

(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2007-529714

(P2007-529714A)

(43) 公表日 平成19年10月25日(2007. 10. 25)

(51) Int. Cl.	F I	テーマコード (参考)
GO 1 N 33/53 (2006.01)	GO 1 N 33/53 M	2 GO 4 5
C 1 2 Q 1/68 (2006.01)	C 1 2 Q 1/68 A	4 B O 2 4
C 1 2 Q 1/02 (2006.01)	C 1 2 Q 1/02	4 B O 6 3
GO 1 N 33/15 (2006.01)	GO 1 N 33/15 Z	
GO 1 N 33/50 (2006.01)	GO 1 N 33/50 Z	
審査請求 未請求 予備審査請求 未請求 (全 42 頁) 最終頁に続く		

(21) 出願番号 特願2007-502088 (P2007-502088)
 (86) (22) 出願日 平成17年3月3日 (2005. 3. 3)
 (85) 翻訳文提出日 平成18年10月17日 (2006. 10. 17)
 (86) 国際出願番号 PCT/US2005/007375
 (87) 国際公開番号 W02005/086770
 (87) 国際公開日 平成17年9月22日 (2005. 9. 22)
 (31) 優先権主張番号 60/550, 662
 (32) 優先日 平成16年3月5日 (2004. 3. 5)
 (33) 優先権主張国 米国 (US)
 (31) 優先権主張番号 60/566, 302
 (32) 優先日 平成16年4月28日 (2004. 4. 28)
 (33) 優先権主張国 米国 (US)
 (31) 優先権主張番号 60/590, 534
 (32) 優先日 平成16年7月22日 (2004. 7. 22)
 (33) 優先権主張国 米国 (US)

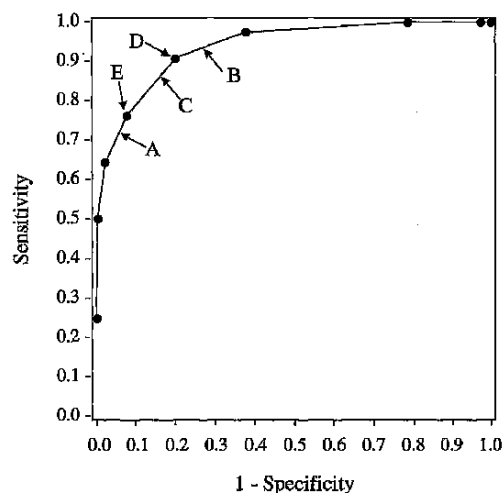
(71) 出願人 504261826
 パーレジェン サイエンス インク.
 アメリカ合衆国 カリフォルニア州 マウ
 ンテン ビュー スターリン コート 2
 O 2 1
 (74) 代理人 100102978
 弁理士 清水 初志
 (74) 代理人 100128048
 弁理士 新見 浩一
 (72) 発明者 コックス デビッド アール.
 アメリカ合衆国 カリフォルニア州 ベル
 モント ホールマーク ドライブ 274
 3

最終頁に続く

(54) 【発明の名称】 遺伝子分析のための方法

(57) 【要約】

個体が多因子性形質を発症または示す尤度を評価するいくつかの方法を記述する。方法は、個体に関するスコアを算出するために遺伝子型を用いて、複数の二対立遺伝子多型座位で個体に関する複数の遺伝子型を決定する段階、およびスコアを少なくとも一つの閾値と比較する段階を含む。



【特許請求の範囲】

【請求項 1】

以下の段階を含む、個体が多因子形質を発症または示す尤度を評価する方法：

a) 複数の二対立遺伝子多型座位で個体の遺伝子型を決定する段階であって、複数のそれぞれが関連対立遺伝子と非関連対立遺伝子とを有し、さらに遺伝子型が関連対立遺伝子に関してホモ接合、および非関連対立遺伝子に関してヘテロ接合、ホモ接合からなる群より選択される段階；

b) 段階a) において決定された遺伝子型に基づいて個体のスコアを算出する段階；および

c) スコアと少なくとも一つの閾値との比較を行う段階であって、該比較を用いて個体の多因子形質を発症または示す尤度を評価する、および個体の適当な治療コースを決定する段階。

10

【請求項 2】

多因子形質を示す症例群および多因子形質を示さない対照群で関連試験を行い、その結果、対照群より症例群において有意により多く存在する多型座位の対立遺伝子の組を決定することによって複数の二対立遺伝子多型座位に関する関連対立遺伝子および非関連対立遺伝子を同定する段階をさらに含み、ここで該対立遺伝子の組またはそのサブセットが関連対立遺伝子である、請求項1記載の方法。

【請求項 3】

症例群および対照群がそれぞれ少なくとも50個体を含む、請求項2記載の方法。

20

【請求項 4】

症例群および対照群の少なくとも一つが少なくとも100個体を含む、請求項2記載の方法。

【請求項 5】

症例群および対照群の少なくとも一つが少なくとも200個体を含む、請求項2記載の方法。

【請求項 6】

症例群および対照群の少なくとも一つが少なくとも500個体を含む、請求項2記載の方法。

【請求項 7】

症例群および対照群が、哺乳類、爬虫類、両生類、魚類、鳥類、甲殻類、昆虫、植物、細菌、ウイルス、または太古代生物である個体を含む、請求項2記載の方法。

30

【請求項 8】

症例群および対照群がヒトである個体を含む、請求項2記載の方法。

【請求項 9】

症例群および対照群が関連試験を行う前にマッチしている、請求項2記載の方法。

【請求項 10】

関連試験を行う段階が、以下の段階をさらに含む、請求項2記載の方法：

a) 複数の二対立遺伝子多型座位を含む一組の多型座位において症例群および対照群の遺伝子型を同定する段階；

40

b) 症例群および対照群のそれぞれに関する多型座位の組のそれぞれに関して相対的対立遺伝子頻度を計算する段階；

c) 多型座位の組のそれぞれに関して、症例群に関して計算した相対的対立遺伝子頻度を、対照群に関して計算した相対的対立遺伝子頻度と比較して、その結果、多型座位の組のサブセットを同定する段階であって、該サブセットのそれぞれが対照群より症例群に関して有意に異なる相対的対立遺伝子頻度を有する段階；ならびに

d) 対照群より症例群においてより多く存在するサブセットのそれぞれに関する対立遺伝子を決定する段階であって、該対立遺伝子が該関連対立遺伝子の一つである段階。

【請求項 11】

多型座位の組が少なくとも約500の多型座位を含む、請求項10記載の方法。

50

【請求項 1 2】

多型座位の組が少なくとも約1000の多型座位を含む、請求項10記載の方法。

【請求項 1 3】

多型座位の組が少なくとも約10,000の多型座位を含む、請求項10記載の方法。

【請求項 1 4】

多型座位の組が少なくとも約100,000の多型座位を含む、請求項10記載の方法。

【請求項 1 5】

多型座位の組が少なくとも約1,000,000の多型座位を含む、請求項10記載の方法。

【請求項 1 6】

多型座位の組が、個体のゲノムにおける一つまたはそれ以上の染色体からの多型座位を含む、請求項10記載の方法。 10

【請求項 1 7】

多型座位の組が、個体のゲノムにおけるあらゆる染色体からの多型座位を含む、請求項10記載の方法。

【請求項 1 8】

多型座位の組が、個体のゲノムにおけるあらゆる染色体からの複数の多型座位を含む、請求項10記載の方法。

【請求項 1 9】

関連試験が、個体の遺伝子型同定法を用いて行われる、請求項2記載の方法。

【請求項 2 0】

関連試験が、プールされた遺伝子型同定法を用いて行われる、請求項2記載の方法。 20

【請求項 2 1】

個体の遺伝子型同定法を用いて症例群および対照群で第二の関連試験を行い、その結果、第二の関連試験に基づいてどの関連対立遺伝子が対照群より症例群において有意により多く存在するかを決定することにより、関連対立遺伝子を確認する段階をさらに含み、ここで第二の関連試験に基づいて対照群より症例群において有意により多く存在する関連対立遺伝子を確認された関連対立遺伝子である、請求項20記載の方法。

【請求項 2 2】

多因子形質を示す第二の症例群および多因子形質を示さない第二の対照群で第二の関連試験を行い、その結果、どの関連対立遺伝子が第二の対照群より第二の症例群において有意により多く存在するかを決定することにより、関連対立遺伝子を確認する段階をさらに含み、ここで第二の対照群より第二の症例群において有意により多く存在する関連対立遺伝子を確認された関連対立遺伝子である、請求項2記載の方法。 30

【請求項 2 3】

以下の段階を含む方法によって少なくとも一つの閾値の一つを決定する段階をさらに含む、請求項2記載の方法：

- a) 症例群および対照群のそれぞれのメンバーに関するスコアを計算する段階；
- b) 一連のリスクカットオフ値を選択する段階；
- c) 一連のリスクカットオフ値のそれぞれに関して、感度、特異性、PPV、NPV、精度、相対リスク、LR+、およびLR-の少なくとも一つを含む値の組を算出する段階； 40
- d) 該値の組に基づいて少なくとも一つの閾値の一つとして、一連のリスクカットオフ値の一つを選択し、それによって少なくとも一つの閾値の一つを決定する段階。

【請求項 2 4】

症例群および対照群のそれぞれのメンバーに関するスコアを計算する段階が、以下の段階を含む、請求項23記載の方法：

- a) 複数の二対立遺伝子多型座位で、関連対立遺伝子に関してホモ接合、および非関連対立遺伝子に関してヘテロ接合、ホモ接合からなる群より選択される遺伝子型を各メンバーに関して決定する段階；
- b) 関連対立遺伝子ではない対立遺伝子に関してホモ接合である遺伝子型を有する多型座位のそれぞれにゼロの値を割付する段階； 50

- c) ヘテロ接合である遺伝子型を有する多型座位のそれぞれに1の値を割付する段階；
- d) 関連対立遺伝子に関してホモ接合である遺伝子型を有する多型座位のそれぞれに2の値を割付する段階；
- e) 全ての多型座位に関して段階a) からc) において決定された値を合計し、それによって症例群および対照群のそれぞれのメンバーに関するスコアを計算する段階。

【請求項 2 5】

一連のリスクカットオフ値を選択する段階が以下の段階を含む、請求項23記載の方法：
症例群および対照群のそれぞれのメンバーに関して計算されたスコアから最高スコアを同定する段階；

1から最高スコアまでである、リスクカットオフ範囲を決定する段階；

10

リスクカットオフ範囲の中から一連の値を選択して、それによって該一連のリスクカットオフ値を選択する段階。

【請求項 2 6】

リスクカットオフ範囲の中から一連の値を選択する段階が、以下の段階からなる群より選択される方法を含む、請求項25記載の方法：

リスクカットオフ範囲内のあらゆる値を選択する段階；

リスクカットオフ範囲内のあらゆるn番目の値を選択する段階；

リスクカットオフ範囲を除いて百分率にし、リスクカットオフ範囲のあらゆるn番目のパーセント値で値を選択する段階；

リスクカットオフ範囲の上部または下部からより、リスクカットオフ範囲の中央部分から多数の値を選択する段階；および

20

リスクカットオフ範囲の中央部分からより、リスクカットオフ範囲の上部または下部から多数の値を選択する段階。

【請求項 2 7】

所定のリスクカットオフ値に関して、感度が、所定のリスクカットオフ値より大きいスコアを有する症例群のメンバーの比率を決定することによって算出され、比率が所定のリスクカットオフ値に関する感度である、請求項23記載の方法。

【請求項 2 8】

所定のリスクカットオフ値に関して、特異性が、所定のリスクカットオフ値に等しいまたはそれ未満であるスコアを有する対照群のメンバーの比率を決定することによって算出され、比率が所定のリスクカットオフ値に関する特異性である、請求項23記載の方法。

30

【請求項 2 9】

少なくとも一つの閾値の一つを決定する段階が、多因子形質の少なくとも一つおよび個体に関する予備の臨床知識を用いることをさらに含む、請求項23記載の方法。

【請求項 3 0】

予備の臨床知識に、多因子形質の重症度が含まれる、請求項29記載の方法。

【請求項 3 1】

予備の臨床知識に、多因子形質の発生率が含まれる、請求項29記載の方法。

【請求項 3 2】

少なくとも一つの閾値の一つを決定する段階がさらに、c) において算出された感度および特異性に基づいてROC曲線を用いる段階を含み、ROC曲線のグラフによる表示がプロットと呼ばれる、請求項23記載の方法。

40

【請求項 3 3】

少なくとも一つの閾値の一つを決定する段階が、プロットの左上角に最も近いROC曲線の部分から少なくとも一つの閾値の一つを選択する段階をさらに含む、請求項32記載の方法。

【請求項 3 4】

該部分がROC曲線の約20%を含む、請求項33記載の方法。

【請求項 3 5】

ROC曲線上の任意の他のデータポイントより、プロットの左上角により近いROC曲線上の

50

データポイントに対応するリスクカットオフ値を、少なくとも一つの閾値の一つとして選択する段階をさらに含み、ROC曲線上のそれぞれのデータポイントが異なるリスクカットオフ値に対応する、請求項32記載の方法。

