおよびXB LASTプログラム(バージョン2.0)に取り込まれる。
BLASTプログラムおよびGapped BLASTプログラムを利用する場合、そ
れぞれのプログラム(例えば、NB LAST)の初期状態のパラメータが使用され得る。
ワールドワイドウェブ(ncbi.nlm.nih.gov.)上のウェブサイトを参照
。1つの実施態様では、配列比較のパラメータはスコア=100、語長=12に設定し得
るし、または、異なり得る(例えば、W=5またはW=20)。アルゴリズムの別の例は
BLAT[Kent, W. J. Genome Res. 12:656-64(200
2)]である。

【0293】

他の例は、MyersおよびMiller、CABIOS(1989)のアルゴリズム
、Torellis A.およびRobotti C.、Comput. Appl. Bi
osci. 10:3-5(1994)に記載されたADVANCEおよびADAM、およ
びPearson W.およびLipman, D.、Proc. Natl. Acad. S
ci. USA、85:2444-48(1988)に記載されたFASTAを含む。

【0294】

別の実施態様では、2つのアミノ酸配列間の同一性(%)は、GCGソフトウェアパッ
ケージ(アクセルリス(Accelrys)、英国、ケンブリッジ)におけるGAPプロ
グラムを使用して達成され得る。

【0295】

本発明はまた、本願明細書で定義されるLDブロックC09、LDブロックC10A、
LDブロックC10B、LDブロックC11A、LDブロックC11B、LDブロックC1
4、LDブロックC18、LDブロックC22A、およびLDブロックC22Bのすべ
てもしくは一部分を含むヌクレオチド配列を含む(またはこれらからなる)核酸;または本
願明細書中の表4(配列番号:1~562)に記載された多型マーカーのうちの少なくと
も1つを含む核酸;またはPAXS、TUB、SERPINH1、RAD51L1、FH
OD3およびTNRC6B遺伝子のうちのいずれか1つのヌクレオチド配列;あるいは本
願明細書で定義されるLDブロックC09、LDブロックC10A、LDブロックC10B
、LDブロックC11A、LDブロックC11B、LDブロックC14、LDブロックC
18、LDブロックC22A、およびLDブロックC22Bのすべてもしくは一部分を含
むヌクレオチド配列の補完物;または本願明細書中の表4(配列番号:1~562)に記
載された多型マーカーのうちの少なくとも1つを含む核酸の補完物;またはPAXS、T
UB、SERPINH1、RAD51L1、FHOD3およびTNRC6B遺伝子のうち
のいずれか1つのヌクレオチド配列の補完物を含む(またはこれらから構成される)ヌク
レオチド配列に、高度にストリンジェントな条件下でハイブリダイズする断片または部分
を含む、単離した核酸分子を提供する。

【0296】

本発明の核酸断片は、長さが少なくとも約15、少なくとも約18、20、23または
25ヌクレオチドであり、30、40、50、100、200、500、1000、10
, 000またはこれより多いヌクレオチドであり得る。

【0297】

本発明の核酸断片は、本願明細書記載のようなアッセイにおけるプローブまたはプライ
マーとして使用される。「プローブ」または「プライマー」は、塩基特異的様式で核酸分
子の相補鎖にハイブリダイズする、オリゴヌクレオチドである。DNAおよびRNAに加
えて、このようなプローブおよびプライマーは、Nielsen P.ら、Scienc
e 254:1497-1500(1991)において記載されたように、ポリペプチド
核酸(PNA)を含む。プローブまたはプライマーは、核酸分子の少なくとも約15、典
型的には約20~25、およびある実施態様では約40、50または75の、核酸分子の
連続的ヌクレオチドにハイブリダイズするヌクレオチド配列領域を含む。1つの実施態
様では、当該プローブまたはプライマーは、本願明細書記載の少なくとも1つの多型マーカー

10

20

30

40

50

一の少なくとも1つの対立遺伝子もしくは少なくとも1つのハプロタイプ、またはその補完物を含む。特定の実施態様では、プローブまたはプライマーは、100またはより少ないヌクレオチド（例えば、ある実施態様では、6～50ヌクレオチド、または、例えば、12～30ヌクレオチド）を含み得る。他の実施態様では、当該プローブまたはプライマーは、近接するヌクレオチド配列もしくは当該近接するヌクレオチド配列の補完物と、少なくとも70%の同一性、少なくとも80%の同一性、少なくとも85%の同一性、少なくとも90%の同一性、または少なくとも95%の同一性である。別の実施態様では、当該プローブまたはプライマーは、近接するヌクレオチド配列または当該近接するヌクレオチド配列の補完物に、選択的にハイブリダイズできる。しばしば、当該プローブまたはプライマーは、さらに、標識（例えば、放射性同位元素、蛍光標識、酵素標識、酵素共因子標識、磁性標識、スピン標識、エピトープ標識）を含む。

10

【0298】

上記などの本発明の核酸分子は、当業者に周知の標準的な分子生物学技術を使用して同定され、単離され得る。増幅したDNAは、標識（例えば、放射標識）され得、ヒト細胞由来のcDNAライブラリーのスクリーニングのためのプローブとして使用され得る。cDNAは、mRNA由来であり得、適したベクター中に含まれ得る。対応するクローンが単離され得、DNAは*in vivo*での切除の後に得ることができ、クローン化された挿入断片は、一方向または両方向で、当該技術分野で承認されている方法によって、適切な分子量のポリペプチドをコードする正しい読み枠を同定し、配列決定され得る。これらの方法または同様の方法を使用して、ポリペプチドおよびポリペプチドをコードするDNAは単離され、配列決定され、さらに特徴付けられ得る。

20

【0299】

(抗体)

本発明はまた、変異対立遺伝子によってコードされる変異アミノ酸配列（例えば、アミノ酸置換を含む）または対応する非変異対立遺伝子または野生型対立遺伝子によってコードされる参照アミノ酸配列のいずれかを含むエピトープに結合する抗体を提供する。用語「抗体」は、本願明細書で使用される場合、免疫グロブリン分子および免疫グロブリン分子の免疫学的に活性な部分、すなわち、特異的に抗原に結合する抗原結合部位を含む分子を意味する。本発明のポリペプチドに特異的に結合する分子は、そのポリペプチドまたはその断片に結合するが、試料中、例えば、自然に当該ポリペプチドを含む生物学的試料中の他の分子には実質的に結合しない分子である。免疫グロブリン分子の免疫学的に活性な部分の例は、抗体をペプシンなどの酵素で処理することによって生じ得るF(ab)およびF(ab')₂断片を含む。本発明は、本発明のポリペプチドに結合するポリクローナル抗体およびモノクローナル抗体を提供する。用語「モノクローナル抗体」または「モノクローナル抗体の組成物」は、本願明細書で使用される場合、本発明のポリペプチドの特定のエピトープと免疫反応できる1種のみ抗原結合部位を含む抗体分子の集団を意味する。従って、モノクローナル抗体の組成物は、典型的には、免疫反応する本発明の特定のポリペプチドに対する単一の結合アフィニティを示す。

30

【0300】

ポリクローナル抗体は、適した被験者に、所望の免疫原（例えば、本発明のポリペプチドまたはその断片）で免疫することによって、上記のように、調製され得る。免疫された被験者における抗体力価は、酵素結合免疫吸着検定法(ELISA)などの標準的な技術によって、固定化されたポリペプチドを使用して、経時的にモニターされ得る。所望であれば、ポリペプチドに対して作られた抗体分子は、哺乳類から（例えば、血液から）単離され得、さらに、IgG分画を得るために、プロテインAクロマトグラフィーなどの周知の技術によって精製され得る。免疫後の適切な時期に、例えば、抗体力価が最も高いときに、抗体産生細胞が被験者から入手され、標準的な技術(KohlerおよびMilstein、Nature 256:495-497(1975))に最初に記載された融合細胞法、ヒトB細胞の融合細胞法(Kozborら、Immunol. Today 4:72(1983))、EBV融合細胞法(Coleら、Monoclonal Anti

40

50

bodies and Cancer Therapy, Alan R. Liss, 1985, Inc., 77-96頁) またはトリオーマ (trioma) 法など) によってモノクローナル抗体を調製するために使用され得る。融合細胞を産生する技術は周知である (概括的に、Current Protocols in Immunology (1994) Coliganら、(編集) John Wiley & Sons社、New York, NY参照)。簡潔に言えば、不死の株化細胞 (典型的にはミエロ-マ) は、上記の免疫原で免疫された哺乳類から得たリンパ球 (典型的には脾細胞) と融合され、得られた融合細胞の培養上清は、本発明のポリペプチドに結合するモノクローナル抗体を産生する融合細胞を同定するためにスクリーニングされる。

【0301】

リンパ球および不死化株化細胞を融合するために使用される多くの周知の手順のいずれも、本発明のポリペプチドに対するモノクローナル抗体を生じる目的で適用され得る (例えば、Current Protocols in Immunology、前出; Galfreら、Nature 266:55052 (1977); R. H. Kenneth、Monoclonal Antibodies: A New Dimension In Biological Analyses中、Plenum Publishing Corp., New York, New York (1980); および Lerner、Yale J. Biol. Med. 54:387-402 (1981) 参照)。さらに、当業者は、同様に有用であるこのような方法の多くのバリエーションがあることを理解するだろう。

【0302】

モノクローナル抗体分泌融合細胞の調製の代わりに、本発明のポリペプチドに対するモノクローナル抗体が、当該ポリペプチドとの組み換えコンビナトリアル免疫グロブリンライブラリー (例えば、抗体ファージディスプレイライブラリー) をスクリーニングし、それにより、当該ポリペプチドに結合する免疫グロブリンライブラリーメンバーを単離することによって、同定および単離され得る。ファージディスプレイライブラリーを生じ、かつスクリーニングするためのキットは、市販されている (例えば、ファルマシア社 (Pharmacia) 組み換えファージ抗体システム、カタログ番号 27-9400-01; およびストラタジーン社 (Stratagene) SurfZAP (商標) ファージディスプレイキット、カタログ番号 240612)。その上、方法および試薬の例、特に、抗体のディスプレイライブラリーの生成およびスクリーニングにおける使用において受け入れられる例は、例えば、米国特許第 5,223,409号明細書、国際公開第 92/18619号パンフレット、国際公開第 91/17271号パンフレット、国際公開第 92/20791号パンフレット、国際公開第 92/15679号パンフレット、国際公開第 93/01288号パンフレット、国際公開第 92/01047号パンフレット、国際公開第 92/09690号パンフレット、国際公開第 90/02809号パンフレット; Fuchsら、Bio/Technology 9:1370-1372 (1991); Hayら、Hum. Antibod. Hybridomas 3:81-85 (1992); Huseら、Science 246:1275-1281 (1989); および Griffithsら、EMBO J. 12:725-734 (1993) において見出され得る。

【0303】

その上、ヒト部分および非ヒト部分の両方を含み、標準的な組み換えDNA技術を使用して産生され得る組み換え抗体 (キメラのモノクローナル抗体およびヒト化モノクローナル抗体など) は、本発明の範囲内である。このようなキメラのモノクローナル抗体およびヒト化モノクローナル抗体は、当該技術分野で公知の組み換えDNA技術によって産生され得る。

【0304】

通常は、本発明の抗体 (例えば、モノクローナル抗体) は、標準的な技術 (アフィニティークロマトグラフィーまたは免疫沈降など) によって本発明のポリペプチド (例えば P

10

20

30

40

50

A X 5、T U B、S E R P I N H 1、R A D 5 1 L 1、F H O D 3 および T N R C 6 B 遺伝子のうちのいずれか1つによってコードされるポリペプチド)を単離するのに使用され得る。ポリペプチド特異的な抗体は、細胞からの自然のポリペプチドの精製および宿主細胞で発現した組み換えによって産生されたポリペプチドの精製を促進できる。さらに、本発明のポリペプチドに特異的な抗体は、当該ポリペプチドの存在量および発現様式を評価するために、(例えば、細胞溶解物、細胞上清、または組織試料中の)当該ポリペプチドを検出するのに使用され得る。抗体は、臨床試験過程の一部として、例えば、所定の治療計画の有効性の判定などのため、組織中のタンパク質レベルをモニターするために診断的に使用され得る。抗体は、その検出を促進するための検出可能な物質と結合され得る。検出可能な物質の例は、様々な酵素、補欠分子族、蛍光物質、発光物質、生物発光物質、および放射性物質を含む。適した酵素の例は、西洋わさびペルオキシダーゼ、アルカリホスファターゼ、ガラクトシダーゼ、またはアセチルコリンエステラーゼを含む。適した補欠分子族複合体の例は、ストレプトアビジン/ビオチンおよびアビジン/ビオチンを含む。適した蛍光物質の例は、ウンベリフェロン、フルオレセイン、フルオレセインイソチオシアネート、ローダミン、ジクロロトリアジニルアミンフルオレセイン(dichlorotriazinylamine fluorescein)、塩化ダンシルまたはフィコエリトリンを含む。発光物質の例は、ルミノールを含む。生物発光物質の例は、ルシフェラーゼ、ルシフェリン、およびエクオリンを含み、ならびに、適した放射性物質の例は¹²⁵I、¹³¹I、³⁵Sまたは³Hを含む。

10

20

【0305】

抗体はまた、薬理ゲノミクス分析において有用であってもよい。このような実施態様では、本発明に関する核酸によってコードされた変異タンパク質に対する抗体(少なくとも1つの本発明の多型マーカを含む核酸によってコードされた変異タンパク質など)は、改変された治療様式を必要とする個体の同定のために使用され得る。

【0306】

抗体は、さらに、疾病の活性段階などの疾病の状態、またはタンパク質の機能に関連した疾病(特に乳癌)に対する素因のある個体における変異タンパク質の発現の評価のために有用であり得る。本願明細書記載の少なくとも1つの多型マーカまたはハプロタイプを含む核酸によってコードされた本発明の変異タンパク質に特異的な抗体は、変異タンパク質の存在のスクリーニング、例えば、変異タンパク質の存在によって示された乳癌への素因についてのスクリーニング、に使用され得る。

30

【0307】

抗体は他の方法で使用され得る。従って、抗体は、電気泳動の移動度、等電点、トリプシンまたは他のプロテアーゼ消化による分析と組み合わせて、タンパク質(本発明の変異タンパク質など)を評価するための診断の手段として、または当業者に知られた他の物理的なアッセイにおける使用において有用である。抗体はまた、組織適合試験に使用してもよい。そのような1つの実施態様では、特定の変異タンパク質は、特定の組織型の発現に関連し、変異タンパク質に特異的な抗体は、次いで、特定の組織型の同定に使用され得る。

【0308】

変異タンパク質を含むタンパク質の細胞内局在はまた、抗体を使用して判定され得、様々な組織の細胞のタンパク質の異常な細胞内局在を評価するために適用され得る。このような使用は遺伝子検査に適用され得るが、特定の治療様式をモニターすることにも適用され得る。治療が、変異タンパク質の発現レベルもしくは存在または変異タンパク質の異常な組織分布もしくは発生上の発現を矯正することを目的にしている場合、当該変異タンパク質またはその断片に特異的な抗体は治療有効性をモニターするために使用され得る。

40

【0309】

抗体は、さらに、例えば、結合分子またはパートナーへの、変異タンパク質の結合をブロックすることによる変異タンパク質の機能の阻害において有用である。このような使用はまた、治療が変異タンパク質の機能の阻害に関連する治療上の構成においても適用され

50

得る。抗体は、例えば、結合をブロックまたは競合的に阻害し、それによりタンパク質の活性を調節（すなわち、刺激または拮抗）するのに使用され得る。抗体は、特定の機能に必要とされる部位を含む特定のタンパク質断片、または、細胞または細胞膜に関連するインタクトなタンパク質に対して調製され得る。in vivoにおける投与について、抗体はさらなる治療上の負荷量（放射性核種、酵素、免疫原性エピトープ、または細菌毒素（ジフテリアまたは、リシンなどの植物毒素）を含む細胞毒など）に関連してもよい。in vivoでの抗体またはその断片の半減期は、ポリエチレングリコールとの結合によるペグ化によって増加されてもよい。

【0310】

本発明は、さらに、本願明細書記載の方法における抗体を使用するためのキットに関する。これは、試験試料における変異タンパク質の存在を検出するためのキットを含むが、これに限定されない。1つの好ましい実施態様は、抗体（標識された抗体または標識できる抗体など）および生物学的試料中の変異タンパク質を検出するための化合物もしくは薬剤、試料中の変異タンパク質の量もしくは存在および/もしくは不存在を判定する手段、ならびに試料中の変異タンパク質の量を標準物質と比較する手段、ならびに当該キットの使用についての説明書を含む。

10

【0311】

本発明は、これから下記の非限定的な実施例によって例証される。

【実施例】

【0312】

（乳癌のリスクと関連している9つの染色体の部位上の変異の同定）

乳癌の感受性に関連している、共通のSNPの対立遺伝子について広く検索するため、発明者らは、イルミナ ヒトHap300マイクロアレイ技術を使用して、ゲノム全体でのSNP関連研究を行った。遺伝子型の同定は、約1835のアイスランド人の乳癌患者および30,320の対照について行った。品質管理チェックに通らなかったSNPを除いた後、311,524のSNPが残り、これらを乳癌との関連について試験した。この結果は、ゲノムの制御方法を使用して、個体間の関連性および潜在的な集団の層別化について調整された[DevlinおよびRoeder、(1999)、Biometrics、55、997-1004]（方法の項を参照）。シグナルは、P値によって順位付けした。ゲノム上の種々の部位から、中でも注目すべきは染色体9、10（2つの領域AおよびB）、11（2つの領域AおよびB）、14、18および22（2つの領域AおよびB）から得たSNPのセットは、最も高い順位を示した。これらのマーカーを含む興味ある領域（LDブロックと呼ぶ）は、LDブロック表に定義されている；本願明細書の全ての座標は、NCBI ビルド36からのものである。この領域のイルミナSNPの遺伝子型の同定から得た結果は、表1に示されている。

20

30

【0313】

高順位の9つのマーカー；例えばrs2005154（C09）、rs2184380（C10A）、rs2224696（C10B）、rs2242503（C11A）、rs12291026（C11B）、rs999737（C14）、rs9956546（C18）、rs11912922（C22A）およびrs6001954（C22B）に関連するシグナルをさらに検討するため、発明者らは、これらのSNPについて、ケンタウルス（Centaurus）アッセイを作製し、検証した。当該SNPアッセイは、約450のアイスランド人の乳癌患者および5000を超える対照のさらなる試料の遺伝子型を同定するために使用した。アイスランド人の乳癌患者（2280）および対照（35650）についてのイルミナアッセイおよびケンタウルスアッセイからの合わせたデータを表2に示す。すべてのSNPは、合わせたアイスランド人のコホートにおいて乳癌との有意な関連を示し、これは表1の結果を伴う当初の観察を裏付ける。

40

【0314】

（方法）

（患者および対照の選択）

50

研究の被験者からの血液試料および医療情報の収集は、インフォームドコンセントおよびヘルシンキ宣言に従った倫理審査委員会の承認を得て行った。

【0315】

アイスランド：乳癌診断の記録は、アイスランド人の癌登録簿（ICR）から入手した。当該ICRは、1955年1月1日から、アイスランドで診断された、浸潤性乳腺腫瘍および乳管癌または非浸潤性小葉癌の全ての症例を含む。2006年12月末までにICRに記入された診断を受けたアイスランドに住んでいる全ての有病の症例は、本研究への参加適格がある。当該ICRは、この期間に診断された4785の個体の記録を含んでいた。同意、試料および良好な遺伝子型は、約2280の患者から入手した。これらのうち、遺伝子型は、1835の患者についてのイルミナ Hap300チップ、および445の患者についてのケンタウルスアッセイに由来した。およそ35500のアイスランド人の対照は、ディコーデジェネティクス（deCODE genetics）での進行中のイルミナに基づくゲノム全体での関連研究から選択された個体からなっていた。ICRにおける乳癌の診断を受けた個体は除外した。男性および女性の両方が含まれていた。アイスランド人の対照（および下記の外国の再現対照群）において、表6に記載されたSNPの頻度について、性別間で有意差はなかった。従って、発明者らは、これらの対照群は、検討下でのSNPの集団頻度の適切な表示を与えると考えた。

10

【0316】

スペイン：スペイン人の研究患者は、2006年3月から2007年8月の間に、サラゴサ病院の癌科から動員された。遺伝子型の同定は、十分に、約825の患者において行

20

【0317】

スウェーデン：スウェーデン人の試料セットは、家族性患者の系列および継続患者の系列からなっていた。家族性乳癌の動員群は、乳癌の家族歴の検討についてカロリンスカ大学病院（ストックホルム）の癌遺伝子のカウンセリングクリニックを紹介された347の乳癌患者からなっていた。各患者は、別個の家庭の出身である。BRCA変異のスクリーニングについての現在の判断基準に適合する全ての症例は、陰性の結果が出た。継続性乳癌の動員群は、フディング（Huddinge）病院およびゼーダー（Soeder）病院の癌科（南ストックホルムの集団を網羅）で、1998年10月から2000年5月まで、原発性湿潤性乳癌について外科的に治療された、482の継続的に動員された患者からなっていた。家族歴は、動員のための患者の選択において考慮しなかった。対照は、両方の性別の1302の血液提供者および448の癌のない個体であった。全ての対照は、カロリンスカ大学病院（ストックホルム）で収集した。いずれの試験されたSNPについて、家族性の系列と継続性の系列との間に有意な不均一性のエビデンスはなかった。

30

【0318】

オランダ：2005～2006年の期間に乳癌と診断された女性の患者は、総合癌センター（Comprehensive Cancer Centre、オランダ、東ナイメーヘン）が保有する局所性癌の登録簿から選択された。この癌センターは、集団に基づく癌登録簿を保有し、オランダ東部（130万の住民の領域）を網羅する。70歳よりも前に乳癌と診断された全ての患者は、本研究に参加するよう、招かれた。東総合癌センターは、癌登録簿における全ての患者についての臨床上のデータおよび病理学上のデータを収集した。これらの標準的な癌登録簿のデータには、患者が治療された病院の医療ファイルからの抽出による、より詳細なデータが追加されている。対照は、ラートボウト（Radboud）大学ナイメーヘン医療センターによる、2002～2003年の調査において収集された。この調査（ナイメーヘン生物医学研究）は、ナイメーヘンの集団における年齢を層別化した確率標本に基づく。この群から、患者集団に対する頻度によって年齢をマッチさせた、2034の対照個体を選択し、遺伝子型を同定した。

40

【0319】

50

CGEMS (The Cancer Genetics Marker of Susceptibility、感受性の癌遺伝子マーカー) : これは、米国国立癌研究所の計画であり、この米国国立癌研究所は、イルミナプラットフォームを使用した約530,000のSNPで遺伝子型を同定された1145の患者および1142の対照に基づく乳癌の感受性についてのゲノム全体でのSNPの関連研究について公共にデータを公開している。これらのデータは、<https://caintegrator.nci.nih.gov/cgems/>で入手できる。

【0320】

(遺伝子型の同定)

約1840のアイスランド人患者および3020の対照は、イルミナ Hap300SNPアレイ上で、以前に[Staceyら、(2007)、Nat Genet 39:865-9]で記載されたように遺伝子型を同定された。結果を表1に示す。全ての他の遺伝子型の同定は、表2および3に示されたSNPのために作製された、ナノゲン(Nanogen)ケンタウルスアッセイ[Kutyavinら、(2006)、Nucleic Acids Res、34、e128]を使用して行われた。ケンタウルスSNPアッセイは、HapMap CEU試料の遺伝子型の同定、および公表されたデータとの遺伝子型の比較によって検証された。アッセイは、HapMapデータと、1.5%のミスマッチを示した場合、拒絶された。アイスランド人の症例試料の約10%がイルミナおよびナノゲンプラットフォームの両方で遺伝子型を同定され、観察されたミスマッチ率は、0.5%よりも低かった。全ての遺伝子型の同定は、ディコーディジェネティクスの設備で行われた。全ての物理的な座標は、NCBI ビルド36に従って与えられた。

10

20

【0321】

(イルミナ遺伝子型の同定)

DNA試料について、国際HapMap計画の第I相に由来する317,503のSNPを含むイルミナ インフィニウム(Infinium) ヒトHap300 SNPビーズマイクロアレイ(イルミナ社(Illumina)、米国、カリフォルニア州、サンディエゴ)についての製造者の説明書に従い、遺伝子型を同定した。このチップは、 $r^2 = 0.8$ の共通のSNPについてのユタ州のCEPH(CEU)HapMap試料の約75%のゲノムの網羅を与える(BarrettおよびCardon、(2006)、Nat Genet、38、659-62)。チップ上のSNPの全体数のうち、5979は、これらが単形性(すなわち、合わせた患者および対照のセットにおける少ない対立遺伝子頻度が0.001未満であった)であったためか、低い($< 95\%$)収率を有していたためか、または対照においてハーディ・ワインベルグ平衡からの非常に有意な歪みを示したためか($P < 1 \times 10^{-10}$)、のいずれかの理由で不適であるとみなした。これらの問題のSNPの全てを分析から除去した。従って、311,524のSNPが、関連分析において使用された。SNPの98%を下回る全体のコールレート(call rate)を有するいずれのチップもまた、ゲノム全体での関連分析から除外した。

30

【0322】

(ケンタウルスSNP遺伝子型同定)

ケンタウルスアッセイ[Kutyavinら、(2006)、Nucleic Acids Res、34、e128]を、すべての9つの変異体:例えばrs2005154、rs2184380、rs2224696、rs2242503、rs12291026、rs999737、rs9956546、rs11912922およびrs6001954について設計し、HapMap CEU試料の遺伝子型を同定し、遺伝子型を公表されたデータと比較することで検証した。当該分析は、HapMapデータについて $< 1.5\%$ のミスマッチを与えた。表2は、これらのSNPについての配列構成についての配列番号の参照、例えば配列番号:1、配列番号:2、配列番号:3、配列番号:4、配列番号:5、配列番号:6、配列番号:7、配列番号:8および配列番号:9を示す。

40

【0323】

(統計学的方法)

50

発明者らは、複合モデルを仮定して、すなわちヒトが保有する2つの対立遺伝子の相対リスクが乗算で表されると仮定してSNP対立遺伝子のオッズ比(OR)を算出した。保有者頻度よりも対立遺伝子の頻度が、当該マーカースに表われた。関連したP値は、NEMOソフトウェアパッケージにおいて実行されるように、標準的な尤度比のカイ2乗統計量で算出した[Gretarsdottirら、(2003)、Nat Genet、35、131-8]。信頼区間は、ORの評価が対数正規分布を有すると仮定して算出した。

【0324】

いくらかのアイスランド人の患者および対照は、群内および群間の両方で関連しており、カイ2乗検定統計量が、1超の平均値および 0.675^2 超の中央値を有する原因となっている。発明者らは、関連性および潜在的な集団の層別化を説明する、ゲノム全体でのSNPセットについての観察されたカイ2乗統計量の平均値を算出することによって、ゲノム制御方法を使用してアイスランド1についての膨張因子を評価した[DevlinおよびRoeder、(1999)、Biometrics、55、997-1004]。アイスランド2については、マーカースのゲノム全体のセットについて分類されておらず、当該膨張因子は、アイスランド人の系統学を通じて遺伝子型をシミュレートすることで評価した[Grantら、(2006)、Nat Genet、38、320-3]。評価された膨張因子は、アイスランド1では1.105、アイスランド2では1.11であった。アイスランド1およびアイスランド2の試料セットの共同解析において評価された膨張因子は1.08であり、これをシミュレーションによって得た。