【請求項36】

以下をさらに含む、請求項32記載の方法：

a) プロットの左上角に最も近いROC曲線上の位置を決定して、該位置に対応する感度および特異性を決定する段階；

b) その感度および特異性が該位置に対応する感度および特異性に最も近いリスクカットオフ値を同定するために、症例群および対照群のそれぞれのメンバーのスコアを分析する段階であって、その感度および特異性が該位置に対応する感度および特異性に最も近いリスクカットオフ値が、少なくとも一つの閾値の一つである段階。 10

【請求項37】

所定のリスクカットオフ値に関して相対リスクが、以下の段階を含む方法によって算出される、請求項23記載の方法：

a) 所定のリスクカットオフ値と少なくとも同じ大きさであるスコアを有する症例群のメンバーの百分率を決定する段階；

b) 所定のリスクカットオフ値と少なくとも同じ大きさであるスコアを有する対照群のメンバーの百分率を決定する段階；

c) 相対リスクを算出するために、a) において決定された百分率をb) において決定された百分率で除する段階。 20

【請求項38】

相対リスクに多因子形質の発生率を掛けることによって、所定のリスクカットオフ値に等しいスコアを有する所定の個体が、該多因子形質を発症または示すリスクを計算する段階をさらに含む、請求項37記載の方法。

【請求項39】

所定のリスクカットオフ値に関して、PPVが、所定のリスクカットオフ値より高いスコアを有する症例群のメンバーの数を、所定のリスクカットオフ値より大きいスコアを有する症例群および対照群のメンバーの数で除することによって計算される、請求項23記載の方法。

【請求項40】 30

所定のリスクカットオフ値に関して、NPVが、所定のリスクカットオフ値より低いスコアを有する対照群のメンバーの数を、所定のリスクカットオフ値より低いスコアを有する症例群および対照群のメンバーの数で除することによって算出される、請求項23記載の方法。

【請求項41】

多型座位がSNPである、請求項1記載の方法。

【請求項42】

個体が哺乳類、爬虫類、両生類、魚類、鳥類、甲殻類、昆虫、植物、細菌、ウイルス、および太古代生物からなる群より選択される、請求項1記載の方法。

【請求項43】 40

個体がヒトである、請求項1記載の方法。

【請求項44】

スコアを算出する段階がさらに以下の段階を含む、請求項1記載の方法：

a) 関連対立遺伝子ではない対立遺伝子に関してホモ接合である遺伝子型を有する多型座位のそれぞれにゼロの値を割付する段階；

b) ヘテロ接合である遺伝子型を有する多型座位のそれぞれに1の値を割付する段階；

c) 関連対立遺伝子に関してホモ接合である遺伝子型を有する多型座位のそれぞれに2の値を割付する段階；

d) 全ての多型座位に関して段階a) からc) の段階において決定された値を合計し、それによって個体に関するスコアを算出する段階。 50

【請求項 4 5】

多因子形質が疾患である、請求項1記載の方法。

【請求項 4 6】

スコアが少なくとも一つの閾値の一つより大きい場合、疾患を予防するための手段を講じる段階をさらに含む、請求項45記載の方法。

【請求項 4 7】

疾患が薬物に対する有害反応の公知のリスクを付与する、請求項45記載の方法。

【請求項 4 8】

スコアが少なくとも一つの閾値の一つより大きい場合、個体を薬物による治療から除外する段階をさらに含む、請求項47記載の方法。

10

【請求項 4 9】

個体が疾患の家族歴を有する、請求項45記載の方法。

【請求項 5 0】

個体が疾患の症状を示す、請求項45記載の方法。

【請求項 5 1】

多因子形質が薬物に対する反応である、請求項1記載の方法。

【請求項 5 2】

反応が薬物に対する有効な反応の欠如である、請求項51記載の方法。

【請求項 5 3】

スコアが少なくとも一つの閾値の一つより大きい場合に、個体を薬物による治療から除外する段階をさらに含む、請求項52記載の方法。

20

【請求項 5 4】

反応が薬物の投与によって引き起こされた有害事象である、請求項51記載の方法。

【請求項 5 5】

スコアが少なくとも一つの閾値の一つより大きい場合に、個体を薬物の投与から除外する段階をさらに含む、請求項54記載の方法。

【請求項 5 6】

反応が後発品に対する有効な反応であって、後発品が、少なくとも一つの商標付きの薬物を含む薬物のクラスに存在し、個体が後発品に対して有効な反応を有する尤度を用いて、該後発品が個体に投与されるか否かが決定される、請求項51記載の方法。

30

【請求項 5 7】

反応が後発品に対する有効な反応であって、後発品が少なくとも一つの商標付きの薬物を含む薬物のクラスに存在し、個体の、後発品に対して有効な反応を有する尤度を用いて、商標付きの薬物による治療が償還されるか否かが決定される、請求項51記載の方法。

【請求項 5 8】

比較によって、スコアが少なくとも一つの閾値の一つより大きいこと、および個体が多因子形質を示す可能性があることと評価されることが判明する、請求項1記載の方法。

【請求項 5 9】

比較によって、スコアが少なくとも一つの閾値の一つより小さいまたは等しいこと、および個体が多因子形質を示す可能性が低いと評価されることが判明する、請求項1記載の方法。

40

【請求項 6 0】

比較によって、スコアが少なくとも一つの閾値の第一の値より小さいまたは等しいこと、および少なくとも一つの閾値の第二の値より大きいことが明らかになり、さらに個体の適当な治療コースを決定するためにさらなる因子を用いる段階を含む、請求項1記載の方法。

【請求項 6 1】

さらなる要因に、多因子形質に関する情報、個体に関する情報、可能性がある治療選択肢に関する情報、個体からの入力、および管理機関からの入力からなる群の少なくとも一つの要因が含まれる、請求項60記載の方法。

50

【請求項 6 2】

生体試料における請求項1記載の関連対立遺伝子を検出するように設計された核酸プローブを含む診断または予後アッセイ。

【請求項 6 3】

プローブが固体基板に結合している、請求項62記載のアッセイ。

【請求項 6 4】

以下の段階を含む、多因子形質を発症するまたは示す個体の尤度を評価する方法：

a) 複数の二対立遺伝子多型座位で個体の遺伝子型を決定する段階であって、複数のそれぞれが関連対立遺伝子および非関連対立遺伝子を有し、さらに遺伝子型が関連対立遺伝子に関してホモ接合、および非関連対立遺伝子に関してヘテロ接合、ホモ接合からなる群より選択される、ならびにさらに多因子形質を示す症例群と多因子形質を示さない対照群で関連試験を行い、その結果、対照群より症例群において有意により多く存在する多型座位の対立遺伝子の組を決定することによって、関連対立遺伝子および非関連対立遺伝子が同定され、ここで、対立遺伝子の組が関連対立遺伝子であり、およびさらに症例群および対照群が関連試験を行う前にマッチする段階； 10

b) a) において決定された遺伝子型に基づいて個体のスコアを算出する段階であって、スコアを算出する段階が関連対立遺伝子ではない対立遺伝子に関してホモ接合である遺伝子型を有する多型座位のそれぞれにゼロの値を割付する段階；ヘテロ接合である遺伝子型を有する多型座位のそれぞれに1の値を割付する段階；関連対立遺伝子に関してホモ接合である遺伝子型を有する多型座位のそれぞれに2の値を割付する段階；および多型座位の全てに割付された値を合計し、それによって個体のスコアを算出する段階；ならびに 20

c) スコアと少なくとも一つの閾値との比較を行う段階であって、比較が、個体の適切な治療コースを決定するために用いられ、少なくとも一つの閾値のそれぞれが、症例群および対照群のそれぞれのメンバーのスコアを計算する段階；一連のリスクカットオフ値を選択する段階；感度、特異性、PPV、NPV、精度、相対リスク、LR+、LR-、多因子形質に関する臨床データ、個体に関する臨床データ、可能性がある治療選択肢に関する臨床データ、および少なくとも一つの管理機関からの入力 of 少なくとも一つを含む情報を編集する段階；一連のリスクカットオフ値の一つを、該情報に基づいて少なくとも一つの閾値のそれぞれとして選択し、それによって少なくとも一つの閾値のそれぞれを決定する段階を含む方法によって決定される段階。 30

【発明の詳細な説明】

【背景技術】

【0001】

発明の背景

ヒトの染色体を構成するDNAは、体内の全てのタンパク質の産生を指示する説明書を提供する。これらのタンパク質は、生命にとって必須の機能を行う。タンパク質をコードするDNA配列に変化が起これば、コードされるタンパク質の変化または変異が起これ、このように、細胞の正常な機能に影響を及ぼす。環境は疾患においてしばしば重要な役割を果たすが、個体のDNAにおける変化および/または変異は、心血管、代謝、感染疾患、癌および自己免疫障害を含む、ほぼ全てのヒト疾患に直接関係している。その上、遺伝学、特にヒトの遺伝学を知ることによって、多くの疾患がいくつかの遺伝子またはその産物の複雑な相互作用に起因することが理解されるようになった。例えば、I型およびII型糖尿病は、それぞれが自身の変異パターンを有する多数の遺伝子に連鎖している。 40

【0002】

さらに、ヒトの遺伝学を知ることによって、薬理遺伝学の分野である、薬物反応のような個体間の変動が限定的に理解されるようになった。半世紀以上前、有害薬物反応は、二つの薬物代謝酵素、すなわち血漿中のコリンエステラーゼおよびグルコース-6-リン酸デヒドロゲナーゼにおけるアミノ酸の変化に相関した。それ以降、注意深い遺伝子分析によって、35を超える薬物代謝酵素における配列多型（変化）、薬物標的25個、および薬物輸送体5個が、薬物の有効性または安全性レベルの障害に関連した（Evans and Relling, Scie 50

nce 296:487~91(1999) (非特許文献1))。臨床において、そのような情報は、薬物毒性を防止するために用いられている；例えば患者は6-メルカプトプリンまたはアザチオプリンの代謝の減少を引き起こすチオプリンメチルトランスフェラーゼ遺伝子における遺伝的差に関して日常的にスクリーニングされる。しかし、認められた薬物毒性が、今日まで確認された薬理遺伝学マーカーの組によって適切に説明されているのは、ごく小さい割合に過ぎない。毒性問題よりさらに一般的であるのは、数例の個体に関して安全および/または有効であることが証明された薬物が、他の個体では治療有効性が不十分であるか、または予期しない副作用を有するかのいずれかであることが判明した場合であろう。

【0003】

如何なる二人のヒトもその遺伝子構成は99.9%類似であり、そのゲノムのDNA配列のほとんどが同一である。しかし、個体間にはDNA配列に変化がある。例えば、DNAの多塩基の枝の欠失、DNAの枝の挿入、コードまたは非コード領域における反復DNA要素の数の変化、および「一塩基多型 (SNPs)」と呼ばれるゲノムにおける一つの窒素塩基位置での変化が存在する。ヒトDNA配列の変化は、疾患に対する感受性または抵抗性および個体が特定の治療または治療レジメにどのように反応するかを含む、個体間で認められた差の大きい部分を占める。

【0004】

多因子形質または複合形質は、遺伝子、環境要因、およびその相互作用のような多数の要因によって影響を受ける。しばしば、一つより多い遺伝子および/または環境要因の組み合わせによって、同じ多因子形質が起こり、この複雑さのために、誰がそのような形質を発症するであろうかを決定することは難しくなる。さらに、それぞれの要因の関与は典型的に、他の要因の関与と同一ではない。すなわち、例えば、非常に強い関与を有する可能性がある要因もあれば、関与が非常に弱い可能性がある要因もある。多因子形質の生物学的基礎をさらにより複雑にするのは、因子の関与が、他の任意の因子の関与と相加的、相乗的、またはそれから完全に独立している可能性がある点である。そのような複合形質は、心血管疾患、糖尿病、肥満、および高コレステロールのような共通の疾患を表す。他の複合形質には、個体が薬物または他の薬物治療レジメに反応する方法のような表現型が含まれる。

【0005】

近年、疾患の遺伝的基礎に関する研究によって、疾患のいくつかの遺伝子試験が開発された。しかし、これらの遺伝子試験は、健康な人が一般的な多因子疾患を発症する確率を予測するためには有用ではないであろう。一般的な多因子形質 (例えば、疾患) に関する遺伝子試験は、不完全な浸透度および関係する各遺伝子の個々の関与が低いことにより、実践では有用でないであろう (Holtzman and Marteau, 2000 (非特許文献2) ; Vineis et al., 2001 (非特許文献3)) と多くの人が論じている。しかし、これらの議論は、大部分、個体が形質を示すか否かを予測するために単一の座位を用いることに基づいている (Beaudet 1999 (非特許文献4) ; Evans et al. 2001 (非特許文献5))。必要なことは、そのそれぞれが多因子形質の発現における要因である多数の座位での個体の遺伝子型に基づいて、多因子形質を発症または示す個体のリスクを決定するための信頼できるアプローチである。

【0006】

【非特許文献1】 Evans and Relling, Science 296:487~91(1999)