【0325】

全てのP値は、両側として報告した。

【0326】

10

20

【表 4】

表1：アイスランド人の症例および対照試料において乳癌に関連して見出された、イルミナ インフィニウム ヒトHap300 チップから得られたSNP。この表は、SNPのrs-名称、乳癌との関連についてのP値、関連するリスク、罹患者および対照における関連する対立遺伝子の数および頻度、リスクのある対立遺伝子の同一性、関連するマーカーの染色体およびその染色体上（NCBI ビルド36）の位置、ならびに当該SNPの隣接配列を示す配列番号の参照を示す。対立遺伝子コードは1=A；2=C；3=G；4=Tである。

SNP	P値	OR	罹患者数	罹患者頻度	対照	対照頻度	対立 遺伝子	染色体	ビルド36に おける位置	配列 番号
rs2800256	0.1454055	1.0545	1835	0.323706	30311	0.312197	4	C09	36,806,714	12
rs2851696	0.1540843	1.0513	1836	0.630174	30363	0.618434	1	C09	36,809,706	15
rs2851695	0.4950720	1.024	1833	0.603382	30308	0.597697	4	C09	36,810,262	14
rs2800266	0.0940407	1.0605	1833	0.625477	30301	0.61163	1	C09	36,810,604	13
rs7848675	0.0424981	1.0789	1837	0.715024	30383	0.699306	1	C09	36,814,086	18
rs2800252	0.5399746	1.0477	1837	0.948013	30370	0.94567	3	C09	36,820,616	11
rs4880018	0.8570986	1.0154	1837	0.958628	30345	0.958016	2	C09	36,833,844	17
rs3780136	0.0670526	1.0972	1831	0.873839	30243	0.863258	2	C09	36,835,973	16
rs7852051	0.0515094	1.077	1837	0.279532	30380	0.264845	4	C09	36,841,358	19
rs1051193	0.2874298	1.056	1837	0.128198	30380	0.122235	4	C09	36,846,268	10
rs2005154	0.0009612	1.2026	1837	0.902831	30353	0.885398	4	C09	36,846,828	1
rs2031557	0.0127968	1.0884	1831	0.513381	30326	0.492201	1	C10	8,651,020	32
rs1325881	0.5529599	1.0381	1837	0.080022	30376	0.077314	1	C10	8,653,900	28
rs1079565	0.6949928	1.0221	1837	0.104246	30380	0.102222	2	C10	8,656,133	22
rs1125575	0.0423179	1.0728	1834	0.41385	30330	0.396917	3	C10	8,659,854	25
rs1090540	0.7733530	1.016	1837	0.106696	30378	0.105191	2	C10	8,663,595	23
rs7921734	0.3601882	1.0528	1836	0.103486	30381	0.098812	2	C10	8,668,422	40
rs7069464	0.4623737	1.0515	1836	0.067266	30375	0.064181	3	C10	8,677,428	38
rs4747806	0.0000274	1.1555	1837	0.595808	30379	0.560568	4	C10	8,678,241	35
rs1090541	0.0328102	1.0772	1837	0.39657	30382	0.378925	2	C10	8,685,729	24
rs1241208	0.0297294	1.1385	1835	0.092916	30225	0.082548	1	C10	8,691,311	27
rs6602329	0.0004012	1.1281	1836	0.541939	30380	0.511899	4	C10	8,700,101	36
rs4454616	0.0008570	1.1202	1837	0.47877	30369	0.450542	1	C10	8,705,153	33
rs4584486	0.6199739	1.0318	1837	0.079477	30388	0.077218	4	C10	8,708,704	34
rs1050836	0.1309007	1.0717	1837	0.167665	30388	0.15822	1	C10	8,714,527	20
rs7083359	1.76E-06	1.1923	1837	0.704954	30380	0.6671	2	C10	8,722,685	39
rs827402	0.7315722	1.0268	1829	0.051668	30287	0.050385	4	C10	8,726,332	41
rs827405	0.0002462	1.1624	1837	0.228634	30375	0.203177	4	C10	8,729,009	42
rs1334553	0.6242464	1.0218	1837	0.183723	30377	0.180515	3	C10	8,733,175	30
rs706771	0.0249670	1.0848	1837	0.33043	30365	0.312679	1	C10	8,736,452	37
rs1334549	4.86E-08	1.2428	1834	0.773991	30247	0.733726	1	C10	8,744,989	29
rs2184380	2.13E-08	1.2496	1836	0.773693	30351	0.732332	3	C10	8,754,080	2

【 0 3 2 7 】

【表 5】

SNP	P値	OR	罹患数	罹患頻度	対照	対照頻度	対立 遺伝子	染色体	ビルド36に おける位置	配列 番号
rs1537603	0.0063633	1.0978	1833	0.455265	30293	0.432229	3	C10	8,774,301	31
rs1225037	5.32E-06	1.1696	1837	0.593087	30380	0.554789	4	C10	8,802,201	26
rs1075214	0.4819822	1.0344	1837	0.147795	30375	0.143588	1	C10	8,815,736	21
rs4749829	0.9501223	1.0042	1828	0.068107	30108	0.067839	3	C10	9,078,140	61
rs1912584	0.2426291	1.0945	1835	0.052589	30237	0.048269	4	C10	9,081,164	56
rs1775551	0.2937532	1.0477	1837	0.82417	30336	0.817313	1	C10	9,093,049	55
rs962993	0.2774384	1.0386	1836	0.393519	30288	0.384525	4	C10	9,093,138	67
rs7916897	0.2492789	1.0426	1832	0.674945	30120	0.66572	3	C10	9,101,110	64
rs2197415	0.2826170	1.0382	1834	0.39313	30349	0.38423	4	C10	9,102,862	57
rs1444788	0.2798396	1.0491	1837	0.82417	30368	0.81711	4	C10	9,104,126	54
rs2476175	0.8776101	1.0064	1837	0.217746	30376	0.216668	4	C10	9,114,133	58
rs2797266	0.3487364	1.0331	1836	0.396786	30364	0.389013	2	C10	9,129,166	59
rs860419	0.2755681	1.0382	1835	0.432425	30218	0.423258	4	C10	9,132,868	65
rs1276604	0.0466943	1.119	1835	0.104087	30349	0.094056	2	C10	9,133,679	52
rs1079569	0.1235447	1.064	1810	0.767956	30063	0.756711	2	C10	9,148,549	44
rs7067548	0.1945938	1.0599	1834	0.829062	30143	0.820655	1	C10	9,150,794	62
rs961945	0.1230478	1.0542	1834	0.445747	30375	0.432741	3	C10	9,166,649	66
rs1278142	0.0040593	1.1397	1837	0.173925	30379	0.15593	3	C10	9,167,711	53
rs2224696	0.0010448	1.1626	1837	0.170659	30377	0.150377	4	C10	9,168,781	3
rs1049094	0.4748388	1.0506	1836	0.936275	30382	0.933266	2	C10	9,183,517	43
rs1125606	0.0894290	1.0937	1837	0.120849	30372	0.111649	2	C10	9,187,738	45
rs372712	0.0755905	1.0628	1837	0.43767	30369	0.422734	1	C10	9,194,043	60
rs7898051	0.6463330	1.0192	1837	0.788242	30318	0.785045	3	C10	9,204,690	63
rs1149923	0.0228098	1.0815	1796	0.496381	28769	0.476815	4	C10	9,222,722	46
rs1149928	0.4126463	1.0305	1836	0.314815	30374	0.308372	3	C10	9,231,471	47
rs1149938	0.9114461	1.0038	1834	0.493184	30279	0.492239	4	C10	9,240,108	48
rs1149945	0.8799003	1.0052	1837	0.43767	30368	0.436397	3	C10	9,247,627	49
rs1149952	0.4948711	1.0248	1802	0.358213	29500	0.35261	2	C10	9,256,962	50
rs1242987	0.9668504	1.0014	1835	0.415259	30374	0.414911	3	C10	9,263,430	51
rs1540210	0.4912002	1.0262	1836	0.288126	30368	0.282847	3	C11	74,887,046	78
rs1790144	0.1544257	1.0705	1837	0.151062	30376	0.14253	4	C11	74,889,549	79
rs1793399	0.0609546	1.0906	1837	0.164671	30385	0.153085	2	C11	74,898,153	80
rs618202	0.0017623	1.1675	1837	0.142896	30362	0.124959	4	C11	74,916,133	86
rs633473	0.0053293	1.113	1836	0.740741	30362	0.71965	2	C11	74,930,524	87
rs1229102	0.0007803	1.1935	1837	0.12466	30384	0.106602	3	C11	74,932,878	5
rs1938800	0.1182986	1.0798	1836	0.141885	30372	0.132787	1	C11	74,933,770	81
rs662279	0.6695122	1.0159	1837	0.303756	30370	0.300428	2	C11	74,939,444	90

10

20

30

40

【 0 3 2 8 】

【表 6】

SNP	P値	OR	罹患数	罹患頻度	対照	対照頻度	対立遺伝子	染色体	ビルド36における位置	配列番号
rs4945015	0.0274668	1.2223	1800	0.965556	29750	0.958218	1	C11	74,944,615	82
rs608585	0.9731233	1.0015	1836	0.826797	30383	0.826581	3	C11	74,946,352	84
rs640649	0.0066177	1.1668	1837	0.104246	30371	0.090695	4	C11	74,949,499	88
rs652054	0.8326146	1.008	1836	0.714052	30368	0.712428	2	C11	74,953,180	89
rs609309	0.2081718	1.0523	1837	0.774633	30209	0.765616	3	C11	74,964,989	85
rs585433	0.3707179	1.0312	1834	0.43675	30358	0.429211	1	C11	74,970,249	83
rs3794018	0.1981440	1.0455	1835	0.414169	30364	0.403422	2	C11	8,054,010	74
rs4758040	0.0129157	1.0895	1834	0.588604	30343	0.567709	4	C11	8,057,749	76
rs4578424	0.1068004	1.0722	1837	0.81301	30373	0.802176	2	C11	8,064,861	75
rs1881237	0.0243448	1.1965	1835	0.955041	30380	0.946675	4	C11	8,067,055	71
rs2242503	0.0001113	1.1522	1837	0.697877	30352	0.667205	2	C11	8,075,048	4
rs2242504	0.0277748	1.1826	1836	0.950436	30375	0.941909	2	C11	8,079,519	73
rs2049684	0.0030921	1.1133	1837	0.683179	30387	0.659509	3	C11	8,127,695	72
rs1083998	0.0239367	1.0939	1837	0.765378	30383	0.748873	1	C11	8,137,470	68
rs1104179	0.0101003	1.0987	1837	0.692433	30363	0.672035	4	C11	8,168,818	69
rs1881229	0.3637320	1.0428	1837	0.840501	30372	0.834798	3	C11	8,184,540	70
rs7951027	0.7460144	1.0117	1837	0.336962	30383	0.334365	1	C11	8,190,452	77
rs737387	0.7483800	1.011	1835	0.452316	30359	0.449603	1	C14	68,037,135	111
rs757369	0.2989297	1.0362	1836	0.559913	30378	0.551139	3	C14	68,043,299	112
rs1541390	0.6290307	1.021	1837	0.195427	30382	0.192186	3	C14	68,049,682	98
rs1023529	0.3683977	1.0323	1834	0.366957	30379	0.359607	2	C14	68,059,561	92
rs8012659	0.6278097	1.0364	1837	0.056886	30383	0.054998	4	C14	68,060,343	114
rs916962	0.3596030	1.0387	1836	0.215959	30363	0.209597	1	C14	68,063,331	115
rs1162127	0.0099335	1.1003	1837	0.707403	30367	0.687226	1	C14	68,066,219	96
rs1013444	0.6226353	1.0171	1837	0.418345	30370	0.414225	3	C14	68,067,388	91
rs727392	0.6708335	1.0309	1837	0.060425	30382	0.058719	2	C14	68,071,554	110
rs1115875	0.0561406	1.0802	1835	0.234877	30355	0.221298	2	C14	68,076,508	94
rs1710581	0.8079818	1.0104	1837	0.803484	30382	0.80184	4	C14	68,078,080	99
rs2842331	0.2212033	1.0527	1808	0.213772	29730	0.205264	2	C14	68,080,443	107
rs2158358	0.3127803	1.0405	1835	0.251226	30275	0.243832	2	C14	68,081,725	101
rs2842329	0.3199784	1.04	1827	0.250137	30229	0.242846	2	C14	68,087,348	106
rs2244814	0.3779656	1.0343	1822	0.276345	29822	0.26965	3	C14	68,095,623	102
rs2525503	0.0007823	1.1213	1837	0.550898	30365	0.522427	3	C14	68,098,356	105
rs1710583	0.0138473	1.1027	1837	0.247142	30373	0.229398	1	C14	68,098,494	100
rs1115875	0.8363170	1.0071	1837	0.468699	30376	0.466948	1	C14	68,101,178	95
rs999737	0.0017229	1.1326	1836	0.766885	30385	0.743887	2	C14	68,104,435	6
rs6573841	0.3860408	1.0386	1837	0.816277	30382	0.810529	2	C14	68,107,274	109