【非特許文献2】 Holtzman and Marteau, 2000

【非特許文献3】 Vineis et al., 2001

【非特許文献4】 Beaudet 1999

【非特許文献5】 Evans et al. 2001

【発明の開示】

【0007】

概要

本出願は、複数の二対立遺伝子多型座位での個体の遺伝子型に基づいて個体のスコアを

10

20

30

40

50

決定する段階、およびそのスコアを少なくとも一つの閾値と比較する段階によって、多因子形質を発症または示す個体のリスクを決定する方法を開示する。特定の態様において、多型座位のそれぞれに関して、個体の遺伝子型は、関連する対立遺伝子に関してホモ接合、無関係な対立遺伝子に関してホモ接合、またはヘテロ接合であってもよい。個体のスコアが閾値より大きい場合、個体は、多因子形質を発症または示すリスクを有すると見なされ、個体のスコアが閾値に等しいまたは閾値より低い場合、個体は、多因子形質を発症または示すリスクを有しないと見なされるであろう。個体のスコアが一つの閾値より大きい、もう一つの閾値より低いまたは等しい場合、個体は、多因子形質を発症または示す中間のリスクを有すると見なされるであろう。

【0008】

10

本出願はさらに、本明細書において「関連対立遺伝子」と呼ばれる多因子形質に関連している二対立遺伝子多型座位の対立遺伝子を同定する方法を開示する。方法は、多因子形質を示す個体群（「症例群」）の遺伝子組成を、多因子形質を示さない個体群（「対照群」）の遺伝子組成と比較する関連試験を行う段階、および対照群の遺伝子組成より症例群の遺伝子組成において有意に多く存在する対立遺伝子を関連対立遺伝子として同定する段階を含む。特定の態様において、第一の症例群と第一の対照群とによる第一の関連試験において同定された関連する対立遺伝子は、第二の症例群と第二の対照群とによる第二の関連試験を行うことによって確認される。

【0009】

本発明はまた、多遺伝子試験において用いるための閾値を決定する方法も開示する。一つの局面において、閾値は、症例群からのスコアの組および対照群からのスコアの組に基づく一連のリスクカットオフ値を分析することによって決定される。閾値の決定は、閾値としてそれぞれのリスクカットオフ値を用いる多遺伝子試験に関する感度、特異性、PPV、NPV、精度、LR+およびLR-が含まれるがそれらに限定されるわけではない情報；多因子形質、可能性がある治療の選択肢、および試験する個体に関する臨床データ；ならびに少なくとも一つの管理機関からの入力を用いることを含む。

20

【0010】

本発明はさらに、生体試料における関連対立遺伝子を検出するように設計された核酸プローブを含む診断または予後アッセイを開示する。特定の態様において、診断または予後アッセイのプローブは、固体基板に結合している。

30

【0011】

詳細な説明

全般

本発明の特定の態様は、例えば疾患もしくは他の障害の発症、または薬物に対する正もしくは負の反応であってもよい、多因子形質を発症または示す個体の素因を非常に確実に決定するための方法を提供する。この決定は、そのそれぞれが、多因子形質の発現に関与している遺伝的要因である、複数の遺伝子座での個体の遺伝子型に基づく。方法はさらに、それぞれの遺伝的要因が多因子形質の発現に影響を及ぼす程度または方法を知らずにそのような決定を行う利益を提供する。代わりに、本発明の方法は多数の遺伝的要因の累積効果に依存し、それによって当業者は、多因子形質の発生に関連していることが決定されている複数の遺伝子座での個体の遺伝子型に基づいて、個体が多因子形質を発症または示す尤度の正確な予測を行うことができる。

40

【0012】

多因子形質は、複数の遺伝的要因、環境的要因、およびそれらの相互作用によって影響を受ける。さらに、それぞれの要因の関与は典型的に、他の要因の関与と同一ではない。すなわち、例えば強い関与を有する可能性がある要因もあれば、関与が弱い可能性がある要因もある。多因子形質の生物学的基礎をさらにより複雑にすることに、因子の関与は、他の任意の因子と相加的、相乗的である、または他の因子から完全に独立している可能性がある。特定の態様において、本明細書に提示の方法は、それぞれの要因が多因子形質に及ぼす効果の程度に依存せず、要因の効果が相加的、相乗的、または独立しているか否か

50

にも依存しない。そのような態様において、方法は、個体がそのような多因子形質を発症する「リスク」（例えば、確率、尤度）を計算する場合に、それぞれの要因の影響の程度を考慮する必要がない。特定の態様において、方法は、ある要因を含む、それに遺伝的に連鎖している、またはその要因の産物である、任意の遺伝子、RNA、またはタンパク質の発現プロファイルに関する知識を必要としない。特定の態様において、方法は、多因子形質に影響を及ぼす可能性がある環境的要因を知る必要はない。特定の態様において、本明細書に示した方法は、それぞれの遺伝的要因の個々の関与が他のあらゆる遺伝的要因の関与と同じである、個々の関与が多因子形質の基礎となる全ての遺伝的要因に関して単に相加的である、および多因子形質の発現に対する環境的要因の関与を知らずに個体のリスクを評価してもよい、という一連の仮定に依存する。他の態様において、方法は、例えばそれぞれの因子が多因子形質に及ぼす効果の程度、および/または多因子形質を発症する個体のリスクを計算する場合に、要因が相加的、相乗的、拮抗的、独立、またはそうでなければ上位性であるか否かのような、要因のさらなる特徴を考慮に入れる。いくつかの態様において、発現データを同様に検討してもよい。さらなる態様において、個体が多因子形質を発症するリスクを計算する場合、多因子形質に影響を及ぼす可能性がある環境的要因に関する知識を考慮する。

10

【0013】

本発明の特定の態様は、多因子形質に関連した多型座位の組を同定するために関連試験を行うための方法を提供する。同様に、関連するどの多型座位の組を、多因子形質に関する多遺伝子試験に含めるかを決定する方法と共に、そのような試験の特定の特徵、例えば感度、特異性、陽性予測値、陰性予測値、相対的リスク、尤度比、精度等を決定するための手段を提供する。多因子形質を発症または示す個体の素因を決定するために、多遺伝子試験において関連する多型座位の組を用いる方法がさらに提供される。一つの態様において、多因子形質は疾患であり、その疾患を発症する可能性があると同定された個体に、疾患の発症を治療または予防するために、治療または他の医学的介入を供してもよい。もう一つの態様において、本発明の方法は、提案された医学治療の有効性を予測するために用いられ、治療が有効でない可能性がある場合、それを患者に投与しない。もう一つの態様において、多因子形質は、薬物治療に反応した有害事象の発現である。有害事象を示す可能性があると同定された個体を、薬物治療レジメから除外してもよく、または薬物（例えば、最後の手段として）によって治療する場合には、有害事象を予測するためにさらなるモニタリングを利用してもよい。なおさらなる態様において、本明細書に開示の方法は医薬品の開発のために用いられ、特に試験の参加に関して適当な患者を選択することによって、薬物の有効性および安全性を増加させるために用いられる。

20

30

【0014】

当業者に容易に明らかであるように、本発明の方法は、対象多因子形質を有する、または発症のリスクを有する個体を同定するために役立つツールとして用いられ、本明細書において示された方法は、形質、試験される個体、および個体が選択される集団に関する臨床情報のみならず、他の臨床検査およびさらには医師の臨床での「直感」と共に用いてもよい。遺伝子試験は典型的に、医師を補助するために用いられるのであって、医師の決定を左右するためではない。本質的に、試験によって不正確に同定された個体に対するリスクと、正確に同定された個体に対する利益とのバランスをとりながら、例えば形質（例えば、疾患）および可能性がある治療の選択肢に関する臨床知識、診断試験、試験が開発された集団、および試験される特定の患者の特徵を用いて、診断または予後試験を用いる方法を決定しなければならないのは、医師である。もう一つの局面において、医師はまた、試験によって「陰性」と不正確に同定された個体に対するリスクと比較して、試験によって「陽性」と不正確に同定された個体に対するリスクを考慮してもよい（例えば、そのような治療を必要とする患者に対する治療を控えることが、それを必要としない患者に治療を行うことより有害となるか？）。

40

【0015】

これから、本発明の様々な態様および特定の応用に対してより詳細に参照する。本発明

50

を、様々な態様および応用に関連して詳細に記述するが、そのような態様および応用は本発明を制限すると意図されないと理解されるであろう。反対に、本発明は、代用、改変、および同等物も含むと意図され、それらも本発明の趣旨および範囲に含まれるであろう。

【0016】

関連試験

本発明の一つの局面において、多因子形質の発現に関連した多型座位の組およびそれらの多型座位に対応する関連対立遺伝子は、関連試験を行うことによって同定され、関連対立遺伝子はさらに、症例群または対照群のメンバーでない個体が、多因子形質の発症または発現に対して遺伝的に素因を有するか否かを決定するために用いられる。多因子形質は、疾患もしくは他の医学的障害の発現、それに対する感受性、もしくは抵抗性、薬物もしくは他の治療レジメに対する反応、またはもう一つの身体的もしくは精神的特徴のような、任意のタイプの表現型形質であってもよい。例えば、一つの態様において、多因子形質は疾患であって、関連試験は、疾患を示す個体群（症例）の遺伝子組成を、疾患を示さない個体群（対照）の遺伝子組成と比較する。多因子である疾患の例には、喘息および他の肺疾患、乾癬、失読症、不妊、痛風、白内障、肥満、糖尿病、消化管障害、癌、心血管疾患、卒中、高血圧症、注意欠陥障害、統合失調症、躁うつ病、骨粗鬆症、免疫系の障害、多発性硬化症、アテローム性動脈硬化症、およびてんかんが含まれるがそれらに限定されるわけではない。口唇/口蓋裂、先天性心欠損、および神経管欠損のような、特定の発達異常も同様にこの分類に含まれる。もう一つの態様において、他因子形質は、薬物に対する反応であり、関連試験は、薬物に対して特定の反応を示す個体群（症例）の遺伝子組成を、特定の反応を示さない個体群（対照）の遺伝子組成と比較する。一つの局面において、薬物反応は、薬物の有効性に関連してもよい。例えば、薬物は、症例群における個体にとって非常に有効であって、対照群における個体にとって有効性が不良であってもよく、またはこの逆であってもよい。もう一つの局面において、薬物反応は、薬物の投与に反応した有害事象に関連してもよい。例えば、症例群の個体は薬物に反応して有害事象を示してもよく、対照群の個体は、有害事象を示さなくてもよい。特定の多因子形質と組み合わせて本発明の方法を用いることを記述する様々な例が本明細書において提供されるが、これらの例は、本発明の範囲を制限すると解釈されず、それらは、その発現が複数の遺伝子座位を含む任意の多因子形質と共に本明細書において示される方法を用いることを含む。

【0017】

典型的に、個体少なくとも50例、好ましくは少なくとも100例が症例および対照群の双方に存在する。いくつかの試験において、症例群および対照群の少なくとも一つに、個体少なくとも200例または少なくとも500例が存在する。しばしば、対照群の個体は症例群の個体より多い。特定の態様において、症例群および対照群の個体は哺乳類であるが、症例群および対照群はまた、例えば細菌、真菌、原生生物、ウイルス、太古代生物のような非哺乳類個体、および爬虫類、両生類、魚類、鳥類、甲殻類、昆虫、および植物のような他の真核生物を含んでもよい。いくつかの態様において、症例および対照群における個体はヒトである。

【0018】

典型的に、症例群および対照群の組成は、検討される多因子形質とは別の特徴に関して類似でなければならない。例えば、一つの態様において、類似の年齢の類似の数の男女がそれぞれの群に関して選択されるであろう。特定の態様において、環境の危険因子は症例群および対照群の組成に影響を及ぼす可能性がある。例えば、肺癌に関する試験に関して喫煙者のみ（または非喫煙者のみ）を症例および対照群を含むように選択してもよい。本発明のいくつかの態様において、症例および対照群の構成体を、関連試験を行う前に、二つの群の集団の構造が「マッチ」するように調節する。集団の構造（または、「集団の階層化」）は、集団における個体の遺伝子組成の不均一性を指す。例えば、主にイタリア人で構成される症例群の集団構造は、二つの群の民族起源が異なるために、主にメキシコ人で構成される対照群の集団構造とは異なる。群をマッチさせないで関連試験を行うと、イタリア人祖先に関連するが、メキシコ人祖先には関連しない遺伝子座位が、誤って試験中

10

20

30

40

50

の多因子形質に関連するようになることがある。症例および対照群の集団構造をマッチさせることによって、当業者は、対象の多因子形質に関連しない、症例群および対照群のあいだの遺伝的な差を制御することができる。したがって、その後の関連試験によって同定された群のあいだの遺伝的な差は、対象の多因子形質に偶然に関連した座位である可能性がより高い。関連試験を行う前に症例群および対照群をマッチさせる方法は、2003年4月30日に提出された、「Method for Identifying Matched Groups」と題する米国実用特許出願第10/427,696号および2003年8月26日に提出された、「Matching Strategies for Genetic Association Studies in Structured Populations」と題する米国特許仮出願第60/497,771号において詳細に記述されている。