10

20

30

40

【表 7】

SNP	P値	OR	罹患数	罹患頻度	対照	対照頻度	対立 遺伝子	染色体	ビルド36に おける位置	配列 番号
rs8009944	0.0319682	1.0851	1833	0.28096	30307	0.264757	2	C14	68,109,341	113
rs2285883	0.2432548	1.0508	1836	0.203431	30344	0.195525	1	C14	68,113,723	104
rs1290999	0.0505573	1.0811	1835	0.243869	30359	0.229784	1	C14	68,115,629	97
rs2256224	0.3955036	1.0321	1836	0.704793	30374	0.698179	1	C14	68,122,500	103
rs1048381	0.0804077	1.1317	1837	0.063963	30381	0.056943	1	C14	68,125,552	93
rs4899246	0.4842718	1.0287	1836	0.773148	30366	0.768145	1	C14	68,129,955	108
rs559325	0.5198271	1.0224	1837	0.581383	30369	0.575982	4	C18	32,116,417	121
rs1296146	0.7436940	1.0159	1833	0.14539	30306	0.143437	3	C18	32,119,988	117
rs4799863	0.5350597	1.0215	1837	0.434404	30382	0.429185	3	C18	32,127,413	119
rs9956546	0.0000702	1.2177	1836	0.144608	30375	0.121909	1	C18	32,136,446	7
rs680423	0.1990102	1.0817	1835	0.917711	30339	0.911582	4	C18	32,136,736	122
rs7243619	0.2091339	1.0454	1837	0.642352	30371	0.632083	4	C18	32,139,089	125
rs686784	0.5894600	1.0287	1837	0.120033	30374	0.117074	4	C18	32,142,076	123
rs492392	0.0498932	1.0841	1836	0.222767	30363	0.209103	1	C18	32,142,779	120
rs7240720	0.0271451	1.1376	1837	0.097441	30354	0.086677	3	C18	32,142,893	124
rs1272559	0.9364776	1.0031	1837	0.268645	30383	0.268045	3	C18	32,142,971	116
rs4531816	0.1177158	1.0929	1837	0.903375	30374	0.895338	4	C18	32,143,579	118
rs114607	0.4201611	1.0281	1835	0.426158	30375	0.419391	1	C22	38,706,329	126
rs2958658	0.7075628	1.0145	1837	0.731355	30360	0.728524	4	C22	38,718,953	127
rs3021222	0.0121019	1.0919	1837	0.62847	30368	0.607729	3	C22	38,722,550	128
rs3021274	0.0110828	1.0932	1837	0.631464	30372	0.610497	2	C22	38,725,030	129
rs1191292	0.0005138	1.1298	1837	0.383234	30385	0.35483	4	C22	38,733,117	8
rs6001722	0.6476742	1.0178	1836	0.739379	30361	0.735961	3	C22	38,758,379	130
rs9623117	0.0169431	1.1073	1837	0.808383	30374	0.792092	4	C22	38,782,065	134
rs6001749	0.0153071	1.1091	1837	0.808383	30382	0.791834	1	C22	38,805,272	131
rs6001762	0.0654330	1.085	1836	0.824891	30380	0.812788	2	C22	38,825,419	132
rs6001794	0.1618358	1.0563	1837	0.255035	30382	0.244767	4	C22	38,859,361	133
rs1108997	0.0832280	1.0697	1837	0.260751	30365	0.247966	4	C22	38,873,554	136
rs9611280	0.6602329	1.0241	1837	0.890038	30384	0.887688	3	C22	38,882,065	169
rs1108997	0.8500487	1.0072	1837	0.272727	30378	0.271298	4	C22	38,892,173	137
rs9611286	0.9363221	1.0031	1834	0.272083	30286	0.271479	2	C22	38,914,908	170
rs8137636	0.0133127	1.0931	1836	0.669662	30296	0.64969	1	C22	38,929,577	166
rs138021	0.8870477	1.0054	1837	0.272455	30365	0.271382	3	C22	38,942,920	147
rs739182	0.0242520	1.0817	1835	0.397275	30343	0.378638	4	C22	38,945,222	163
rs138027	0.8504863	1.0072	1837	0.272727	30372	0.271303	3	C22	38,946,058	148
rs138039	0.0145794	1.0918	1836	0.66939	30368	0.649681	3	C22	38,951,234	149
rs1305673	0.6304407	1.0295	1837	0.087099	30385	0.084812	4	C22	38,976,018	140

10

20

30

40

【 0 3 3 0 】

【表 8】

SNP	P値	OR	罹患数	罹患頻度	対照	対照頻度	対立遺伝子	染色体	ビルド 36 における位置	配列番号
rs9611302	0.6750492	1.023	1837	0.890038	30378	0.887797	3	C22	38,980,230	171
rs733381	0.1003993	1.0658	1837	0.262657	30348	0.250494	3	C22	38,999,594	162
rs139921	0.0014652	1.1167	1834	0.607688	30282	0.581088	3	C22	39,056,708	150
rs470113	0.0017857	1.128	1835	0.27248	30377	0.249268	3	C22	39,059,560	155
rs1248469	0.0009138	1.1368	1836	0.270153	30368	0.24562	1	C22	39,066,418	139
rs8135371	0.4303623	1.037	1837	0.164126	30370	0.159203	2	C22	39,087,174	164
rs9611324	0.0008177	1.1381	1836	0.270697	30357	0.245924	2	C22	39,111,108	172
rs2072856	0.9278854	1.0031	1837	0.537289	30365	0.536522	3	C22	39,146,695	152
rs5757949	0.0045246	1.1225	1836	0.781863	30315	0.761521	4	C22	39,150,097	158
rs9306345	0.0014298	1.1507	1837	0.189439	30377	0.168812	3	C22	39,162,321	167
rs4507196	0.0163338	1.0897	1834	0.662214	30343	0.642735	1	C22	39,179,650	154
rs8135759	0.6212354	1.017	1836	0.535948	30354	0.531759	2	C22	39,188,950	165
rs9607721	0.0013221	1.1517	1837	0.189984	30380	0.16919	3	C22	39,192,066	168
rs6001932	0.0170002	1.1289	1837	0.133097	30380	0.119717	1	C22	39,207,581	161
rs3827382	0.0210574	1.1247	1835	0.132425	30372	0.119502	1	C22	39,211,349	153
rs5750948	0.5730917	1.0346	1837	0.914262	30353	0.911557	1	C22	39,216,690	156
rs9611367	0.0038990	1.1159	1837	0.731083	30360	0.708992	4	C22	39,244,895	173
rs6001954	0.0053978	1.1537	1829	0.131219	30092	0.115762	3	C22	39,251,626	9
rs5758001	0.0228856	1.0817	1833	0.423622	30282	0.40458	1	C22	39,276,253	159
rs5758008	0.0060923	1.1104	1837	0.735166	30371	0.714283	1	C22	39,289,206	160
rs5750960	0.4684336	1.0255	1836	0.603214	30342	0.597176	2	C22	39,289,866	157
rs1022533	0.0084935	1.1059	1832	0.735535	30222	0.715489	2	C22	39,302,493	135
rs1700203	0.0175391	1.1304	1834	0.12868	30350	0.115552	4	C22	39,326,313	151
rs133036	0.4223034	1.0282	1837	0.604518	30371	0.59784	2	C22	39,342,384	141
rs133047	0.3858756	1.0467	1836	0.883715	30385	0.878937	2	C22	39,357,765	142
rs1215920	0.0507958	1.0818	1837	0.236799	30368	0.222883	2	C22	39,372,037	138
rs133067	0.5254346	1.0263	1837	0.780348	30374	0.775861	4	C22	39,404,114	143
rs133072	0.7468165	1.0113	1826	0.59885	30062	0.596151	3	C22	39,405,489	144
rs133074	0.6315619	1.0164	1834	0.505725	30331	0.501648	2	C22	39,408,419	145
rs133076	0.0627455	1.0659	1836	0.572712	30373	0.557024	2	C22	39,411,507	146

10

20

30

【 0 3 3 1 】

【表 9】

表 2 : アイスランド人の症例および対照試料において乳癌に関連して見出され、かついくつかのコーカサス人のコホートで再現することが見出された、イルミナ インフィニウム ヒト H a p 3 0 0 チップから得られたアンカー SNP。この表に示されたアイスランド人のデータは、合わせられたイルミナおよびケンタウルスアッセイを使用したデータの分析によるものである。

SNP	対立遺伝子	染色体	ビルド 36 における位置	P 値	OR	罹患数	罹患頻度	対照	対照頻度	配列番号
rs2005154	4	C09	36,846,828	5.21E-04	1.19	2282	0.90	35552	0.89	1
rs2184380	3	C10	8,754,080	1.33E-06	1.19	2282	0.77	35928	0.73	2
rs2224696	4	C10	9,168,781	3.48E-04	1.16	2277	0.17	35631	0.15	3
rs2242503	2	C11	8,075,048	9.39E-04	1.11	2280	0.69	35582	0.67	4
rs12291026	3	C11	74,932,878	3.98E-04	1.18	2280	0.12	35625	0.11	5
rs999737	2	C14	68,104,435	8.79E-05	1.15	2279	0.77	35638	0.74	6
rs9956546	1	C18	32,136,446	2.11E-06	1.23	2282	0.15	36015	0.12	7
rs11912922	4	C22	38,733,117	3.33E-05	1.14	2285	0.38	35631	0.35	8
rs6001954	3	C22	39,251,626	1.28E-03	1.16	2267	0.13	35210	0.12	9

10

20

【 0 3 3 2 】

【表 10】

表3：個々のおよび合わされたコーカサス人のコホートについて示された、染色体9、10、11、14、18および22上の9つの変異についての乳癌との関連

集団	P	OR	罹患数	罹患頻度	対照数	対照頻度	対立遺伝子	SNP	染色体	候補遺伝子
全コーカサス	1.80E-04	1.14	NA	NA	NA	NA	4	rs2005154	C09	PAX5
	8.30E-04	1.13	NA	NA	NA	NA	4	rs2005154	C09	PAX5
CGEMSを除く全コーカサス	7.05E-02	1.22	1144	0.88	1140	0.85	4	rs2005154	C09	PAX5
	1.02E-01	1.16	735	0.88	2017	0.86	4	rs2005154	C09	PAX5
アイスランド	5.21E-04	1.19	2282	0.90	35552	0.89	4	rs2005154	C09	PAX5
	9.92E-01	1.00	633	0.89	1502	0.89	4	rs2005154	C09	PAX5
スウェーデン	4.52E-01	1.07	827	0.86	1738	0.86	4	rs2005154	C09	PAX5
全コーカサス	3.00E-04	1.11	NA	NA	NA	NA	3	rs2184380	C10	
	2.50E-04	1.11	NA	NA	NA	NA	3	rs2184380	C10	
CGEMSを除く全コーカサス	8.81E-01	0.96	1144	0.79	1141	0.80	3	rs2184380	C10	
	8.80E-01	1.01	741	0.80	2019	0.80	3	rs2184380	C10	
オランダ	1.33E-06	1.19	2282	0.77	35928	0.73	3	rs2184380	C10	
	1.89E-01	0.88	459	0.75	986	0.77	3	rs2184380	C10	
アイスランド	2.47E-01	1.09	820	0.82	1717	0.80	3	rs2184380	C10	
全コーカサス	3.70E-04	1.12	NA	NA	NA	NA	4	rs2224696	C10	
	1.10E-03	1.12	NA	NA	NA	NA	4	rs2224696	C10	
CGEMSを除く全コーカサス	1.25E-01	1.19	1144	0.16	1141	0.14	4	rs2224696	C10	
	2.56E-01	0.90	738	0.12	2025	0.13	4	rs2224696	C10	
オランダ	3.48E-04	1.16	2277	0.17	35631	0.15	4	rs2224696	C10	
	3.32E-01	1.10	632	0.15	1507	0.14	4	rs2224696	C10	
スウェーデン	6.07E-02	1.18	831	0.15	1743	0.13	4	rs2224696	C10	
全コーカサス	2.20E-05	1.08	NA	NA	NA	NA	2	rs2242503	C11	TUB
	1.90E-03	1.08	NA	NA	NA	NA	2	rs2242503	C11	TUB
CGEMSを除く全コーカサス	3.93E-03	1.08	1144	0.67	1140	0.65	2	rs2242503	C11	TUB
	9.86E-01	1.00	739	0.67	2005	0.67	2	rs2242503	C11	TUB
オランダ	9.39E-04	1.11	2280	0.69	35582	0.67	2	rs2242503	C11	TUB
	2.79E-01	1.09	636	0.74	1485	0.72	2	rs2242503	C11	TUB
アイスランド	3.04E-01	1.07	814	0.67	1729	0.66	2	rs2242503	C11	TUB
スウェーデン										
全コーカサス	7.00E-04	1.13	NA	NA	NA	NA	3	rs12291026	C11	SERPINH1
	1.60E-03	1.12	NA	NA	NA	NA	3	rs12291026	C11	SERPINH1
CGEMSを除く全コーカサス	1.91E-01	1.18	1139	0.13	1138	0.11	3	rs12291026	C11	SERPINH1
	8.65E-01	0.98	737	0.12	2012	0.12	3	rs12291026	C11	SERPINH1
オランダ	3.98E-04	1.18	2280	0.12	35635	0.11	3	rs12291026	C11	SERPINH1
アイスランド										

【表 1 1】

スペイン	9.31E-02	1.18	641	0.14	1518	0.12	3	rs12291026	C11	SERPINH1
スウェーデン	7.97E-01	1.03	829	0.10	1740	0.10	3	rs12291026	C11	SERPINH1
全コーカサス	1.70E-07	1.15	NA	NA	NA	NA	2	rs999737	C14	RAD51L1
CGEMSを除く全コーカサス	3.50E-06	1.14	NA	NA	NA	NA	2	rs999737	C14	RAD51L1
CGEMS	1.31E-02	1.19	1143	0.79	1141	0.76	2	rs999737	C14	RAD51L1
オランダ	3.18E-01	1.08	740	0.78	2006	0.77	2	rs999737	C14	RAD51L1
アイスランド	8.79E-05	1.15	2279	0.77	35638	0.74	2	rs999737	C14	RAD51L1
スペイン	2.49E-02	1.21	638	0.81	1514	0.78	2	rs999737	C14	RAD51L1
スウェーデン	8.90E-02	1.13	831	0.80	1733	0.78	2	rs999737	C14	RAD51L1
全コーカサス	3.90E-06	1.17	NA	NA	NA	NA	1	rs9956546	C18	FHOD3
CGEMSを除く全コーカサス	1.50E-05	1.17	NA	NA	NA	NA	1	rs9956546	C18	FHOD3
CGEMS	1.04E-01	1.21	1144	0.13	1140	0.11	1	rs9956546	C18	FHOD3
オランダ	4.71E-01	1.07	740	0.12	2022	0.11	1	rs9956546	C18	FHOD3
アイスランド	2.11E-06	1.23	2282	0.15	36015	0.12	1	rs9956546	C18	FHOD3
スペイン	7.07E-01	1.04	458	0.14	993	0.13	1	rs9956546	C18	FHOD3
スウェーデン	2.56E-01	1.11	819	0.12	1732	0.11	1	rs9956546	C18	FHOD3
全コーカサス	5.00E-06	1.11	NA	NA	NA	NA	4	rs11912922	C22	TNRC6B
CGEMSを除く全コーカサス	3.40E-05	1.11	NA	NA	NA	NA	4	rs11912922	C22	TNRC6B
CGEMS	4.72E-02	1.16	1143	0.32	1141	0.29	4	rs11912922	C22	TNRC6B
オランダ	4.80E-01	0.96	741	0.32	2011	0.33	4	rs11912922	C22	TNRC6B
アイスランド	3.33E-05	1.14	2285	0.38	35631	0.35	4	rs11912922	C22	TNRC6B
スペイン	3.77E-02	1.16	639	0.34	1529	0.30	4	rs11912922	C22	TNRC6B
スウェーデン	6.66E-02	1.12	826	0.36	1702	0.33	4	rs11912922	C22	TNRC6B
全コーカサス	8.90E-06	1.17	NA	NA	NA	NA	3	rs6001954	C22	TNRC6B
CGEMSを除く全コーカサス	3.70E-05	1.17	NA	NA	NA	NA	3	rs6001954	C22	TNRC6B
CGEMS	9.30E-02	1.22	1101	0.11	1105	0.09	3	rs6001954	C22	TNRC6B
オランダ	1.35E-01	1.16	741	0.12	2017	0.10	3	rs6001954	C22	TNRC6B
アイスランド	1.28E-03	1.16	2267	0.13	35210	0.12	3	rs6001954	C22	TNRC6B
スペイン	6.68E-02	1.22	640	0.11	1533	0.10	3	rs6001954	C22	TNRC6B
スウェーデン	1.26E-01	1.16	825	0.12	1724	0.10	3	rs6001954	C22	TNRC6B