【0019】

遺伝子型同定アッセイにおいて用いるために、核酸試料を症例群および対照群の個体から採取する。核酸試料は、DNAまたはRNAであってもよく、例えば全血、精液、唾液、類液、便、尿、汗、口腔、皮膚、および毛髪のような様々な生体試料から得てもよい。特定の局面において、核酸試料はゲノムDNA試料を含む。試料核酸は、当業者に公知の任意の技術を用いて分析のために調製してもよい。好ましくは、そのような技術によって、核酸分子における一つまたはそれ以上の位置で一つまたはそれ以上の多型の有無を決定するために十分に純粋な核酸分子が産生される。そのような技術は一般的に公知であり、例えば、Sambrook, et al., 「Molecular Cloning : A Laboratory Manual」 (Cold Spring Harbor Laboratory, New York) (2001) および Ausubel, et al., 「Current Protocols in Molecular Biology」 (John Wiley and Sons, New York) において認められるであろう。

【0020】

一つまたはそれ以上の対象核酸を増幅および/または標識してから、核酸における一つまたはそれ以上の多型の有無を決定してもよい。ポリメラーゼ連鎖反応 (PCR) 技術を含む、当業者に公知の任意の増幅技術を、本発明の特定の方法及び共に用いてもよい。PCRは、当業者に公知の材料および方法を用いて行ってもよい。一般的に「PCR Technology : Principles and Applications for DNA Amplification」 (ed. H.A. Erlich, Freeman Press, NY, NY, 1992) ; 「PCR Protocols : A Guide to Methods and Applications」 (eds. Innis, et al., Academic Press, San Diego, CA, 1990) ; Matilla et al., Nucleic Acids Res. 19:4967(1991) ; Eckert et al., PCR Methods and Applications 1:17(1991) ; PCR (eds. McPherson et al., IRL Press, Oxford) ; 米国特許第4,683,202号を参照されたい。他の適した増幅法には、リガーゼ連鎖反応 (LCR) (Wu and Wallace, Genomics 4:560(1989) および Landegren et al., Science 241:1077(1988) を参照されたい)、転写増幅 (Kwoh et al., Proc. Natl. Acad. Sci. USA 86:1173(1989))、self-sustained sequence replication (Guatelli et al., Proc. Nat. Acad. Sci. USA 87:1874(1990)) および nucleic acid-based sequence amplification (NASBA) が含まれる。さらに、2002年3月26日に提出された、「Methods for Genomic Analysis」と題する継続中の米国特許出願第10/106,097号、2002年1月9日に提出された「Algorithms for Selection of Primer Pairs」と題する第10/042,406号、2002年9月5日に提出された「Methods for Amplification of Nucleic Acids」と題する第10/236,480号、2002年6月17日に提出された「Methods for Storage of Reaction Cocktails」と題する第10/174,101号、2003年5月28日に提出された「Liver Related Disease Compositions and Methods」と題する第10/447,685号、2004年3月4日に提出された「Apparatus and Methods for Analyzing and Characterizing Nucleic Acid Sequences」と題する第10/768,788号、および2003年4月30日に提出された「Method for Identifying Matched Groups」と題する第10/427,696号に開示された方法は、本発明の特定の方法において用いるために、核酸を増幅、標識、またはさらに操作 (すなわち断片化) するために適している。

【0021】

関連試験において、多型であることが知られている遺伝子座 (例えば、SNPs) を、症例群および対照群のそれぞれにおいて各個体に関して遺伝子型を同定して、相対的対立遺伝子頻度を、群に存在する遺伝子型に基づいて群のそれぞれに関する座位のそれぞれに関し

10

20

30

40

50

て計算する。すなわち、多型座位10個を遺伝子型同定した場合、症例および対照群のそれぞれに関して10づつである、20の相対的対立遺伝子頻度が決定される。所定の多型座位に関して、症例群の相対的対立遺伝子頻度を対照群の頻度と比較して、多型座位が対照群より症例群において有意に異なる相対的対立遺伝子頻度を有する場合、それは症例群と対照群を区別する多因子形質に関連する可能性がある座であると同定される（「関連座」）。特定の態様において、相対的対立遺伝子頻度の有意な差は、約5%より大きい、約8%より大きい、約10%より大きい、約12%より大きい、または約15%より大きい差である。症例集団においてよりしばしば存在する対立遺伝子は、「関連対立遺伝子」と呼ばれることがあり、対象集団においてよりしばしば存在する対立遺伝子は、「非関連対立遺伝子」と呼ばれることがある。同定された関連座位（およびしたがって、二対立遺伝子関連座位の関連対立遺伝子）は、試験中の多因子形質（例えば、疾患）に関与するまたは関与する座位と連鎖不平衡にある多型座位の数に応じて広く変化するであろう。例えば、疾患の発現が遺伝子10個を含む場合、同定された関連座位の数は、関連試験において遺伝子型を同定された多型座位が疾患を引き起こす遺伝子10個の対立遺伝子と連鎖不平衡にある数に依存するであろう。典型的に、多因子形質の発現に係る座位の数は約5～数百個に及ぶが、これより多いまたは少なくともよい。症例および対照群の相対的対立遺伝子頻度を用いて関連試験を行うための方法に関する詳細な記述に関しては、その双方が「Apparatus and Methods for Analyzing and Characterizing Nucleic Acid Sequences」と題する2003年4月3日に提出された米国特許出願第60/460,329号、2004年1月30日に提出された第10/768,788号を参照されたい。

10

20

【0022】

個体の遺伝子型を同定は当業者に公知の任意の技術を用いて行ってもよい。好ましい技術によって、試料の取り扱いを最小限にして多数の変種の迅速で正確な定量を行うことができる。適した技術のいくつかの例は、直接DNAシーケンシング、キャピラリー電気泳動、ハイブリダイゼーション、対立遺伝子特異的プローブまたはプライマー、一本鎖構造多型分析、核酸アレイ、ビーズアレイ、制限断片長多型分析、クレバーゼ断片長多型分析、ランダム増幅多型DNA、リガーゼ検出反応、ヘテロ二本鎖または断片分析、質量分析を伴うディフュゼンシャルシーケンシング、原子力顕微鏡、ピロシーケンシング、FRET（例えば、TaqMan (Applied Biosystems, Inc., Foster City, CA)）および分子ビーコン (Stratagene, La Jolla, CA) アッセイ、ならびに当技術分野で周知の他の技術を含むがこれらに限定されない。DNAシーケンシングに関するいくつかの方法は周知であり、一般的に当技術分野で利用可能である。例えば、Sambrook, et al., 「Molecular Cloning: A Laboratory Manual」 (Cold Spring Harbor Laboratory, New York) (2001) および Ausubel, et al., 「Current Protocols in Molecular Biology」 (John Wiley and Sons, New York) (1997)、Twyman, et al. (2003)、 「Techniques Patents for SNP Genotyping」, Pharmacogenomics 4(1):67~79; および Kristensen et al., (2001) 「High-Throughput Methods for Detection of Genetic Variation」, BioTechniques 30(2):318~332を参照されたい。例えばSNPsを検出するために核酸アレイ (DNAチップ) を用いることの詳細に関しては、Lipshultzらに交付された米国特許第6,300,063号およびCheeらに対する米国特許第5,837,832号、HuSNPマッピングアッセイ、試薬キットおよびユーザーマニュアル、Affymetrixパーツ番号90094号 (Affymetrix, Santa Clara, CA) を参照されたい。

30

40

【0023】

症例または対照群に関する相対的対立遺伝子頻度は、集団における個々の個体における各対立遺伝子の正確な量を決定するために、集団における個体の全てを個々に遺伝子型を同定することによって、直接決定してもよい。複数の個体を個々に遺伝子型を同定する方法は、2003年1月27日に提出された、「Apparatus and Methods for Determining Individual Genotypes」と題する米国特許出願第10/351,973号、および2004年2月24日に提出された、「Improvements to Analysis Methods for Individual Genotyping」と題する米国特許出願（まだ割付されていない）代理人整理番号第100/1046-20号において詳細に記述されている。または、プールされた遺伝子型同定を用いて、症例群および対照群のそれぞれ

50

に関する相対的対立遺伝子頻度を決定してもよい。プールされた遺伝子型同定の場合、症例群からの核酸試料を共にプールして（症例プール）、対照群からの核酸試料を共にプールして（対照プール）、症例および対照プールの分析を通して症例群および対照群に関する相対的対立遺伝子頻度を決定する。プールされた遺伝子型同定の方法は、その双方が「Apparatus and Methods for Analyzing and Characterizing Nucleic Acid Sequences」と題する、2003年4月3日に提出された米国特許出願第60/460,329号、および2004年1月30日に提出された第10/768,788号に詳細に考察されている。

【0024】

遺伝子座

「SNP」または「一塩基多型」という用語は、例えば、生存している生物のDNAにおける一つの窒素塩基位置での個体間の遺伝子変化を指す。SNPsは、ゲノム中に認められ、個体間の遺伝子変化の多くは、SNP座での変化により、しばしばこの遺伝子の変化によって、個体間の表現型の変化が起こる。本発明において用いられるSNPおよびそのそれぞれの対立遺伝子は、公共のデータベース（U.C. Santa Cruz Human Genome Browser Gateway (<http://genome.ucsc.edu/cgi-bin/hgGateway>)）もしくはNCBI dbSNPウェブサイト (<http://www.ncbi.nlm.nih.gov/SNP/>) のような多くの起源に由来してもよく、または2002年3月26日に提出された、「Methods for Genomic Analysis」と題する米国特許出願第10/106,097号、および2002年10月31日に提出された「Human Genomic Polymorphisms」と題する第10/284,444号に記述されるように、実験的に決定してもよい。SNPsを用いることは、本明細書に示した態様のいくつかにおいて記述されるが、他の二対立遺伝子遺伝子マーカーも同様に用いてもよいと理解されるであろう。二対立遺伝子遺伝子マーカーは、二つの多型型または対立遺伝子を有するマーカーである。先に述べたように、形質に関連する二対立遺伝子遺伝子マーカーに関して、対照群と比較して症例群の遺伝子組成においてより多く存在する対立遺伝子を「関連対立遺伝子」と呼び、他の対立遺伝子を「非関連対立遺伝子」と呼んでもよい。このように、所定の形質（例えば、疾患または薬物反応）に関連するそれぞれの二対立遺伝子多型に関して、対応する関連対立遺伝子が存在する。本明細書において示される方法と共に用いてもよい他の二対立遺伝子多型には、多ヌクレオチド変化、挿入、欠失、および転座が含まれるがそれらに限定されるわけではない。本明細書においてDNAという言葉には、アンプリコンのようなDNAの誘導體、RNA転写物、cDNA、DNA類似体等が含まれてもよいとさらに認識されるであろう。関連試験においてスクリーニングされる多型座位は、二倍体または半数体状態であってもよく、理想的には、ゲノム全体の部位に由来するであろう。

【0025】

本発明のいくつかの態様において、関連試験は、SNP少なくとも約100個、SNP少なくとも約500個、SNP少なくとも約1000個、SNP少なくとも約10,000個、SNP少なくとも約100,000個、またはSNP少なくとも約1,000,000個をスクリーニングすることを含む。特定の態様において、多因子形質に関連すると考えられるゲノムの一つまたはそれ以上の部位に存在するSNPをスクリーニングする。他の態様において、一つまたはそれ以上の染色体上のSNPをスクリーニングする。なおさらなる態様において、ゲノムにおけるあらゆる染色体からのSNPをスクリーニングする。他の態様において、ゲノムにおけるあらゆる染色体からの多数のSNPをスクリーニングする。他の態様において、遺伝子のコード領域または調節領域に存在するSNPをスクリーニングする。さらなる態様において、遺伝子の異なる対立遺伝子発現に関連することが判明しているSNPをスクリーニングする。（ディファレンシャル対立遺伝子発現は、遺伝子の一つの対立遺伝子が、ヘテロ接合である同じ遺伝子のもう一つの対立遺伝子より高いレベルで発現される場合に起こり、これは2003年5月13日に提出された、「Allele-specific Expression Patterns」と題する米国特許出願第10/438,184号に詳細に記述されている。）特定の態様において、全ての公知のSNP（今日まで約300万個）をスクリーニングする。他の態様において、スクリーニングされていないSNPのサブセットの対立遺伝子組成を予測するために用いてもよいSNPのサブセットをスクリーニングする。本明細書に示される方法によってスクリーニングされるSNPは、個体において

二倍体または半数体状態のいずれかであってもよい。

【0026】

本明細書において示される方法によって同定された関連するSNP（および従って関連する対立遺伝子）の数は、いくつかの基準に依存する。第一に、これは疾患の発現に係る遺伝子座の数に依存する。例えば、多因子疾患に関する遺伝的基礎がごく少数の座位を含む場合、関連SNPおよび関連対立遺伝子の数は典型的に、その遺伝的基礎が何百もの座位を含む多因子疾患の場合に認められる数より少ないであろう。さらに、同定された関連SNPおよび関連対立遺伝子の数は、関連試験においてスクリーニングされたSNPの数に依存する。例えば、症例群および対照群においてSNP 100個のみをスクリーニングする関連試験は、SNP 100万個をスクリーニングする試験より、多数の関連SNPを発見する可能性はより低いであろう。典型的に、本明細書に示した方法は、関連SNP/関連対立遺伝子約10個～数百個を同定するが、それより多くまたは少なく同定することもある。 10

【0027】

関連対立遺伝子の組のバリデーション

一つの態様において、関連する対立遺伝子の同定を確認するために、第二の症例および第二の対照集団を用いて、関連試験を繰り返す。この第二の関連試験は、第一の関連試験からのそれらの関連対立遺伝子がおも、新しい組の症例および対照の相対的対立遺伝子頻度に基づいて関連対立遺伝子として同定されるか否かを決定し、「同じ結果が得られた」それらがそれによって関連SNPであると確認されている。特定の態様において、第一の関連試験においてスクリーニングされた多型座位を、第二の確認関連試験において同様にスクリーニングする。他の態様において、第一の関連試験においてスクリーニングされた多型座位のサブセットを、第二の確認関連試験においてスクリーニングする。特定の態様において、第二の関連試験においてスクリーニングされる多型座位の組は、第一の関連試験によって同定された関連多型座位を含む。例えば、第一の関連試験においてSNP 30,000個が疾患の発生に関連していると同定される場合、次に、それらのSNP 30,000個を、疾患を示す個体の第二の症例群および疾患を示さない第二の対照群が選択される第二の関連試験においてスクリーニングする。特定の態様において、第二の症例群は、同じ基準に従って第一の症例群として選択され、第二の対照群は同じ基準に従って第一の対照群として選択される。一つの局面において、第一および第二の症例群ならびに第一および第二の対照群は、共通のメンバーを有しない。プールされたまたは個々の遺伝子型同定法を用いて、 20 30

【0028】

他の局面において、プール遺伝子型同定を用いて関連試験を行う場合、プール遺伝子型同定法によって決定された関連対立遺伝子の組を、症例および対照群のあらゆる個体における関連SNPの組を個々に遺伝子型同定すること、および相対的対立遺伝子頻度を再計算して再比較することによって確認してもよい。個々の遺伝子型同定データに基づいて対照群と比較して症例群において有意に高い対立遺伝子頻度を有する最初のプール遺伝子型同定分析に基づいて同定された関連対立遺伝子は、それによって関連対立遺伝子であると確認される。このバリデーション段階を、プール遺伝子型同定法を利用する第一の関連試験に関して用いてもよく、またはプール遺伝子型同定法を用いる第二の確認関連試験に 40

【0029】

関連SNPの組を同定するための試験デザインにおいて、一つより多いバリデーション法を用いてもよい。例えば、本発明の一つの態様において、疾患を示す個体の症例集団および疾患を示さない個体の対照集団に関して、第一の関連試験を行う。プール遺伝子型同定法を用いて、症例群および対照群のあいだで有意差を示す相対的対立遺伝子頻度を有するSNP約30,000個を同定するために、SNP座約150万個において症例および対照群を遺伝子型同定する。次のバリデーション段階において、個々の遺伝子型同定法に基づいて、対照群より症例群において有意に異なる相対的対立遺伝子頻度を有するSNP約300個を同定するために、「プールされた」関連試験において同定されたSNP約30,000個のそれぞれにおいて 50

、症例および対照群を個々に遺伝子型を同定する。このように、これらのSNP約300個を、個々の遺伝子型同定によって確認する。さらなるバリデーション段階において、個々の遺伝子型同定段階によって確認されたSNP約300個が、第二の症例群および第二の対照群による個々の遺伝子型同定法に基づく第二の関連試験を行うことによってさらに確認される、第二の関連試験を行う。第二の関連試験において同じ結果が得られたそれらのSNPは、疾患に関する関連SNPであると分類され、対照群より症例群において多く存在する関連SNPの対立遺伝子を、関連対立遺伝子と呼ぶ。

【0030】

リスクカットオフを決定するための関連対立遺伝子の利用

本発明の一つの態様において、それぞれの疾患関連SNP座における症例群および対照群の個体の遺伝子型を用いて、症例群を対照群と区別する多因子形質の発症に関して個体の素因を決定するために用いられる一連のカットオフ値を発見する。 10

【0031】

一つの局面において、それぞれの関連するSNP位置での遺伝子型を、症例群および対照群における全ての個体から得る。先に考察したように、関連試験の際に個々の遺伝子型を同定を行う場合、関連試験の際の関連SNP位置に関して収集した遺伝子型同定データを用いてもよい。しかし、個々の遺伝子型が決定されていない場合、症例および対照群のそれぞれのメンバーを、関連するSNPの組に関して個々に遺伝子型を同定しなければならない。例えば、二対立遺伝子SNPの場合、二倍体個体は、関連対立遺伝子に関してホモ接合である、非関連対立遺伝子に関してホモ接合である、およびヘテロ接合である（関連対立遺伝子1個と非関連対立遺伝子1個とを有する）異なる三つの遺伝子型の一つを有してもよい。 20

本明細書に示した方法は、また半数体生物に適用してもよく、または二倍体生物における半数体座（例えば、ヒトにおけるY染色体座）に適用してもよい。半数体座の場合、それぞれの可能性がある対立遺伝子に関して一つずつの二つのみの遺伝子型が存在するであろう。

【0032】

もう一つの局面において、症例および対照群におけるそれぞれの個体に、関連SNP座のそれぞれにおいてその遺伝子型に基づくスコアを割付する。本発明の一つの態様において、各関連対立遺伝子に1ポイントを与え、それにより関連対立遺伝子に関してホモ接合である各SNP遺伝子型には2ポイントが与えられ、ヘテロ接合である各SNP遺伝子型には1ポイントが与えられ、非関連対立遺伝子に関してホモ接合である各SNP遺伝子型にはゼロポイントが与えられる。半数体座に関する一つの態様において、関連対立遺伝子を有する各SNP遺伝子型は1ポイントが与えられ、非関連対立遺伝子を有する各SNP遺伝子型にはゼロポイントが与えられる。半数体座に関するもう一つの態様において、関連対立遺伝子を有する各SNP遺伝子型には2ポイントが与えられ、非関連対立遺伝子を有する各SNP遺伝子型にはゼロポイントが与えられる。 30

所定の個体に関して、関連SNPの全てのポイントを合計して、その個体に関するスコアを提供する。例えば、関連SNP 100個を遺伝子型同定する場合、個体に関する最大スコアは200であり、このことは、あらゆる関連SNP位置において関連対立遺伝子2個を有することを意味する。言い換えれば、個体はあらゆるSNP位置において関連対立遺伝子に関してホモ接合である。最小スコアは、任意の関連SNP位置で関連対立遺伝子を有しない個体の場合のゼロであり、またはあらゆるSNP位置において非関連対立遺伝子に関してホモ接合である。スコアを、症例および対照群におけるあらゆる個体に関して計算する。例えば、関連SNP 100個を個体102例の症例集団および個体405例の対照集団に関して調べた。症例群における最低スコアは42であり、最高スコアは97であった；対照群に関して最低スコアは23であり、最高スコアは79であった。これは、それぞれの要因が多因子形質に及ぼす影響の程度に依存せず、要因の影響が相加的、相乗的、または独立しているか否かにも依存しない個体に関してスコアを決定するための態様の例である。 40

さらに、この態様は、任意の因子を含む、遺伝的に連鎖している、またはその産物である任意の遺伝子、RNAまたはタンパク質の発現プロファイルに関する知識を必要としない。さらに、態様は、多因子形質に影響を及ぼす可能性がある環境的要因に関する知識を必要 50

としない。

【0033】

本発明の他の態様において、個体のスコアを作製する場合に、例えばそれぞれの因子が多因子形質に及ぼす影響の程度、および/または個々のスコアを計算する場合に、要因が相加的、相乗的、拮抗的、独立的、またはそうでなければ上位性であるか否か、のような他の要因を考慮に入れる。例えば、第一の対立遺伝子は第二の対立遺伝子より個体のリスクに対して2倍高い影響を有する可能性がある場合、第一の対立遺伝子に割付される値は、個体のスコアを計算する場合に、第二の対立遺伝子に割付される値より2倍高いであろう。いくつかの態様において、所定の組の要因が異なる上位性相互作用の組み合わせを示す可能性がある。他の態様において、発現データを同様に検討してもよい。例えば、癌の発症の基礎となる要因である遺伝子（またはそのRNAもしくはタンパク質産物）の特定の対立遺伝子の発現レベルは、癌の発症のリスクの増加または減少を予測する可能性がある。さらなる態様において、多因子形質に影響を及ぼす可能性がある環境的要因に関する知識を、個体のスコアを計算する場合に考慮しなければならない。そのような態様において、一つまたはそれ以上の環境的要因の存在に関連するポイントを個体のスコアの計算に用いる。

【0034】

もう一つの局面において、一連のリスクカットオフ値を決定する。リスクカットオフ値は、個体が多因子形質を発症または示す可能性を同定するために遺伝子試験において用いられる仮説上の閾値を表す。例えば、閾値より高いスコアを有する個体は、多因子形質を示す可能性があるとして診断される可能性があり、閾値より低いスコアを有する個体は、多因子形質を示す可能性がより低いと診断される可能性がある。または、多因子形質を示す個体のリスクを決定するために、多数の閾値を用いてもよい。

【0035】

一連のリスクカットオフ値は、それらが症例群または対照群のメンバーであるか否かによらず、1から関連試験において個体に関して計算した最高スコアまでの範囲である。一つの例において、個体に関する最高スコアは97ポイントであり、そこからリスクカットオフ値を決定する範囲（リスクカットオフ範囲）は、1～97のあいだであった。特定の局面において、リスクカットオフ値は、リスクカットオフ範囲から選択されるが、特定のリスクカットオフ値の選択はいくぶん任意である。いくつかの態様において、リスクカットオフ範囲におけるあらゆるスコアが選択される。他の態様において、あらゆるn番目のスコア（例えば、5または10番目の）が選択される。なおさらなる態様において、範囲を百分率にして、n番目の百分率を選択する。いくつかの態様において、より多くのリスクカットオフ値が、範囲の上の部分または下の部分より、スコアの完全な範囲の中央部分から選択される、またはその逆である。例えば、スコアの完全な範囲が1～97の場合、20から80までの10番目のスコア毎にリスクカットオフ値を選択して、この範囲の中央部分をよりよく評価するために、さらなるリスクカットオフ値55および65が付加された（表1を参照されたい）。

【0036】

その後の段階において、リスクカットオフ値のそれぞれを、症例群および対照群における個体に関して計算したスコアと比較する。詳しく述べると、症例（「罹患」）および対照（「非罹患」）個体を用いて、多因子形質を示す可能性がある個体と、多因子形質を示さない可能性がある個体とを区別するために、どのリスクカットオフ値が最善の感度、特異性、陽性予測値（PPV）、陰性予測値（NPV）、精度、またはその組み合わせを提供するかを決定して、それによって関連SNPを用いて多遺伝子試験に関する良好な閾値となるであろうリスクカットオフ値を同定する。さらに、適当な閾値の同定はさらに、臨床情報（例えば、多因子形質、試験個体群、または被験個体に関して）、および/または本発明の実践者の外部団体（例えば、米国食品医薬品局（FDA））との協合作用を用いることを含んでもよい。この閾値は、症例および対照群の個体に関する閾値およびスコアに基づいて計算された感度、特異性、PPV、NPV、および精度と共に、遺伝子試験、例えば診断試験に

発展させてもよい。

【0037】

二つのクラスの遺伝子試験は可能性のある二つの結果を有する。陽性試験結果は、個体
が対象形質を示すまたは示す可能性があることを示し、陰性試験結果は、個体が対象形質
を示さないおよび示さない可能性があることを示している。そのため、遺伝子試験の信頼
性は、試験の結果が、個体を形質に関して「陽性」または「陰性」と正しく同定する回数
に関連する。真の陽性（TP）および真の陰性（TN）は、個体を陽性（例えば、「罹患」）
および陰性（例えば、「非罹患」）と正確に同定する試験結果である。偽陽性（FP）は、
それらが実際に形質に関して陰性である場合に、個体を陽性であると不正確に分類する試
験結果である。同様に、偽陰性（FN）は、それらが実際に形質に関して陽性である場合に
個体を陰性であると不正確に分類する試験結果である。TP、NP、FP、およびFNの測定値は
、遺伝子試験に関して感度、特異性、PPV、およびNPVを計算するために用いられる。 10

【0038】

試験の「感度」は、試験が罹患個体または対象形質を発症するであろう個体を正確に同
定できる能力の測定である。感度が1に近ければ、試験は罹患個体の同定においてより正
確である。詳しく述べると、感度は、試験によってそのように正確に診断される罹患個体
の比率を指し、罹患していると正確に同定された個体数（TP）を罹患個体の総数（TP+FN
）で除した値として計算される。ほとんどの罹患個体が遺伝子試験によってそのように同
定されるように、高い感度が好ましい。試験の「特異性」は、試験が非罹患個体、または
対象形質を発症しない個体を正確に同定することができる測定値である。特異性が1に近
づけば、試験は、非罹患個体の同定においてより正確である。詳しく述べると、特異性は
、試験によってそのように正しく同定される非罹患個体の比率を指し、非罹患（TN）であ
ると正確に同定された個体数を非罹患個体の総数（TN+FP）で除した値として計算される
。罹患していると不正確に同定される個体数が最小限となるように、高い特異性が好まし
い。このように、所定のリスクカットオフ値に関して、感度は、リスクカットオフ値より
高いスコアを有する症例個体の比率として計算され、特異性は、リスクカットオフ値より
低いまたはそれに等しいスコアを有する対照個体の比率（または、 $1 - (\text{リスクカットオフ値より高いスコアを有する対照個体の比率})$ ）として計算される。 20