対立遺伝子オッズ比は、複合モデル下で算出した。すべてのP値は両側であり、関連性および他の可能性について調整されている。

【 0 3 3 4 】

10

20

30

40

【表 1 2】

表 4 : 9つのアンカー-SNPについての代用マーカー。マーカーに1Mb間隔で隣接する、HapMap CEUデータセット (<http://www.hapmap.org>) 中のアンカーマーカーに対してR²が0.2よりも大きいマーカーが選択された。示されたのは、関連するSNPの名称、当該マーカーに対するR²およびD'の値、ならびに対応するP値、ならびにNCBIビルド36における代用マーカーの位置ならびに当該マーカーの隣接配列を含む配列番号の参照である。

アンカー	代用	D'	R ²	P 値	染色体	ビルド36での位置	配列番号
rs2005154	rs2005154	1	1	-	C09	36,846,828	1
rs2005154	rs4878662	1	1	1.07E-16	C09	36,851,331	174
rs2005154	rs4880019	1	0.900826	3.20E-15	C09	36,853,450	175
rs2184380	rs2184380	1	1	-	C10	8,754,080	2
rs2184380	rs10466295	1	1	1.53E-28	C10	8,747,556	176
rs2184380	rs10508363	1	1	1.53E-28	C10	8,736,508	177
rs2184380	rs10508364	0.948936	0.856732	1.01E-22	C10	8,808,518	178
rs2184380	rs10508365	0.948936	0.856732	1.01E-22	C10	8,808,985	179
rs2184380	rs10795670	1	0.209877	4.71E-09	C10	8,765,045	180
rs2184380	rs10905411	0.874714	0.387964	3.45E-11	C10	8,678,685	181
rs2184380	rs10905414	0.870352	0.356105	3.50E-10	C10	8,680,134	182
rs2184380	rs10905415	0.91879	0.237472	4.68E-08	C10	8,680,223	183
rs2184380	rs10905430	0.844581	0.2373	1.96E-07	C10	8,704,489	184
rs2184380	rs10905437	1	0.304695	1.26E-11	C10	8,735,801	185
rs2184380	rs10905439	0.947367	0.897401	6.59E-23	C10	8,753,333	186
rs2184380	rs10905440	1	0.951417	1.10E-26	C10	8,755,150	187
rs2184380	rs10905443	1	1	1.53E-28	C10	8,757,369	188
rs2184380	rs10905444	1	1	1.53E-28	C10	8,757,533	189
rs2184380	rs10905445	1	1	1.53E-28	C10	8,758,377	190
rs2184380	rs10905446	1	1	1.53E-28	C10	8,758,390	191
rs2184380	rs10905447	1	1	1.53E-28	C10	8,758,955	192
rs2184380	rs10905454	0.895085	0.796777	4.82E-19	C10	8,807,933	193
rs2184380	rs11255764	0.885251	0.566844	9.13E-15	C10	8,676,745	194
rs2184380	rs11255776	0.844286	0.245302	2.21E-07	C10	8,701,838	195
rs2184380	rs11255777	0.847467	0.247154	1.12E-07	C10	8,702,017	196
rs2184380	rs11255778	0.848675	0.247226	9.24E-08	C10	8,703,356	197
rs2184380	rs11255779	0.840039	0.241626	3.17E-07	C10	8,703,789	198
rs2184380	rs11255790	0.937236	0.429299	1.11E-12	C10	8,722,186	199
rs2184380	rs11255795	1	1	1.93E-28	C10	8,731,572	200
rs2184380	rs11255797	1	1	1.53E-28	C10	8,737,159	201
rs2184380	rs11255800	1	1	3.91E-27	C10	8,742,581	202
rs2184380	rs11255804	0.94445	0.792459	3.17E-19	C10	8,747,207	203
rs2184380	rs11255805	1	1	1.53E-28	C10	8,747,459	204
rs2184380	rs11255820	1	1	1.53E-28	C10	8,759,525	205
rs2184380	rs11255821	1	1	9.41E-28	C10	8,759,540	206
rs2184380	rs11255822	1	1	2.45E-28	C10	8,759,613	207
rs2184380	rs11255832	1	0.949367	2.82E-26	C10	8,770,466	208
rs2184380	rs11255836	1	1	1.53E-28	C10	8,773,212	209
rs2184380	rs11255840	1	0.95	4.16E-26	C10	8,779,199	210
rs2184380	rs11255858	0.948936	0.856732	1.01E-22	C10	8,790,639	211
rs2184380	rs11255862	0.948936	0.856732	1.01E-22	C10	8,793,888	212
rs2184380	rs11255869	0.948925	0.855817	1.27E-22	C10	8,801,802	213
rs2184380	rs11255870	0.948936	0.856732	1.01E-22	C10	8,803,268	214
rs2184380	rs11255871	0.948936	0.856732	1.01E-22	C10	8,803,763	215
rs2184380	rs11255882	0.884394	0.542127	3.04E-14	C10	8,810,853	216
rs2184380	rs11255884	0.933202	0.665922	4.29E-14	C10	8,811,968	217
rs2184380	rs12049705	0.885669	0.565013	5.63E-15	C10	8,679,446	218
rs2184380	rs12218610	0.856757	0.38242	1.22E-09	C10	8,706,615	219
rs2184380	rs12250379	0.926804	0.305824	1.61E-09	C10	8,802,201	26
rs2184380	rs12259226	0.946719	0.85359	2.31E-21	C10	8,806,948	220
rs2184380	rs1325874	1	1	1.53E-28	C10	8,740,707	221
rs2184380	rs1334549	1	1	1.53E-28	C10	8,744,989	29
rs2184380	rs1334550	1	1	1.53E-28	C10	8,744,845	222
rs2184380	rs1334559	0.948936	0.856732	1.01E-22	C10	8,796,810	223

【 0 3 3 5】

【表 1 3】

アンカー	代用	D'	R ²	P 値	染色体	ビルド36での位置	配列番号
rs2184380	rs1360749	1	1	1.53E-28	C10	8,759,215	224
rs2184380	rs1413678	0.924198	0.284149	5.54E-09	C10	8,794,683	225
rs2184380	rs1413683	0.944299	0.850224	5.27E-20	C10	8,804,953	226
rs2184380	rs1537601	0.949468	0.901005	7.18E-24	C10	8,785,004	227
rs2184380	rs1537602	0.949474	0.9015	5.73E-24	C10	8,784,898	228
rs2184380	rs1537603	1	0.201238	7.61E-09	C10	8,774,301	31
rs2184380	rs17407711	1	1	1.53E-28	C10	8,769,001	229
rs2184380	rs17407781	1	1	1.93E-28	C10	8,773,088	230
rs2184380	rs17407830	0.943872	0.805805	5.02E-20	C10	8,775,487	231
rs2184380	rs17408204	0.948913	0.854883	1.59E-22	C10	8,790,006	232
rs2184380	rs17408337	0.948936	0.856732	1.01E-22	C10	8,791,931	233
rs2184380	rs17408580	0.946832	0.854684	4.55E-22	C10	8,794,975	234
rs2184380	rs17484150	0.874978	0.400207	3.72E-11	C10	8,707,041	235
rs2184380	rs17485426	1	1	1.53E-28	C10	8,743,013	236
rs2184380	rs17485998	1	1	1.53E-28	C10	8,759,878	237
rs2184380	rs17486082	1	1	1.53E-28	C10	8,767,474	238
rs2184380	rs17486795	0.948936	0.856732	1.01E-22	C10	8,792,167	239
rs2184380	rs17486816	0.948936	0.856732	1.01E-22	C10	8,792,191	240
rs2184380	rs1970170	0.949389	0.899443	2.68E-23	C10	8,785,426	241
rs2184380	rs1999638	1	0.95	4.16E-26	C10	8,783,743	242
rs2184380	rs2031561	0.874358	0.384902	4.27E-11	C10	8,683,505	243
rs2184380	rs2182292	0.872839	0.372331	7.65E-11	C10	8,683,484	244
rs2184380	rs2388821	1	0.318315	4.75E-12	C10	8,736,687	245
rs2184380	rs2388825	1	1	1.53E-28	C10	8,756,643	246
rs2184380	rs2388826	0.946678	0.850726	4.53E-21	C10	8,805,023	247
rs2184380	rs2892613	1	1	1.53E-28	C10	8,756,702	248
rs2184380	rs4112287	0.948936	0.856732	1.01E-22	C10	8,791,551	249
rs2184380	rs4112288	0.948936	0.856732	1.01E-22	C10	8,791,525	250
rs2184380	rs4345867	1	0.283927	4.34E-11	C10	8,736,698	251
rs2184380	rs4454616	0.911606	0.211521	4.36E-07	C10	8,705,153	33
rs2184380	rs4747806	0.870907	0.357423	1.65E-10	C10	8,678,241	35
rs2184380	rs4749805	1	1	7.43E-28	C10	8,751,692	252
rs2184380	rs4749807	1	1	1.53E-28	C10	8,751,763	253
rs2184380	rs4749812	1	0.201238	7.61E-09	C10	8,777,570	254
rs2184380	rs6602328	0.8255	0.21783	3.95E-06	C10	8,700,007	255
rs2184380	rs6602329	0.843632	0.234214	2.42E-07	C10	8,700,101	36
rs2184380	rs7069110	1	0.318315	4.75E-12	C10	8,735,184	256
rs2184380	rs7080765	0.922785	0.313611	2.97E-09	C10	8,797,824	257
rs2184380	rs7083359	0.937236	0.429299	1.11E-12	C10	8,722,685	39
rs2184380	rs7477023	1	1	1.53E-28	C10	8,753,932	258
rs2184380	rs7904921	0.948936	0.856732	1.01E-22	C10	8,792,331	259
rs2184380	rs7912413	1	1	2.03E-24	C10	8,762,126	260
rs2184380	rs7912704	1	0.201238	7.61E-09	C10	8,771,181	261
rs2184380	rs7912831	1	0.201238	7.61E-09	C10	8,771,261	262
rs2184380	rs827389	1	0.417761	1.41E-13	C10	8,740,261	263
rs2184380	rs9665623	0.948936	0.856732	1.01E-22	C10	8,797,122	264
rs2224696	rs2224696	1	1	-	C10	9,168,781	3
rs2224696	rs10905509	1	0.461682	4.95E-12	C10	9,160,368	265
rs2224696	rs11256045	0.807004	0.329629	6.26E-08	C10	9,157,239	266
rs2224696	rs12761213	0.88785	0.589831	1.21E-10	C10	9,106,345	267
rs2224696	rs12761461	0.906382	0.751312	1.87E-13	C10	9,158,901	268
rs2224696	rs12766048	1	0.66736	4.52E-12	C10	9,133,679	52
rs2224696	rs12772042	1	0.919893	4.43E-18	C10	9,164,748	269
rs2224696	rs12776383	1	0.727891	1.42E-12	C10	9,227,620	270
rs2224696	rs12778120	0.621966	0.221569	0.000014	C10	9,253,860	271
rs2224696	rs12780218	0.906382	0.751312	1.87E-13	C10	9,159,062	272
rs2224696	rs12781427	1	0.919192	4.98E-18	C10	9,167,711	53
rs2224696	rs1475189	1	0.830628	2.85E-15	C10	9,147,691	273
rs2224696	rs1573109	0.772073	0.519575	1.43E-08	C10	9,141,049	274
rs2224696	rs1573110	0.794588	0.591124	4.06E-10	C10	9,135,501	275
rs2224696	rs17145088	1	0.830628	2.85E-15	C10	9,116,440	276

10

20

30

40

【 0 3 3 6 】

【表 1 4】

アンカー	代用	D'	R ²	P 値	染色体	ビルド36での位置	配列番号
rs2224696	rs17145095	1	0.830628	2.85E-15	C10	9,116,991	277
rs2224696	rs17145118	1	0.66736	5.30E-12	C10	9,121,203	278
rs2224696	rs17145120	1	0.830628	3.46E-15	C10	9,122,007	279
rs2224696	rs17145151	1	0.830628	2.85E-15	C10	9,130,711	280
rs2224696	rs17145164	1	0.830628	2.85E-15	C10	9,146,874	281
rs2224696	rs17145169	1	0.66736	4.90E-12	C10	9,147,141	282
rs2224696	rs17145188	0.621966	0.221569	0.000014	C10	9,148,524	283
rs2224696	rs17145193	0.883003	0.562854	1.15E-09	C10	9,155,959	284
rs2224696	rs17145221	1	0.89916	5.67E-15	C10	9,168,824	285
rs2224696	rs17363338	0.906382	0.751312	1.87E-13	C10	9,161,882	286
rs2224696	rs1775559	0.820762	0.495886	1.94E-10	C10	9,115,547	287
rs2224696	rs1857230	1	0.270147	1.73E-08	C10	9,127,246	288
rs2224696	rs1891532	1	0.209856	7.58E-08	C10	9,120,032	289
rs2224696	rs1935813	1	0.830628	2.85E-15	C10	9,118,559	290
rs2224696	rs2013364	1	1	1.16E-19	C10	9,168,159	291
rs2224696	rs2025289	0.906382	0.751312	1.87E-13	C10	9,161,060	292
rs2224696	rs2057442	1	0.242991	1.64E-08	C10	9,129,940	293
rs2224696	rs2093625	1	0.310553	1.63E-08	C10	9,126,798	294
rs2224696	rs2093626	1	0.252336	1.09E-08	C10	9,126,888	295
rs2224696	rs2146598	1	0.66736	4.52E-12	C10	9,118,227	296
rs2224696	rs2185817	0.796105	0.599318	1.59E-08	C10	9,140,970	297
rs2224696	rs2397336	1	0.267074	2.33E-08	C10	9,147,920	298
rs2224696	rs2760204	1	0.242991	1.64E-08	C10	9,139,981	299
rs2224696	rs2797266	1	0.242991	1.64E-08	C10	9,129,166	59
rs2224696	rs391733	0.780788	0.203685	0.00001	C10	9,191,976	300
rs2224696	rs4550140	1	0.25	1.23E-08	C10	9,145,853	301
rs2224696	rs7081544	1	0.830628	2.85E-15	C10	9,144,695	302
rs2224696	rs852273	1	0.276161	4.18E-09	C10	9,133,645	303
rs2224696	rs860418	1	0.245283	5.19E-08	C10	9,131,906	304
rs2224696	rs861172	1	0.240684	1.85E-08	C10	9,133,432	305
rs2224696	rs962993	0.885222	0.209643	1.52E-06	C10	9,093,138	67
rs2224696	rs965307	1	0.830628	2.85E-15	C10	9,119,005	306
rs2242503	rs2242503	1	1	-	C11	8,075,048	4
rs2242503	rs10431029	0.723593	0.468793	4.39E-13	C11	8,133,019	307
rs2242503	rs1055233	0.76309	0.517167	1.34E-14	C11	8,084,071	308
rs2242503	rs10734629	0.720552	0.46649	6.01E-13	C11	8,143,253	309
rs2242503	rs10743052	0.717443	0.464136	8.24E-13	C11	8,131,542	310
rs2242503	rs10743053	0.553357	0.295358	9.02E-09	C11	8,158,239	311
rs2242503	rs10743054	0.553357	0.295358	9.02E-09	C11	8,158,291	312
rs2242503	rs10743055	0.553357	0.295358	9.02E-09	C11	8,158,376	313
rs2242503	rs10769872	0.732431	0.432469	1.14E-11	C11	8,110,090	314
rs2242503	rs10769873	0.734884	0.322358	2.52E-09	C11	8,114,186	315
rs2242503	rs10769878	0.717939	0.348534	1.89E-08	C11	8,143,862	316
rs2242503	rs10769882	0.5497	0.229219	5.61E-06	C11	8,154,849	317
rs2242503	rs10839976	0.724865	0.308568	1.44E-08	C11	8,116,468	318
rs2242503	rs10839984	0.734884	0.322358	2.52E-09	C11	8,137,470	68
rs2242503	rs11041740	0.941774	0.461102	4.78E-14	C11	8,074,094	319
rs2242503	rs11041742	0.734884	0.322358	2.52E-09	C11	8,083,499	320
rs2242503	rs11041788	0.539249	0.212989	0.000025	C11	8,164,924	321
rs2242503	rs11041791	0.563934	0.235484	7.51E-07	C11	8,168,818	69
rs2242503	rs11041794	0.644371	0.240003	2.20E-06	C11	8,172,242	322
rs2242503	rs1108277	1	0.705954	1.08E-24	C11	8,077,794	323
rs2242503	rs12146654	0.64592	0.38833	1.53E-08	C11	8,152,036	324
rs2242503	rs12808387	0.553357	0.295358	9.02E-09	C11	8,159,414	325
rs2242503	rs1528125	0.553357	0.295358	9.02E-09	C11	8,156,792	326
rs2242503	rs1569128	0.723593	0.468793	4.39E-13	C11	8,089,461	327
rs2242503	rs1970880	0.544244	0.288701	1.80E-08	C11	8,160,092	328
rs2242503	rs1997262	0.522176	0.272668	3.97E-08	C11	8,158,988	329
rs2242503	rs2049684	0.723593	0.468793	4.39E-13	C11	8,127,695	72
rs2242503	rs2141321	0.789386	0.513662	2.21E-14	C11	8,086,752	330
rs2242503	rs2242501	0.737441	0.335833	2.13E-09	C11	8,074,679	331

10

20

30

40

【 0 3 3 7 】

【表 15】

アンカー	代用	D'	R ²	P 値	染色体	ビルド36での位置	配列番号
rs2242503	rs2272383	0.63086	0.397985	3.43E-11	C11	8,080,075	332
rs2242503	rs3750955	0.784744	0.385656	1.72E-10	C11	8,082,818	333
rs2242503	rs3752898	0.900696	0.245488	5.00E-07	C11	8,073,496	334
rs2242503	rs3849986	0.522176	0.272668	3.97E-08	C11	8,159,303	335
rs2242503	rs3849990	0.665217	0.252525	3.28E-07	C11	8,175,019	336
rs2242503	rs3911309	0.554372	0.202829	2.81E-06	C11	8,173,902	337
rs2242503	rs3911310	0.554372	0.202829	2.81E-06	C11	8,174,324	338
rs2242503	rs4340037	0.782801	0.349689	2.40E-10	C11	8,084,734	339
rs2242503	rs4343012	0.734884	0.322358	2.52E-09	C11	8,089,751	340
rs2242503	rs4385931	0.770418	0.551609	6.82E-16	C11	8,066,224	341
rs2242503	rs4575312	0.669287	0.255625	1.73E-07	C11	8,065,299	342
rs2242503	rs4578424	0.642849	0.214843	2.18E-06	C11	8,064,861	75
rs2242503	rs4636658	0.590173	0.221903	1.11E-06	C11	8,178,862	343
rs2242503	rs4758040	0.834104	0.50807	1.52E-15	C11	8,057,749	76
rs2242503	rs4758042	0.717439	0.45831	2.74E-12	C11	8,085,872	344
rs2242503	rs4758287	0.63975	0.255278	1.71E-07	C11	8,056,489	345
rs2242503	rs4758309	0.544875	0.227845	2.57E-06	C11	8,165,306	346
rs2242503	rs4758310	0.553934	0.235484	7.51E-07	C11	8,165,969	347
rs2242503	rs7103334	0.546807	0.278288	2.54E-08	C11	8,157,501	348
rs2242503	rs7112519	0.535211	0.217411	3.18E-06	C11	8,169,345	349
rs2242503	rs7115706	0.544644	0.223928	1.40E-06	C11	8,169,421	350
rs2242503	rs7122690	0.522176	0.272668	3.97E-08	C11	8,159,397	351
rs2242503	rs7127738	0.527011	0.276798	1.88E-07	C11	8,154,963	352
rs2242503	rs7358396	0.71764	0.42377	2.13E-10	C11	8,141,858	353
rs2242503	rs7479156	0.551079	0.289269	1.57E-08	C11	8,157,993	354
rs2242503	rs7479738	0.553357	0.295358	9.02E-09	C11	8,157,931	355
rs2242503	rs7480804	1	0.236507	1.40E-08	C11	8,066,385	356
rs2242503	rs7481667	0.553357	0.295358	9.02E-09	C11	8,157,701	357
rs2242503	rs7481683	0.553357	0.295358	9.02E-09	C11	8,157,762	358
rs2242503	rs7482611	0.693202	0.457371	2.46E-10	C11	8,106,602	359
rs2242503	rs7927368	0.643226	0.370442	1.47E-10	C11	8,152,196	360
rs2242503	rs7940668	0.643226	0.370442	1.47E-10	C11	8,152,710	361
rs12291026	rs12291026	1	1	-	C11	74,932,878	5
rs12291026	rs1004856	0.795363	0.22224	3.36E-06	C11	74,858,615	362
rs12291026	rs10899091	1	1	1.43E-20	C11	74,933,440	363
rs12291026	rs11236449	1	0.919893	4.43E-18	C11	74,918,745	364
rs12291026	rs11236452	1	0.919893	4.98E-18	C11	74,927,534	365
rs12291026	rs11236454	1	0.906687	6.96E-16	C11	74,932,272	366
rs12291026	rs12362081	1	0.84127	3.43E-16	C11	74,899,292	367
rs12291026	rs1540210	0.800306	0.242864	1.38E-06	C11	74,887,046	78
rs12291026	rs1540211	0.750347	0.450541	1.04E-09	C11	74,887,332	368
rs12291026	rs1557471	0.79757	0.231043	2.21E-06	C11	74,857,322	369
rs12291026	rs1631470	0.799311	0.23844	1.70E-06	C11	74,864,790	370
rs12291026	rs1783551	0.837984	0.701135	2.54E-13	C11	74,908,860	371
rs12291026	rs1783556	0.876315	0.242246	0.000013	C11	74,897,710	372
rs12291026	rs1783559	0.820042	0.540279	1.73E-10	C11	74,894,676	373
rs12291026	rs1790144	0.750347	0.450541	1.04E-09	C11	74,889,549	79
rs12291026	rs1790152	0.79757	0.231043	2.21E-06	C11	74,862,499	374
rs12291026	rs1790307	0.830991	0.514897	7.92E-11	C11	74,898,758	375
rs12291026	rs1793396	0.895234	0.292441	6.02E-08	C11	74,890,048	376
rs12291026	rs1793397	0.750347	0.450541	1.04E-09	C11	74,890,155	377
rs12291026	rs1793398	0.906458	0.369032	1.79E-09	C11	74,893,069	378
rs12291026	rs1793399	0.831094	0.516953	7.12E-11	C11	74,898,153	80
rs12291026	rs1793414	0.792872	0.226475	4.83E-06	C11	74,860,335	379
rs12291026	rs1938800	1	0.748428	2.23E-16	C11	74,933,770	81
rs12291026	rs2853066	0.804722	0.264424	6.22E-07	C11	74,866,805	380
rs12291026	rs499613	0.792872	0.226475	4.83E-06	C11	74,872,133	381
rs12291026	rs504793	0.800306	0.242864	1.38E-06	C11	74,878,212	382
rs12291026	rs514477	0.800306	0.242864	1.38E-06	C11	74,875,650	383
rs12291026	rs549034	0.799814	0.24066	1.53E-06	C11	74,880,782	384
rs12291026	rs550881	0.800306	0.242864	1.38E-06	C11	74,880,612	385

10

20

30

40

【 0 3 3 8 】

【表 16】

アンカー	代用	D'	R ²	P 値	染色体	ビルド36での位置	配列番号
rs12291026	rs581007	0.800306	0.242864	1.38E-06	C11	74,879,594	386
rs12291026	rs589724	0.805181	0.254846	6.34E-07	C11	74,876,829	387
rs12291026	rs600387	1	0.293756	1.11E-09	C11	74,935,675	388
rs12291026	rs606460	0.902797	0.333104	1.27E-08	C11	74,926,105	389
rs12291026	rs617617	1	0.425837	7.51E-12	C11	74,936,378	390
rs12291026	rs618202	0.838016	0.70227	2.28E-13	C11	74,916,133	86
rs12291026	rs628972	0.800306	0.242864	1.38E-06	C11	74,885,999	391
rs12291026	rs640649	0.793071	0.480582	5.31E-09	C11	74,949,499	88
rs12291026	rs662279	1	0.334073	2.08E-10	C11	74,939,444	90
rs12291026	rs667410	0.891296	0.289518	1.94E-07	C11	74,872,345	392
rs12291026	rs667531	0.721513	0.369086	1.84E-07	C11	74,950,364	393
rs12291026	rs670100	0.89343	0.292234	1.02E-07	C11	74,926,922	394
rs12291026	rs670491	0.902797	0.333104	1.27E-08	C11	74,926,833	395
rs12291026	rs682292	0.800306	0.242864	1.38E-06	C11	74,877,949	396
rs12291026	rs7128888	1	1	1.43E-20	C11	74,931,631	397
rs12291026	rs7129014	1	1	1.43E-20	C11	74,931,725	398
rs12291026	rs7129150	1	0.921773	7.97E-18	C11	74,931,825	399
rs12291026	rs947844	1	0.84127	3.43E-16	C11	74,906,532	400
rs999737	rs999737	1	1	-	C14	68,104,435	6
rs999737	rs10134446	1	0.22	1.39E-09	C14	68,067,388	91
rs999737	rs10138140	0.674626	0.269357	4.17E-08	C14	68,011,742	401
rs999737	rs10146772	0.659865	0.223224	3.27E-07	C14	67,948,122	402
rs999737	rs10467820	0.674626	0.269357	4.17E-08	C14	67,988,769	403
rs999737	rs10483812	1	0.914924	1.84E-27	C14	68,087,596	404
rs999737	rs10483813	1	1	1.04E-30	C14	68,101,037	405
rs999737	rs11158749	1	0.224097	1.07E-09	C14	68,067,049	406
rs999737	rs11158751	1	0.340659	4.73E-13	C14	68,101,178	95
rs999737	rs11621276	1	0.87513	3.15E-26	C14	68,066,219	96
rs999737	rs11624097	0.953506	0.831825	5.44E-24	C14	68,055,789	407
rs999737	rs11624164	0.953471	0.82933	9.30E-24	C14	68,055,910	408
rs999737	rs11624333	0.953506	0.831825	5.44E-24	C14	68,049,588	409
rs999737	rs11628293	1	0.954597	2.00E-28	C14	68,107,073	410
rs999737	rs11846360	0.528484	0.243633	1.31E-06	C14	67,997,698	411
rs999737	rs11847185	0.499138	0.226969	2.79E-06	C14	67,997,747	412
rs999737	rs11849916	0.499138	0.226969	2.79E-06	C14	67,957,791	413
rs999737	rs12878761	0.527934	0.255003	2.88E-07	C14	68,138,118	414
rs999737	rs12879200	0.739817	0.228092	4.95E-07	C14	68,137,863	415
rs999737	rs12886864	0.662607	0.259908	1.17E-07	C14	68,017,871	416
rs999737	rs12889251	1	0.876374	2.38E-26	C14	68,063,893	417
rs999737	rs12894230	0.544425	0.273137	2.77E-07	C14	68,137,669	418
rs999737	rs1468279	0.659722	0.26746	2.08E-07	C14	67,994,846	419
rs999737	rs1468280	1	0.310873	3.99E-12	C14	68,101,613	420
rs999737	rs1547012	0.575976	0.303524	1.93E-08	C14	68,133,183	421
rs999737	rs17105675	0.499138	0.226969	2.79E-06	C14	67,998,259	422
rs999737	rs17755657	0.511149	0.219511	2.63E-06	C14	67,950,390	423
rs999737	rs17755734	0.499138	0.226969	2.79E-06	C14	67,979,708	424
rs999737	rs17755752	0.499138	0.226969	2.79E-06	C14	67,979,904	425
rs999737	rs17755925	1	0.876374	2.38E-26	C14	68,081,699	426
rs999737	rs17756000	1	0.914924	1.84E-27	C14	68,089,291	427
rs999737	rs17828691	0.499138	0.226969	2.79E-06	C14	67,956,346	428
rs999737	rs17828721	0.499138	0.226969	2.79E-06	C14	67,979,791	429
rs999737	rs17828763	0.499138	0.226969	2.79E-06	C14	67,984,862	430
rs999737	rs17828907	1	0.914924	1.84E-27	C14	68,080,958	431
rs999737	rs17828955	1	0.914924	1.84E-27	C14	68,086,398	432
rs999737	rs1956534	0.587868	0.237655	4.20E-07	C14	67,963,825	433
rs999737	rs2074563	0.674626	0.269357	4.17E-08	C14	68,004,865	434
rs999737	rs2074565	0.674626	0.269357	4.17E-08	C14	68,013,900	435
rs999737	rs2097800	0.66475	0.25742	2.25E-07	C14	68,006,440	436
rs999737	rs2107340	0.674626	0.269357	4.17E-08	C14	68,000,418	437
rs999737	rs2145157	0.739817	0.228092	4.95E-07	C14	68,137,391	438
rs999737	rs2158357	1	0.956044	8.64E-29	C14	68,098,956	439