【0039】

遺伝子試験の「陽性予測値」（PPV）は、陽性試験の転帰/結果の信頼性を評価して、実
際に対象形質を有する陽性試験結果を有する人々の割合として計算される。言い換えれば
、これは、陽性試験結果が形質を有する個体を正確に同定する確率であり、罹患していると
正確に同定された個体数（TP）を、遺伝子試験によって罹患していると同定された個体
の総数（TP+FP）によって除した値として計算される。多くの場合、罹患していると同定
されたほとんどの個体が実際に罹患しているように、高いPPVが好ましい。例えば、PPV 0
.98は、陽性試験結果を有する個体が、形質を有するまたは発症する機会を98%有するこ
とを意味する。遺伝子試験の「陰性予測値」（NPV）は、陰性試験転帰/結果の信頼性を評
価して、対象形質を有しない陰性試験結果を有する人々の比率として計算される。言い換
えれば、陰性試験結果は、形質を有しない個体を正確に同定する確率であり、非罹患（TF
）と正確に同定された個体数を、非罹患であると同定された個体の総数（TN+FN）で除し
た値として計算される。高いNPVは、非罹患であると同定されたほとんどの個体が実際に
非罹患であることから時に好ましい（例えば、特定の薬物の投与に関連した有害事象に関
してリスクを有する被験者を除外する場合に）。例えば、NPV 0.999は、陰性試験結果を
有する個体が形質を有するまたは発症する確率（例えば、薬物に反応して有害事象を経験
する確率）が0.1%に過ぎないことを意味する。このように、所定のリスクカットオフ値
に関して、PPVは、症例群における実際のリスクカットオフ値より高いスコアを有する全
ての個体の集団として計算してもよく、NPVは、対照群における実際のリスクカットオフ
値より低いまたはそれに等しいスコアを有する全ての個体の比率として計算される。 30 40

【0040】

ある形質の出現率が、調べる集団における形質の頻度であって、所定の時間で総集団に 50

よって除した既存の症例数として計算される。試験の感受性および特異性は、検討される形質の発生率によって影響を受けないが、PPVおよびNPVはいずれも調べる集団における形質の発生率によって非常に影響を受ける；より低い疾患の発生率によって、より低いPPVおよびより高いNPVが得られる。PPVおよびNPVはいずれも、感度（sens）、特異性（spec）、および発生率（prev）の関数として計算してもよい。

$$PPV = (\text{sens})(\text{prev}) / [(\text{sens})(\text{prev}) + (1 - \text{spec})(1 - \text{prev})]$$

$$NPV = (\text{spec})(1 - \text{prev}) / [(\text{spec})(1 - \text{prev}) + (1 - \text{sens})(\text{prev})]$$

【0041】

閾値はまた、遺伝子試験に関する尤度比を用いて選択してもよい。尤度比（LR）は、試験の感度および特異性を一つの測定に組み入れる方法であり、陽性または陰性試験結果に基づいて所定の形質変化を有するまたは発症するオッズがどのくらいであるかに関する指標を示す。感度および特異性は、試験自身の固定の特徴であることから、LRは、PPVおよびNPVとは異なり、集団における形質の発生率とは独立している。LRは、形質を有しない個体において同じ結果が予測されるであろう尤度と比較した、形質を有する個体において所定の試験結果が予測される尤度である。陽性試験結果に関するLR（LR+）は、試験が陽性である場合に、形質を有するまたは発症する個体のオッズがどれくらいであるかの測定を提供し、（1-特異性）によって除した感度として計算される。形質を「含める」ために用いられるよりよい試験は、最大のLR+を有する試験である。陰性試験結果に関するLR（LR-）は、試験が陰性である場合に、形質を有するまたは発症する個体のオッズがどのくらい減少するかに関する手段を提供し、特異性によって除した（1-感度）として計算される。形質を「除外する」ために用いられるよりよい試験は、より小さいLR-を有する試験である。10より大きいまたは0.1未満のLRは通常、高い診断的価値を有すると判断される。LRを、「試験前オッズ」と組み合わせ、調べた個体が対象形質を有するまたは発症するであろう「試験後オッズ」を決定する（試験後オッズ＝試験前オッズ×LR）。試験前オッズは、形質の発生率に関する情報、集団の特徴、および試験される特定の個体に関する情報と共に計算され、試験前に個体が形質を有するまたは形質を発症する尤度を表す。試験後のオッズは、個体が、試験結果を考慮して形質を有するまたは発症する尤度を表す。本発明の一つの態様において、遺伝子試験に関してLRを最大限にする閾値が選択される。

【0042】

遺伝子試験の価値または有用性に関するさらにもう一つの測定値は、試験結果と実際の病的状態との全体的な一致を測定する精度である。精度は、試料の結果の総数で除した真の陽性値と真の陰性値との合計として計算される（ $(TP + TN) / (TP + TN + FP + FN)$ ）。遺伝子試験の精度を用いて、リスクカットオフ値のどの組が多遺伝子試験における有用な閾値であるかを決定してもよい。

【0043】

感度、特異性、PPV、NPVおよび精度を、それぞれのリスクカットオフ値に関して計算し、下記の表1は、症例102例および対照405例を分析した例に関するこれらの値を示す。完全な範囲のスコアから選択されたカットオフ値を、第一の欄に示す。対応するカットオフ値より高いスコアを有する症例個体の数を第二の欄に示す。第三の欄は、対応するカットオフ値より高いスコアを有する対照個体の数を示す。閾値として対応するカットオフ値のそれぞれを用いる試験の感度を、第四の欄に示す。閾値として対応するカットオフ値のそれぞれを用いる試験の特異性を第五の欄に示す。閾値として対応するリスクカットオフ値のそれぞれを用いる試験のPPVおよびNPVをそれぞれ、第六および第七の欄に示す。最後に、閾値として対応するカットオフ値のそれぞれを用いる試験の精度を第八の欄に示す。

【0044】

【表 1】

リスク カットオフ値	症例数 (102例中)	対照数 (405例中)	感度	特異性	PPV	NPV	精度
80	25	0	0.25	1	1	0.84	0.85
70	51	2	0.50	0.995	0.96	0.89	0.90
65	65	8	0.64	0.98	0.89	0.91	0.91
60	79	34	0.77	0.92	0.70	0.94	0.89
55	93	81	0.91	0.80	0.53	0.97	0.82
50	99	154	0.97	0.62	0.39	0.99	0.69
40	102	318	1	0.21	0.24	1	0.37
30	102	394	1	0.03	0.21	1	0.22
20	102	405	1	0	0.20	1	0.20

10

【0045】

最適な条件において、遺伝子試験は、高いPPV、NPV、および精度に関して非常に感度が高く非常に特異的であり、調べた全ての個体が対象形質を有するまたは有しないと正確に同定される。しかし、典型的な状況において、最適なリスクカットオフ値の選択は、例えば特異性、感度、PPV、NPV、および精度、またはそのサブセットの最善の組み合わせに基づいてもよい。表1に示したように、高いリスクカットオフ値を用いることは、感度およびNPVを低下させながら、試験の特異性およびPPVを増加させる。したがって、疾患の発症に関する個体の素因を決定するための遺伝子試験が高いリスクカットオフ値に基づく場合、疾患を発症するリスクが高いと誤診断される個体は非常に少数であるが、実際に高いリスクを有する個体の大部分が同定されないであろう。一方、低いリスクカットオフ値を用いると、感度およびNPVは増加するが、特異性およびPPVは低下して、それによって高いリスクを有するほとんどまたは全ての個体がそのように同定され、低いリスクを有する有意な数の個体も同様に高いリスクを有すると誤って同定されるであろう。したがって、これらの極端な例のいずれも有用ではないが、その代わりに、感度、特異性、PPVおよびNPVのバランスを、検討中の特定の形質、集団および個体に関して決定してもよいことは明らかである。

20

30

【0046】

閾値の決定は多くの要因に依存する。例えば、疾患に関する臨床知識は典型的に、この決定を行うために必要である。さらに、多遺伝子試験の閾値は、管理機関（例えば、FDA）が規制してもよく、例えば可能性がある治療に関する情報、多遺伝子試験の特徴、または特定の患者の種に応じて医師によって変化してもよい。さらに、閾値は、二分して用いてもよくまたは用いなくてもよい。例えば、個体の治療は、個体のスコアが閾値より大きいか否か（例えば、薬物を投与する）、または閾値より小さいもしくは閾値に等しいか否か（例えば、薬物を投与しない）に応じて変化してもよい。または閾値に近いスコアを有する個体は、閾値から離れたスコアを有する個体とは異なるように治療してもよい。例えば、医師は、臨床知識および個体からの入力のようなさらなる要因に基づいて、閾値よりわずかに下のスコアを有する個体に薬物を投与する決定を下してもよい。さらに、スコアと閾値との比較に関して「大きい」対「小さいまたは等しい」という用語を用いることは単なる便宜上の問題であり、本発明のもう一つの態様において、当業者により明確であるように、「より大きいまたは等しい」対「より小さい」を代わりに用いてもよい。

40

【0047】

一つの局面において、閾値の決定は、疾患の重症度に依存する。例えば、形質が重度の疾患の発症に関連する場合、高リスクの個体を同定することはそれらの個体にとって肝要

50

であることから、非常に高い感度を有することを好むであろう。例えば、治療可能な悪性疾患（インサイチュー癌またはホジキン病）は、早期に発見されなければならない。そのため診断には感度のよい試験を用いなければならない。同様に、偽陰性数が確実に低くなるように、重度の疾患に関しては高いNPVを有する試験が好ましい。偽陽性数は理想的なPPVより低い値のために有意となる可能性があることから、例えば非常に精度が高い「標準」試験を用いて試験陽性/罹患と診断されたそれらの個体の状態を確認するために、さらなる試験を行ってもよい。そのため、容易に利用可能な他の確認診断試験が存在する場合には、より低いPPVを有しても許容される可能性がある。例えば、一般的集団における異型頸部細胞の発生率は約1/1000であり、パパニコロー試験の感度および特異性はそれぞれ、0.70および0.90である。これらの値に基づいて、パパニコロー試験のPPVおよびNPVはそれぞれ、0.00696および0.999となり、このことは、パパニコロー試験陽性の人は、真に異型を有する尤度が非常に小さいが、パパニコロー試験陰性の人はほぼ間違いなく無病であるということの意味している。

10

20

40

50

【0048】

特定の局面において、例えば、偽陽性試験結果に関して非常に望ましくない影響が存在する場合には、高い特異性およびPPVが遺伝子試験にとって好ましい。例えば、個体が危険な治療レジメ（移植手術、化学療法、放射線療法、重篤な有害事象を有する薬物、乳房切除等）を受けるか否かを決定するために試験が用いられる場合、試験によって治療を必要とすると同定された個体が、実際に治療を必要とすることは重要である。例えば、心臓移植技法を行わなければ死亡のリスクが高い個体を同定するために、遺伝子試験を開発してもよい。このように、閾値より高いスコアを有する個体は、新しい心臓を移植しなければ死亡する可能性が高いと同定される。そのような試験は、死亡の確率が高い個体のみが心臓移植のために検討されることから、非常に高いPPV（～1.0）を有することが望ましいであろう。これは、心臓移植を行わなければ死亡する個体の有意な数が治療（より低いNPV）から除外されることを意味し、最適に、絶对的にこれを必要としない如何なる人も心臓移植を受けないことを意味する。

【0049】

遺伝子試験に関して適当な閾値を決定するためのもう一つの要因は、全体として集団における疾患の発生率である。例えば、集団において極めてまれな形質を考慮する。特異性0.95は許容されるほど高いように思われるが、これは高いリスクを有しない個体の5%が、形質を発症する高いリスクを有すると誤って診断されることを意味する。このように、集団における発生率が1/10,000である形質に関して、形質を発症するリスクを有すると正確に同定されるあらゆる個体に関して、個体約500例が、「高リスク」（偽陽性）であると誤って診断されるであろう。したがって、まれな、重度でない形質に関してはより高い特異性を有するカットオフを用いること、および一般的な重度の形質に関してより高い感度を有するカットオフを用いることが最も適しているであろう。さらに、上記のように、PPVおよびNPVは、対象形質の発生率に大きく依存する。例えば、疾患の発生率が低い集団から疾患を発症するリスクを有する人を同定するために用いられる遺伝子試験のPPVは、疾患の発生率が高い集団から疾患を発症するリスクを有する個体を同定するために用いられる同じ遺伝子試験のPPVより低いであろう。同様に、疾患の発生率が低い集団から疾患を発症するリスクを有する個体を同定するために用いられる遺伝子試験のNPVは、疾患の発生率が高い集団から疾患を発症するリスクを有する人を同定するために用いられる同じ遺伝子試験のNPVより高いであろう。そのため、一つの集団において個体を試験するために用いる場合、遺伝子試験は非常に高いPPV（または非常に高いNPV）を有するが、対象形質の発生率が異なる場合には、他の集団において有用ではない可能性があり、したがって対象形質の発生率に応じて異なる集団に関して異なる閾値を選択してもよい。要するに、当業者は、本明細書に示した方法のみならず、例えばFDAのような管理機関との協力作用と共に臨床知識および直感を用いて、所定の形質に関する特定の発生率を有する患者集団に関して、感度、特異性、PPV、NPV、精度等のような一つまたはそれ以上の臨床的に有用なパラメータを得るために閾値を選択することができる。

【0050】

本発明の一つの局面において、関連SNPを用いた多遺伝子試験の閾値は、リスクカットオフ値に関して計算した感度および特異性に基づいて、ROC（受信者動作特性）曲線（Hanley et al.(1982)Radiology 143:29~36；およびBeck et al.(1986)Arch. Pathol. Lab. Med. 110:13~20）を用いて決定される。ROC曲線は、遺伝子試験の感度および特異性のあいだの固有のトレードオフに関連し、表1からのデータを用いて作製したROC曲線を示す図1に示されるように、1マイナス（それぞれのリスクカットオフ値の特異性）の関数として感度をプロットすることによって作製される。このように、それぞれのリスクカットオフ値は、ROC曲線上の「データポイント」に対応する。曲線下面積は、遺伝子試験の信頼性の測定値を提供する。罹患および非罹患個体を完全に区別することができる（感度および特異性がそれぞれ1）遺伝子試験に関して、曲線下面積は1である。罹患および非罹患個体を区別することができない遺伝子試験の場合、曲線下面積は0.5である。一般的に、曲線がプロットの左側および上部境界に近づけば近づくほど、遺伝子試験はより正確であり、曲線がROC空間の45度の角度になれば、試験の精度は低下する。ROC曲線を分析するために一般的に用いられるコンピュータープログラムは、公共に入手可能であり、これにはROCKIT、CORROC2、LABROC4、ROCFIT、CLABROC、ROCPWR、LABMRMC、およびPROPROCが含まれ、これらは全て、Kurt Rossman Laboratories for Radiological Image Researchの以下のウェブサイトからダウンロードしてもよい：www-radiology.uchicago.edu/krl/KRL_ROC/software_index.htm#ROC%20calculations%20Auxiliary%20software。特定の態様において、閾値は、そのデータポイントが、プロットの左上角に最も近いROC曲線の一部（例えば、百分率）において認められるリスクカットオフ値から選択される。例えば、データポイントが図1（矢印AとBのあいだ）に示したプロットの左上の角に最も近いROC曲線の20%から選択される場合、閾値はDおよびEとしてそれぞれ示されるように、リスクカットオフ値55および60に対応するデータポイントから選択されるであろう。他の態様において、閾値は、その感度および特異性がプロットの左上の角に最も近いデータポイントによって表されるリスクカットオフ値であると決定される。図1において、このデータポイント（D）は、リスクカットオフ値55に対応する。なおさらなる態様において、閾値は、プロットの左上の角に最も近いROC曲線上の位置から決定される。図1において、この位置は、Cで示され、感度約0.87および特異性約0.84に対応する。この態様において、曲線上のその位置によって表される感度および特異性に対応するリスクカットオフ値が決定され、そのリスクカットオフ値を、関連SNPを用いる遺伝子試験の閾値として用いる。例えば、位置Cは、データポイントDおよびEのあいだに存在することから、閾値として用いるために最適なリスクカットオフ値は、55と60のあいだでなければならない。最適なリスクカットオフ値を決定するために、感度および特異性は、症例群および対照群のスコアに基づいてその範囲に存在する全てのリスクカットオフ値に関して決定される（表2を参照されたい）。その感度および特異性がそれぞれ、0.87および0.84に近いリスクカットオフ値を選択すると、本実施例ではリスクカットオフ値は56、感度0.88、および特異性0.84である。したがって、関連対立遺伝子を用いる多遺伝子試験に関する閾値として56が選択される。

【0051】

本発明のもう一つの態様において、特定の所望の臨床結果に基づいて、閾値を選択してもよい。例えば、特定の治療を行う個体における有害事象の発生率を低下させるための手段として、患者集団を階層化するために、遺伝子試験を開発してもよい。例えば薬物を、有害事象の発生率が4%であるために限定された使用に関して承認してもよいが、有害事象の発生率が少なくとも50%低下すればより広い使用のために承認されうるであろう。この例において、「症例」は、薬物に反応して有害事象を有するであろう個体であり、「対照」は、薬物に曝露されても有害事象を有しないであろう個体である。個体が有害事象を経験するリスクは、関連する座の組でその遺伝子型に基づいて個体に関するスコアを計算する段階、その後、例えば閾値が症例群および対照群のスコアに基づいて遺伝子試験に関するPPV、NPV、感度、特異性等またはそのいくつかの組み合わせの分析によって決定される、遺伝子試験に関する閾値とスコアを比較する段階によって決定される。例えば、閾値

より高い多くの関連対立遺伝子を有する個体は、有害事象を有するリスクが高いと同定される可能性がある。閾値60を用いると、症例の77%および対照の8%が消失するであろう。有害事象の発生率は4%であることが知られていることから、1000例の患者集団は~40例となり、その約31例（77%）が関連対立遺伝子 >60 を有し、その約9例が関連対立遺伝子 ≤ 60 を有するであろう。同じ患者集団が対照~960例を有し、その約77例（8%）が関連対立遺伝子 >60 を有し、その約883例が関連対立遺伝子 ≤ 60 を有するであろう。関連対立遺伝子 ≤ 60 の個体108例を除外した後、除外されなかった個体892例における有害事象の発生率を計算してもよい： $(9/892) \times 100 = 1\%$ 。除外された個体における有害事象の発生率も同様に計算することができる： $(31/108) \times 100 = 29\%$ 。同じコンピューターによる方法を用いて、リスクカットオフ値59および61も同様に、診断試験に関する閾値として評価した。リスクカットオフ値59によって、治療から除外されなかった個体における有害事象の予測発生率1%が得られたが、対照群におけるより多くの個体が除外され（92例）、このことは、有害事象のリスクを有しないより多くの個体が、このリスクカットオフ値を診断試験の閾値として用いた場合に薬物治療を否定されないであろうということを意味している。リスクカットオフ値61では、治療から除外されなかった個体における有害事象の予測発生率は1.2%となり、これはリスクカットオフ値60の場合より高いが、より少数の対照個体が除外され（69例）、このリスクカットオフ値を診断試験において閾値として用いた場合に、有害事象のリスクを有しないより多くの個体が、薬物治療から利益を受けることができるであろうことを意味している。さらに、医師が治療集団における有害事象のリスクを2%またはそれ未満に維持しながら、治療した対照の数を最大限にしたい場合、閾値69は、対照個体10例を除外するに過ぎず、治療集団に有害事象のリスク2%を提供するであろう。さらに、表2に示すように、有害事象のリスクを2%に低下させるために、治療される患者群からの除外に関して十分な症例を同定するための試験に関して0.53が必要であるのみであろう。したがって、そのような診断試験において閾値としてリスクカットオフ値69を用いることは、特定の治療によって治療した個体集団における有害事象の発生率を減少させ、それによって、そのリスク/利益プロフィールを改善して、治療に含まれるであろう有害事象のリスクを有しない個体の総数を最大限にしながら、その表示を広げさせるであろう。明らかに、試験集団における有害事象の特定のリスクの選択は、そのような診断試験に関する閾値を決定するために重要な要因であり、リスクのそのレベルの決定は、そのような診断薬の承認に関与するであろう任意の管理機関（例えば、FDA）との協力において、医師が決定しなければならない。例えば、有害事象の1%のリスクが望ましい場合、閾値60を選択することができるが、これは試験のPPVを犠牲にして（利益を得るであろうより多くの個体（対照）が除外されるであろう）、NPVを増加させるであろう（それによって試験集団における有害事象の実数が減少する）。除外される患者は、例えば異なる薬物によって異なるように治療することができ、または有害事象の厳密なモニタリングを行って薬物を投与する、もしくは有害事象を相殺するであろうもう一つの治療もしくは物質によって治療されるであろう。

【0052】

【表 2】

リスク カットオフ値	症例数 (102例中)	対照数 (405例中)	感度	特異性	PPV	NPV	精度
69	54	4	0.53	0.99	0.93	0.89	0.90
61	74	29	0.73	0.93	0.72	0.93	0.89
60	79	34	0.77	0.92	0.70	0.94	0.89
59	80	39	0.78	0.90	0.67	0.94	0.88
58	81	44	0.79	0.89	0.65	0.95	0.87
57	86	53	0.84	0.87	0.62	0.96	0.86
56	90	64	0.88	0.84	0.58	0.97	0.85
55	93	81	0.91	0.80	0.53	0.97	0.82

10

【0053】

感度、PPV、NPV、精度、尤度比、およびROC曲線の概念、ならびに診断試験に関して適当な閾値を選択する方法は、広く用いられ、当業者に周知である（例えば、Janssens et al.(2004)Am. J. Hum. Genet. 74:585～588；www.bamc.amedd.army.mil/DCI/articles/dci10972.htm；Baum M.(1995)Lancet 346:436～437；Forrest P.(1990)「Breast Cancer : the decision to screen」；Nuffield Provincial Hospitals Trust；Morrison, A.S.(1985)、「Screening in Chronic Disease」Oxford University Press Inc. USA；www.genome.gov/10002404；med.usd.edu/som/genetics/curriculum/1ITEST7.htm；Bauman A.(1990) Australian Prescriber 13:62～64；Walker et al., (1986)Med. J. Aust. 145:185～187；Gilbert R.(2001)Western J. Med. 174:405～409；Frohna, J.G.(2001)「Fostering the Efficient, Effective Use of Evidence-Based Medicine in the Clinic」2nd edition, University of Michigan；Ragland, R.A.(2000)「Studying a Study and Testing a Test」4th edition, Lippincott Williams & Wilkins；www.cebm.net/likelihood_ratios.asp；およびwww.elsevier.com/gej-ng10/22/71/52/140/article.htmlを参照されたい）

例えば、一つの試験において、肝硬変と肝細胞癌とを区別する血清中の α -フェトプロテインに関する最善の閾値を、ROC曲線下面積、尤度比、感度、特異性、PPV、およびNPVに基づいて評価した（Soresi et al(2003)Anticancer Res. 23(2C):1747～1753）。他の試験において、マンモグラフィー、ソノグラフィー、およびMRマンモグラフィーを比較して、これらの技術の一つ、または二つもしくはそれ以上の組み合わせが、感度、特異性、PPV、NPV、および精度に関する測定手段を用いて、浸潤性の癌を検出するために最善の結果を提供するか否かを決定した（Malur et al.(2001)Breast Cancer Res. 3:55～60）。三つの撮像技術を全て組み合わせることによって、感度0.994、特異性0.953、PPV 0.939、NPV 0.996、および精度0.97を有する最善の結果が得られた。さらにもう一つの試験において、二つの臨床試験のROC曲線下面積を比較して、試験の一つまたは試験の双方の組み合わせが、乳腺病変のクラスを同定する場合に最も正確であるか否かを決定した（Buscombe et al.(2001)J. Nuc. Med. 42(1):3～8）。もう一つの試験において、前立腺癌を検出するための前立腺特異抗原（PSA）試験は、4 ng/ml PSAのカットオフに関して、感度0.86および特異性0.33であるが、2 ng/ml PSAまでカットオフを低下させると感度を0.95まで増加させ、特異性を0.20に低下させることが判明した（Hoffman et al.(2002)BMC Fam. Pract. 3(1):19）。全てのリスクカットオフ値を調べて、そのそれぞれの特異性、感度、PPV、NPV、LR+およびLR-値、ならびに精度（またはそのいくつかのサブセット）を計算して、これらのパラメータのあいだの最適なバランス、またはそのいくつかのサブセットを、閾値の決定において用いてもよい。当業者は、例えば、診断、予後、薬物ゲノム学、医薬品開発、テラノスティックス（theranostics）等に関して患者集団を階層化するため

20

30

40

50

の臨床的に有用な手段を得るために、これらの任意の測定値またはその組み合わせを最適にする閾値を選択してもよい。

【0054】

本発明の特定の態様において、一つより多い閾値を決定して、これを用いて個体が多因子形質を示すリスクを階層化してもよい。そのような一つの態様において、選択された第一の閾値は、感度の最適化に基づいてもよく、それは、高リスクであるが試験によって同定されない（偽陰性）個体数を減少させるであろう。第一の閾値を用いて遺伝子試験において試験「陽性」である個体を、特異性の最適化に基づいてもよい第二の閾値を用いて同じ遺伝子試験に供する。この第二の閾値は、試験陽性であるが真に高リスクではない（偽陽性）個体数を減少させるであろう。二つのそのような閾値を連続的に用いることは、方法の精度を増加させるために役立つであろう。