10

20

30

40

【 0 3 3 9 】

【表 17】

アンカー	代用	D'	R ²	P 値	染色体	ビルド36での位置	配列番号
rs999737	rs2189517	1	0.219935	1.39E-09	C14	68,072,741	440
rs999737	rs2253317	0.923817	0.281194	1.03E-08	C14	68,122,156	441
rs999737	rs2257111	0.532356	0.267955	1.67E-07	C14	68,129,446	442
rs999737	rs2257116	0.533815	0.272433	1.21E-07	C14	68,129,544	443
rs999737	rs2257127	0.533815	0.272433	1.21E-07	C14	68,129,821	444
rs999737	rs2331701	0.669937	0.256122	8.50E-08	C14	67,952,384	445
rs999737	rs2331705	0.510976	0.244712	1.44E-06	C14	67,991,800	446
rs999737	rs2331775	0.739817	0.228092	4.95E-07	C14	68,136,385	447
rs999737	rs2525503	1	0.318681	1.88E-12	C14	68,098,356	105
rs999737	rs2525530	1	0.205178	4.00E-09	C14	68,084,205	448
rs999737	rs2842327	1	0.382208	4.73E-13	C14	68,098,285	449
rs999737	rs3784121	0.499138	0.226969	2.79E-06	C14	67,953,441	450
rs999737	rs4531674	0.527934	0.255003	2.88E-07	C14	68,136,496	451
rs999737	rs4899246	0.533815	0.272433	1.21E-07	C14	68,129,955	108
rs999737	rs4902604	0.561463	0.264852	2.16E-07	C14	68,124,238	452
rs999737	rs4902606	0.739817	0.228092	4.95E-07	C14	68,135,119	453
rs999737	rs4902608	0.739817	0.228092	4.95E-07	C14	68,138,576	454
rs999737	rs5004090	0.674626	0.269357	4.17E-08	C14	68,013,998	455
rs999737	rs6573837	0.77856	0.353726	5.97E-10	C14	67,999,862	456
rs999737	rs7140266	1	0.219935	1.39E-09	C14	68,074,558	457
rs999737	rs7146456	1	0.260433	1.04E-10	C14	68,108,734	458
rs999737	rs7153476	1	0.308233	3.65E-12	C14	68,102,983	459
rs999737	rs739874	0.952339	0.761982	7.28E-22	C14	68,044,163	460
rs999737	rs746663	0.61957	0.227187	5.55E-07	C14	67,954,305	461
rs999737	rs8007194	1	0.281975	2.77E-11	C14	68,105,178	462
rs999737	rs8010439	0.587868	0.237655	4.20E-07	C14	67,954,354	463
rs999737	rs8012610	0.659865	0.223224	3.27E-07	C14	67,972,552	464
rs999737	rs9323512	0.674626	0.269357	4.17E-08	C14	68,019,310	465
rs999737	rs9323513	1	0.220532	1.43E-09	C14	68,066,244	466
rs999737	rs9323514	1	0.315043	2.49E-12	C14	68,066,317	467
rs9956546	rs9956546	1	1	-	C18	32,136,446	7
rs9956546	rs16960059	1	0.66736	4.52E-12	C18	32,120,094	468
rs9956546	rs492392	1	0.399199	3.93E-11	C18	32,142,779	120
rs9956546	rs7239113	1	0.66736	4.52E-12	C18	32,133,395	469
rs9956546	rs7240720	0.524852	0.251926	0.000036	C18	32,142,893	124
rs9956546	rs9304157	1	0.211045	0.000182	C18	32,143,156	470
rs11912922	rs11912922	1	1	-	C22	38,733,117	8
rs11912922	rs11089967	1	1	9.56E-34	C22	38,739,085	471
rs11912922	rs11704971	1	1	1.57E-30	C22	38,734,239	472
rs11912922	rs11705454	1	1	6.56E-31	C22	38,731,841	473
rs11912922	rs17406386	1	1	1.37E-33	C22	38,735,844	474
rs11912922	rs17406434	1	1	3.28E-33	C22	38,737,520	475
rs11912922	rs2071771	1	0.214286	1.52E-09	C22	38,745,086	476
rs11912922	rs2958650	1	1	3.28E-33	C22	38,735,877	477
rs11912922	rs2958651	1	0.958241	2.48E-29	C22	38,735,709	478
rs11912922	rs2958659	1	1	9.56E-34	C22	38,730,858	479
rs11912922	rs7284488	1	0.239332	2.31E-10	C22	38,742,497	480
rs11912922	rs7285507	0.747862	0.497836	2.02E-12	C22	38,718,254	481
rs11912922	rs7291782	1	0.924812	2.56E-30	C22	38,745,563	482
rs11912922	rs739145	1	0.24812	1.20E-10	C22	38,728,791	483
rs11912922	rs9611246	1	0.214286	1.52E-09	C22	38,745,655	484
rs11912922	rs9611265	0.957937	0.846507	5.98E-24	C22	38,828,439	485
rs6001954	rs6001954	1	1	-	C22	39,251,626	9
rs6001954	rs10483203	0.922166	0.736052	7.16E-15	C22	39,159,899	486
rs6001954	rs10483204	1	0.865546	9.75E-19	C22	39,200,740	487
rs6001954	rs10483205	1	0.800222	4.35E-17	C22	39,213,545	488
rs6001954	rs10483206	1	0.865546	9.75E-19	C22	39,231,291	489
rs6001954	rs1106673	0.922166	0.736052	7.16E-15	C22	39,155,760	490
rs6001954	rs11913132	0.696801	0.304518	1.56E-07	C22	39,108,724	491
rs6001954	rs12158399	1	0.568501	2.92E-15	C22	39,378,887	492
rs6001954	rs12158872	0.922166	0.736052	7.16E-15	C22	39,162,447	493

10

20

30

40

【 0 3 4 0 】

【表 18】

アンカー	代用	D'	R ²	P 値	染色体	ビルド36での位置	配列番号
rs6001954	rs12159200	1	0.542302	7.76E-15	C22	39,372,037	138
rs6001954	rs12159970	1	0.800222	4.35E-17	C22	39,149,100	494
rs6001954	rs12484697	0.624393	0.257388	1.79E-06	C22	39,066,418	139
rs6001954	rs12627881	0.925127	0.534379	1.11E-12	C22	39,374,539	495
rs6001954	rs133036	1	0.231068	6.07E-09	C22	39,342,384	141
rs6001954	rs133038	1	0.223301	9.26E-09	C22	39,345,215	496
rs6001954	rs16985899	1	0.932127	1.30E-20	C22	39,293,348	497
rs6001954	rs17001846	0.624393	0.257388	1.79E-06	C22	39,073,999	498
rs6001954	rs17001868	0.635691	0.328651	4.59E-07	C22	39,108,177	499
rs6001954	rs17001943	0.922166	0.736052	7.16E-15	C22	39,192,559	500
rs6001954	rs17001974	1	0.865546	9.75E-19	C22	39,207,330	501
rs6001954	rs17001977	1	0.779103	6.11E-16	C22	39,210,159	502
rs6001954	rs17001993	1	0.800222	4.35E-17	C22	39,230,023	503
rs6001954	rs17001994	1	0.788111	8.57E-16	C22	39,230,764	504
rs6001954	rs17001997	1	0.788111	4.07E-16	C22	39,235,018	505
rs6001954	rs17002019	1	0.932127	1.30E-20	C22	39,295,282	506
rs6001954	rs17002020	1	0.930025	1.15E-19	C22	39,295,389	507
rs6001954	rs17002026	1	0.930456	2.22E-19	C22	39,312,527	508
rs6001954	rs17002027	1	0.925154	2.26E-18	C22	39,314,517	509
rs6001954	rs17002034	1	0.865546	9.75E-19	C22	39,326,313	151
rs6001954	rs17002036	1	0.932127	1.30E-20	C22	39,327,057	510
rs6001954	rs17002038	1	0.925548	4.67E-19	C22	39,330,910	511
rs6001954	rs17002069	0.923881	0.509328	4.13E-12	C22	39,380,932	512
rs6001954	rs2075764	0.624393	0.257388	1.79E-06	C22	39,088,527	513
rs6001954	rs2187832	1	0.223301	9.26E-09	C22	39,328,395	514
rs6001954	rs2235318	0.628984	0.275382	9.11E-07	C22	39,130,490	515
rs6001954	rs2280790	0.624393	0.257388	1.79E-06	C22	39,075,127	516
rs6001954	rs2294348	0.694473	0.293289	2.75E-07	C22	39,113,726	517
rs6001954	rs2294350	0.624393	0.257388	1.79E-06	C22	39,114,224	518
rs6001954	rs2294352	0.922166	0.736052	7.16E-15	C22	39,157,265	519
rs6001954	rs2413624	1	0.260204	3.65E-09	C22	39,260,158	520
rs6001954	rs3788577	0.624393	0.257388	1.79E-06	C22	39,072,777	521
rs6001954	rs3788578	0.624393	0.257388	1.79E-06	C22	39,072,867	522
rs6001954	rs3788579	0.624453	0.267064	1.98E-06	C22	39,072,935	523
rs6001954	rs3827381	1	0.865546	9.75E-19	C22	39,211,348	524
rs6001954	rs3827382	1	0.800222	4.35E-17	C22	39,211,349	153
rs6001954	rs4140512	1	0.568501	2.92E-15	C22	39,362,809	525
rs6001954	rs470113	0.624393	0.257388	1.79E-06	C22	39,059,560	155
rs6001954	rs5750957	1	0.265513	1.01E-09	C22	39,265,239	526
rs6001954	rs5750960	1	0.231068	6.07E-09	C22	39,289,866	157
rs6001954	rs5750966	1	0.213336	1.64E-08	C22	39,299,739	527
rs6001954	rs5757976	1	0.223301	9.26E-09	C22	39,245,664	528
rs6001954	rs5757998	1	0.265513	1.01E-09	C22	39,269,856	529
rs6001954	rs5758001	0.90345	0.224527	3.42E-07	C22	39,276,253	159
rs6001954	rs5995849	0.696801	0.304518	1.56E-07	C22	39,107,018	530
rs6001954	rs5995856	0.927089	0.801157	1.72E-16	C22	39,179,008	531
rs6001954	rs5995870	1	1	4.37E-23	C22	39,251,762	532
rs6001954	rs5995871	1	1	5.10E-23	C22	39,252,278	533
rs6001954	rs5995886	1	0.563253	6.45E-15	C22	39,363,747	534
rs6001954	rs6001900	0.693507	0.310492	4.91E-07	C22	39,120,948	535
rs6001954	rs6001910	1	0.800222	4.35E-17	C22	39,153,955	536
rs6001954	rs6001911	0.922166	0.736052	7.16E-15	C22	39,155,139	537
rs6001954	rs6001912	0.922166	0.736052	7.16E-15	C22	39,158,307	538
rs6001954	rs6001913	1	0.800222	4.35E-17	C22	39,166,699	539
rs6001954	rs6001930	1	0.865546	9.75E-19	C22	39,206,180	540
rs6001954	rs6001931	1	0.857313	9.26E-18	C22	39,207,460	541
rs6001954	rs6001932	1	0.865546	9.75E-19	C22	39,207,581	161
rs6001954	rs6001935	1	0.865546	9.75E-19	C22	39,216,939	542
rs6001954	rs6001950	1	0.865546	9.75E-19	C22	39,236,581	543
rs6001954	rs6001974	1	1	4.37E-23	C22	39,309,110	544
rs6001954	rs6001980	1	1	4.37E-23	C22	39,334,330	545

10

20

30

40

【表 19】

アンカー	代用	D'	R ²	P 値	染色体	ビルド36での位置	配列番号
rs6001954	rs6001990	1	0.568501	2.92E-15	C22	39,364,131	546
rs6001954	rs6002000	1	0.568501	2.92E-15	C22	39,381,055	547
rs6001954	rs718193	1	0.932127	1.30E-20	C22	39,354,010	548
rs6001954	rs7292804	0.923881	0.509328	4.13E-12	C22	39,377,502	549
rs6001954	rs7293100	0.92505	0.531874	1.29E-12	C22	39,377,697	550
rs6001954	rs742140	0.624393	0.257388	1.79E-06	C22	39,111,532	551
rs6001954	rs760700	0.624393	0.257388	1.79E-06	C22	39,130,017	552
rs6001954	rs760701	0.628984	0.275382	9.11E-07	C22	39,130,464	553
rs6001954	rs9306345	0.780144	0.423648	8.54E-10	C22	39,162,321	167
rs6001954	rs932379	1	0.932127	1.30E-20	C22	39,282,480	554
rs6001954	rs9607721	0.851982	0.479217	3.33E-11	C22	39,192,066	168
rs6001954	rs9611310	0.624393	0.257388	1.79E-06	C22	39,063,723	555
rs6001954	rs9611311	0.694782	0.31869	2.91E-07	C22	39,063,897	556
rs6001954	rs9611312	0.61091	0.260998	7.16E-06	C22	39,064,736	557
rs6001954	rs9611316	0.673329	0.320213	8.89E-07	C22	39,066,973	558
rs6001954	rs9611318	0.624393	0.257388	1.79E-06	C22	39,069,360	559
rs6001954	rs9611324	0.696801	0.304518	1.56E-07	C22	39,111,108	172
rs6001954	rs9611325	0.624393	0.257388	1.79E-06	C22	39,112,436	560
rs6001954	rs9611328	0.624393	0.257388	1.79E-06	C22	39,121,918	561
rs6001954	rs9611329	0.624393	0.257388	1.79E-06	C22	39,121,984	562

10

【 0 3 4 2 】

【表 2 0】

表5：染色体22上のrs6001954についての乳癌との関連が台湾人を含む個々のコホートについて示されている。合わされたヨーロッパ人の結果およびすべての合わされたコホートの結果も示されている。

試料セット	SNP	対立遺伝子	P値	OR	症例数	症例頻度	対照数	対照頻度	95%信頼区間	Phet
アイスランド	rs6001954	3	7.20E-04	1.18	2321	0.134	35787	0.117		
DKFZ-スウェーデン	rs6001954	3	5.59E-01	1.06	937	0.137	930	0.130		
オランダ	rs6001954	3	2.90E-01	1.11	743	0.116	2031	0.106		
メイヨー(MAYO)-米国	rs6001954	3	7.29E-02	1.17	1508	0.120	1264	0.104		
スペイン	rs6001954	3	2.99E-01	1.10	912	0.108	1864	0.099		
スウェーデン-カロリンスカ	rs6001954	3	1.26E-01	1.16	825	0.119	1724	0.105		
CGEMS	rs6001954	3	9.30E-02	1.22	1101	0.109	1105	0.091		
台湾	rs6001954	3	1.02E-01	1.13	1026	0.243	901	0.221		
全ヨーロッパ人	rs6001954	3	5.20E-06	1.15	8347	N/A	44705	N/A	(1.08, 1.22)	0.95
合わされたすべて	rs6001954	3	1.30E-06	1.15	9373	N/A	45606	N/A	(1.08, 1.21)	0.98

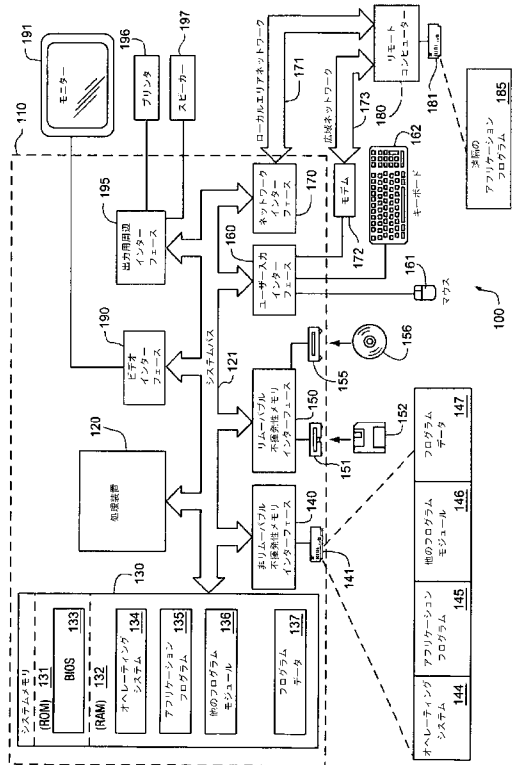
10

20

30

40

【図 1】



【配列表】

2011527565000001.app

【 国際調査報告 】

INTERNATIONAL SEARCH REPORT

International application No
PCT/IS2009/000008

A. CLASSIFICATION OF SUBJECT MATTER INV. C12Q1/68	
According to International Patent Classification (IPC) or to both national classification and IPC	
B. FIELDS SEARCHED	
Minimum documentation searched (classification system followed by classification symbols) C12Q	
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched	
Electronic data base consulted during the international search (name of data base and, where practical, search terms used) EPO-Internal, BIOSIS, EMBASE	
C. DOCUMENTS CONSIDERED TO BE RELEVANT	
Category*	Citation of document, with indication, where appropriate, of the relevant passages
	Relevant to claim No.
X	<p>DATABASE PROBE [Online] NCBI; Distributor: Perlegen Sciences Inc 1 March 2006 (2006-03-01), ANONYMOUS: "Sequence-specific oligonucleotide (SSO) probe for Homo sapiens variation rs999737" XP002548192 retrieved from NCBI Database accession no. Pr005049093 abstract</p> <p style="text-align: center;">----- -/--</p>
<input checked="" type="checkbox"/>	Further documents are listed in the continuation of Box C.
<input checked="" type="checkbox"/>	See patent family annex.
* Special categories of cited documents :	
"A" document defining the general state of the art which is not considered to be of particular relevance	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"E" earlier document but published on or after the international filing date	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
"O" document referring to an oral disclosure, use, exhibition or other means	"&" document member of the same patent family
"P" document published prior to the international filing date but later than the priority date claimed	
Date of the actual completion of the international search 1 October 2009	Date of mailing of the international search report 28/01/2010
Name and mailing address of the ISA/ European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040 Fax: (+31-70) 340-3016	Authorized officer Bradbrook, Derek

2

Form PCT/ISA/210 (second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

International application No.

PCT/IS2009/000008

Box No. 1 Nucleotide and/or amino acid sequence(s) (Continuation of item 1.b of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application and necessary to the claimed invention, the international search was carried out on the basis of:
- a. type of material
- a sequence listing
- table(s) related to the sequence listing
- b. format of material
- on paper
- in electronic form
- c. time of filing/furnishing
- contained in the international application as filed
- filed together with the international application in electronic form
- furnished subsequently to this Authority for the purpose of search
2. In addition, in the case that more than one version or copy of a sequence listing and/or table relating thereto has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that in the application as filed or does not go beyond the application as filed, as appropriate, were furnished.
3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No PCT/IS2009/000008

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>DATABASE PROBE [Online] NCBI; Distributor: Illumina Inc. 11 January 2007 (2007-01-11), ANONYMOUS: "Bead microelement (Bead) probe for Homo sapiens variation rs999737" XP002548193 retrieved from NCBI Database accession no. Pr007592512 abstract</p>	19-24
X	<p>STEEMERS FRANK J ET AL: "Whole genome genotyping technologies on the BeadArray platform." BIOTECHNOLOGY JOURNAL JAN 2007, vol. 2, no. 1, January 2007 (2007-01), pages 41-49, XP002548190 ISSN: 1860-7314 the whole document</p>	19-24
P, X	<p>THOMAS GILLES ET AL: "A multistage genome-wide association study in breast cancer identifies two new risk alleles at 1p11.2 and 14q24.1 (RAD51L1)." NATURE GENETICS MAY 2009, vol. 41, no. 5, May 2009 (2009-05), pages 579-584, XP002548191 ISSN: 1546-1718 the whole document</p>	1-60
A	<p>WO 2006/063285 A (SEQUENOM INC; ROTH RICHARD B [US]; NELSON MATTHEW R [US]; KAMMERER STE) 15 June 2006 (2006-06-15) examples 1-4</p>	1-60
A	<p>HOYAL CAROLYN R ET AL: "Genetic polymorphisms in DPF3 associated with risk of breast cancer and lymph node metastases" JOURNAL OF CARCINOGENESIS, BIOMED CENTRAL, LONDON, GB, vol. 4, no. 1, 19 August 2005 (2005-08-19), page 13, XP021008706 ISSN: 1477-3163 abstract</p>	1-60
A	<p>STACEY SIMON N ET AL: "Common variants on chromosome 5p12 confer susceptibility to estrogen receptor-positive breast cancer" NATURE GENETICS, NATURE PUBLISHING GROUP, NEW YORK, US, vol. 40, no. 6, 1 June 2008 (2008-06-01), pages 703-706, XP002491374 ISSN: 1061-4036 [retrieved on 2008-04-27] abstract</p>	1-60
2	----- -/--	

Form PCT/ISA/210 (continuation of second sheet) (April 2006)

INTERNATIONAL SEARCH REPORT

International application No PCT/IS2009/000008

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>GOLD BERT ET AL: "Genome-wide association study provides evidence for a breast cancer risk locus at 6q22-33" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, NATIONAL ACADEMY OF SCIENCE, WASHINGTON, DC, US, vol. 105, no. 11, 1 March 2008 (2008-03-01), pages 4340-4345, XP002527419 ISSN: 0027-8424 abstract</p> <p>-----</p>	1-60
A	<p>STACEY SIMON N ET AL: "Common variants on chromosomes 2q35 and 16q12 confer susceptibility to estrogen receptor-positive breast cancer" NATURE GENETICS, NATURE PUBLISHING GROUP, NEW YORK, US, vol. 39, no. 7, 1 July 2007 (2007-07-01), pages 865-869, XP002487595 ISSN: 1061-4036 [retrieved on 2007-05-27] abstract</p> <p>-----</p>	1-60

INTERNATIONAL SEARCH REPORT

International application No.
PCT/IS2009/000008**Box No. II Observations where certain claims were found unsearchable (Continuation of Item 2 of first sheet)**

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

1. Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. Claims Nos.:
because they relate to parts of the International application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

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

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

see additional sheet

1. As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of additional fees.
3. As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

see additional sheet(s)

Remark on Protest

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

International Application No. PCT/IS2009 /000008

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

Invention 1. Claims: 1-60 (in part)

A method for determining susceptibility of a human subject to breast cancer comprising determining the presence or absence of at least one allele of at least one polymorphic marker in a sample from the subject, wherein the at least one marker is rs999737 or markers in linkage disequilibrium therewith; related methods for determining risk of developing a second primary tumour in an individual diagnosed with breast cancer, for assessing probability of response to breast cancer therapy, for predicting prognosis of breast cancer or for monitoring progress of treatment for breast cancer; related kit; related use of an oligonucleotide probe; related computer medium and apparatus and method of assessing subject's risk for breast cancer.

Invention 2. Claims: 1-60 (in part)

As invention 1, wherein the at least one marker is rs2005154 or markers in linkage disequilibrium therewith

Inventions 3-9. Claims: 1-60 (in part)

As invention 1, wherein the at least one marker is respectively rs2184380, rs2224696, rs2242503, rs12291026, rs9956546, rs11912922 and rs6001954 or markers in linkage disequilibrium therewith.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/IS2009/000008

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2006063285	A	NONE	

フロントページの続き

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

(特許庁注：以下のものは登録商標)

1. レーザーディスク

Fターム(参考) 4B024 AA12 BA80 CA01 CA09 DA03 HA12
4B063 QA12 QA19 QQ42 QQ58 QR08 QR32 QR55 QR62 QS25 QS34

专利名称(译)	乳腺癌风险评估的遗传变异		
公开(公告)号	JP2011527565A	公开(公告)日	2011-11-04
申请号	JP2011517318	申请日	2009-07-03
[标]申请(专利权)人(译)	解码遗传学私营有限责任公司		
申请(专利权)人(译)	Dikode Jenetekusu leichiefu		
[标]发明人	ステイシーサイモン スレムパトリック		
发明人	ステイシー サイモン スレム パトリック		
IPC分类号	C12Q1/68 C12N15/09 G01N33/53		
CPC分类号	C12Q1/6886 C12Q2600/106 C12Q2600/136 C12Q2600/172		
FI分类号	C12Q1/68.ZNA.A C12N15/00.A G01N33/53.M		
F-TERM分类号	4B024/AA12 4B024/BA80 4B024/CA01 4B024/CA09 4B024/DA03 4B024/HA12 4B063/QA12 4B063/QA19 4B063/QQ42 4B063/QQ58 4B063/QR08 4B063/QR32 4B063/QR55 4B063/QR62 4B063/QS25 4B063/QS34		
代理人(译)	Seihayashi正幸 和义林		
优先权	8746 2008-07-07 IS		
其他公开文献	JP2011527565A5		
外部链接	Espacenet		

摘要(译)

本发明涉及确定为乳腺癌易感变体的遗传变体。描述了包括诊断乳腺癌易感性增加的疾病管理方法，预测治疗反应的方法以及使用这种突变体预测预后的方法。本发明还涉及可用于本发明方法的试剂盒。【选型图】无

IUB コード	意味
A	アデノシン
C	シチジン
G	グアニン
T	チミジン
R	GまたはA
Y	TまたはC
K	GまたはT
M	AまたはC
S	GまたはC
W	AまたはT
B	C GまたはT
D	A GまたはT
H	A CまたはT
V	A CまたはG
N	A C GまたはT (いずれかの塩基)