10

【0055】

個体が多因子形質を示すリスクを分類するために一つより多い閾値が決定され、用いられる本発明のもう一つの態様は、多数の閾値が同じ遺伝子試験において同時に用いられる態様である。そのような試験において、個体のリスクは、個体のスコアがそれより大きい、小さい、またはそれに等しい閾値に基づいて決定される。一つの態様において、閾値少なくとも約2個を用いる、または閾値少なくとも約5個を用いる、または閾値少なくとも約10個を用いる。特定の態様において、所定の多遺伝子試験に関するあらゆる可能性があるスコアを閾値として用いる；他の態様において、特定の範囲のスコアを含んでもよい、またはスコアの範囲中から選択されるスコアが含まれてもよい、可能性があるスコアのサブセットを用いる。例えば、第一の閾値より高いスコアを有する個体が疾患を発症する可能性が非常に高いと分類され、したがって発病を防止するために適当な薬物によって治療されるように、第一の閾値を選択してもよい。第二の閾値より低いスコアを有する個体が、疾患を発症する尤度が非常に低いとして分類され、したがって発症を防止するために治療が行われないうように、第二の閾値を選択してもよい。第一および第二の閾値のあいだのスコアを有するそれらの個体は、疾患を発症する中間の尤度を有すると分類されてもよく、したがって、第一の閾値より高い、または第二の閾値より低いスコアを有する個体とは異なるように治療してもよく、例えば、個体には薬物を投与しなくてもよいが、万一起こっても疾患の発症を検出するためにより厳密にモニターしてもよい。中間のリスクを有する個体の治療は、中間のリスクを有しない（すなわち、「高い」または「低い」リスクがある）個体の治療より、疾患に関する臨床情報、多遺伝子試験、薬物、患者等のような他の情報に大きく依存する可能性がある。

20

30

【0056】

関連座位の組を関連試験によって同定してもよいが、関連座位の必ずしも全てを単一の多遺伝子試験において用いる必要はない。関連座位の組が同定された後、多遺伝子試験において用いられる関連座位の数を調節してもよく、例えばその感度、特異性、相対リスク、尤度比、PPV、NPV、精度、またはその組み合わせに関して試験の値を分析してもよい。例えば、特定の態様において、高い感度と組み合わせた高い相対リスクが好ましい。一つの局面において、本発明の方法を用いて、多遺伝子試験において用いられる関連座位のサブセット（例えば、少なくとも約5、10、15、20、30、または50）を決定してもよい。例えば、症例群と対照群のあいだで最大の対立遺伝子頻度の差を有する関連座位を選択してもよい。いくつかの態様において、少なくとも約8%（0.08）、10%（0.1）、15%（0.15）、または25%（0.25）の対立遺伝子頻度の差を有する座位のみを、多遺伝子試験において用いるために選択する。いくつかの態様において、多遺伝子試験において用いられる関連座位のサブセットは、症例群と対照群からの遺伝子型同定データを用いて得られた多遺伝子試験の特定の特徴を分析することによって決定される。例えば、感度、特異性、相対リスク、尤度比、PPV、NPV、精度、またはその組み合わせを、関連座位の所定のサブセットを用いて仮説上の多遺伝子試験に関して決定してもよい。そのような複数の仮説上の多遺伝子試験をこのように分析してもよく、組み合わせてこれらの特徴の最善の組み合わせを有する多遺伝子試験が得られる関連座位のサブセットを選択してもよい。先に記述したよ

40

50

うに適当な閾値を決定する場合のように、多遺伝子試験の感度、特異性、相対リスク、尤度比、PPV、NPV、精度またはそのサブセットの最善の組み合わせは、例えば、表現型の重症度、表現型の発生率、および集団特異的または患者特異的である他の臨床情報を含む多くの臨床要因に依存する。特定の態様において、多遺伝子試験において用いられる関連座位のサブセットは、関連座位に関する対立遺伝子頻度の差と、得られた多遺伝子試験の特徴との組み合わせに基づいて決定される。このように、本発明の方法を用いて、そのような特徴を測定するためにそのサブセットのみを用いて症例対照試験を行うことなく、関連座位のサブセットを用いて多遺伝子試験の特徴を予測してもよい。

【0057】

本発明のこの局面は重要な実地的な意味を有する。例えば、特定の関連座位が第二のバリデーション関連試験において同じ結果を示さない場合、それらを多遺伝子試験において用いられる関連座位の組から除去してもよく、もう一つの関連試験を行うことなく、「同じ結果を示さない」座位を有しない多遺伝子試験の特徴を決定してもよい。さらに、多数の座位を遺伝子型同定する必要がある多遺伝子試験は、少数の座位を遺伝子型同定する必要がある多遺伝子試験より実施するために高価である。このように、多遺伝子試験において特異的な所望の特徴（例えば、感度、相対リスク等）を維持しながら関連座位の数を減少できることは、そのような試験を行う入手可能性に関して、したがってそのような試験の実際の応用可能性に対して直接の意味を有する。

【0058】

さらに、それぞれの関連SNP対立遺伝子の関与が相加的であって、表現型の効果の程度が類似または異なること、およびそれぞれの関連SNPが互いに独立している（すなわち、独立して分離する）ことを仮定する多遺伝子試験を用いて、世代全体の集団に関して多因子形質を発症するリスクを予測することができる。そのような一つの態様において、症例集団および対照集団における個体におけるSNPの組に関して遺伝子型を決定し、それぞれの集団における各SNPに関して対立遺伝子頻度を決定する。SNPが互いにハーディ-ワインバーク平衡にあると仮定すると、各集団における遺伝子型の頻度は、各集団における対立遺伝子頻度に基づいて決定される。この計算の単純な説明は、一つのみのSNPを遺伝子型同定する例において証明することができる。SNPが小さい対立遺伝子（C）頻度0.3を有する場合、「CC」遺伝子型の頻度は $(0.3)^2$ であり、「GG」遺伝子型の頻度は $(0.7)^2$ であり、「GC」遺伝子型の頻度は $2(0.3)(0.7)$ であろう。したがって、対立遺伝子の一つが、個体が対象表現型形質を発症するであろうリスクを増加させる場合、集団のリスクは、症例および対象集団におけるCC、GGおよびCG遺伝子型の頻度に基づいて決定することができるであろう。そのようなリスクは、直接の集団に当てはまるのみならず、一般的に大きい集団の場合にも当てはまるように、対立遺伝子頻度が世代間で一定である限り、その後の集団にも当てはまるであろう。多遺伝子試験において、複数のSNPを症例および対照集団において遺伝子型同定して、二つの組の遺伝子型を作製する。SNPが互いに独立している場合、複数のSNPに関する対立遺伝子の組を用いて二つの集団に関する遺伝子型頻度の組を計算することができる。それぞれの遺伝子型が、それが含む関連対立遺伝子の数に基づいてそれに関連する特定のリスクを有する場合、集団における個体の全体的なリスクは、遺伝子型の組によって与えられるリスクに基づいて決定することができ、このリスクは、対立遺伝子頻度が相対的に一定である限り、他の集団（例えば、次世代）における個体に関しても同じである。

【0059】

多因子疾患を発症するリスクを有する個体の同定

一つまたはそれ以上の閾値が同定された後、症例または対照群のメンバーではない個体（「試験個体」）を、試験して、個体が対象形質を発症または示すリスクを決定してもよい。本発明の特定の態様において、試験個体は症例および対照群における個体と同じ種である。試験個体は、関連SNP座位のそれぞれにおいて遺伝子型を同定される。SNP座位のそれぞれでの遺伝子型に基づいて、当初の症例および対照群における個体に関してスコアを計算した方法と同じ方法で、試験個体に関してスコアを計算する。本発明の一つの態様に

において、試験個体に関する計算スコアを一つまたはそれ以上の閾値と比較して、個体が疾患を示す可能性があるか否かを決定する。例えば、試験個体が、第一の閾値より大きいスコアを有する場合、試験個体が疾患を発症するまたは示すか否かを考慮してもよく、試験個体のスコアが第二の閾値に等しいまたは小さい場合、試験個体は、疾患を発症するリスクが低いと見なされる可能性がある。第一および第二の閾値は、同じまたは異なる値であってもよい。例えば、第一および第二の閾値の双方に関して55を選択する態様において、55より大きいスコアを有する試験個体は疾患を発症する可能性があるとして診断され、スコア55またはそれ未満を有する試験個体は、疾患を発症する可能性が低いと診断されるであろう。さらに、疾患の発生率ならびに遺伝子試験の感度および特異性に基づいて、試験によって高リスクと同定された人が疾患を有するまたは発症する確率または尤度を計算してもよい（例えば、先に考察したような試験前オッズ）。同様に、試験によってリスクが低いと同定された人が、疾患を有しないおよび疾患を発症しない確率または尤度を計算してもよい。

10

20

30

40

50

【0060】

本発明のもう一つの態様において、個体が疾患を発症するまたは示す尤度をさらに分析するために、試験個体に関して相対リスクを計算する。相対リスクは、どれほど多くの特定の危険因子が明記された転帰のリスクに影響を及ぼすかに関する測定値である。例えば、危険因子に関連した相対リスク2は、その危険因子を有する人が、その危険因子を有しない人より明記された転帰を有するリスクが2倍増加していることを意味する。一つの局面において、疾患の相対リスクは、一般集団における形質（例えば、疾患）のリスクと比較してリスクが倍増している。相対リスクは、疾患に関連するSNPの組でその遺伝子型に基づいて所定のスコアを満たすまたはそれを超える、対照群における個体の百分率に対する症例群における個体の百分率の比を計算することによって計算される。例えば、表1に示すデータを用いて、スコア少なくとも65を有する個体の相対リスクは $(0.64)/(0.02)=32$ であり、このことは、個体が、関連するSNP位置でその対立遺伝子組成に基づいて疾患を発症するリスクが32倍増加していることを意味する。比較すると、スコア少なくとも70の個体の相対リスクは $(0.5)/(0.005)=100$ であり、これは、個体が関連するSNP位置でのその対立遺伝子組成に基づいて疾患を発症するリスクが100倍増加していることを意味する。本発明の一つの局面において、スコアは、関連するSNP座位でその遺伝子型に基づいて試験個体に関して計算され、症例および対照群を分析して、症例個体の何%および対照個体の何%が試験個体と少なくとも同程度に大きいスコアを有するかを決定する。次に、試験個体のスコアと少なくとも同程度に大きいスコアを有する症例個体の百分率を、試験個体のスコアと少なくとも同程度に大きいスコアを有する対照個体の百分率によって除して、試験個体に関する相対的リスクを計算する。

【0061】

先に記述したように、相対リスクは一般的集団において疾患のリスクと比較してリスクの倍増を提供する。したがって、疾患を発症する試験個体のリスクを決定するために、個体の相対リスクを、疾患の発生率に関する臨床情報と組み合わせなければならない。例えば、疾患の発生率が1:100である場合、相対リスク32を有する個体は、疾患を発症する確率32:100または0.32を有する。しかし、発生率1:1,000,000を有する疾患に関して、相対リスク32を有する個体は、疾患を発症する確率32:1,000,000または0.000032を有する。このように、相対リスクがこれらの二つの例において同じであっても、疾患を発症する実際の確率は、これらの二つの疾患に関して非常に異なった。本発明の特定の局面において、試験個体が対象多因子形質を発症するリスクは、一般的集団における多因子形質の発生率を、個体に関して決定した相対リスクに掛けることによって計算される。相対リスクの決定は、広く公知であり、一般的に当業者によって日常的に行われている（Sackett et al. (1991)「Clinical Epidemiology : a basic science for clinical medicine」(second edition) Little Brown, Bostonを参照されたい）。

【0062】

さらに、遺伝子試験のPPVおよびNPVは、試験結果に基づいて個体が疾患を有するまたは

発症するリスクに関する情報を提供することができる。例えば、PPV 0.87およびNPV 0.99の試験を用いて、個体が疾患に関して試験「陽性」である場合、個体が疾患を有するまたは発症する見込みは87%である。同様に、もう一つの個体と同じ試験を用いて疾患に関して試験「陰性」である場合、個体が疾患を有するまたは発症する確率は1%に過ぎない。

【0063】

尤度比は、特定の試験結果が、患者が先に記述したように対象多因子形質を有するまたは有しない尤度をどれほど変化させるかに関する測定手段を提供するために、試験の感度および特異性を利用する。陽性試験結果(LR+)の尤度比(LR)は、(1-特異性)によって除した感度として計算され、陰性試験結果のLR(LR-)は、特異性によって除した(1-感度)として計算される。これらのLR値に試験前の確率を乗じて、試験後のオッズを計算するが、これは疾患の発生率、患者のプール、および特異的患者の危険因子(試験前オッズ)に関する情報、ならびに診断試験(LR)自身に関する情報を組み入れることによって、個体が多因子形質を有するまたは発症する見込みを表す。試験後のオッズを用いて、試験後のオッズを(1+試験後のオッズ)によって除することによって、試験後の確率を計算してもよい。例えば、試験陽性である個体が、発生率1.5%に基づいて試験前オッズ1~66を有して、試験がLR+ 6.6を有する場合、試験後のオッズは0.1で、試験後確率は0.09となり、このことは個体が疾患を有する見込み9%を有することを意味している。同様に、試験陰性である個体が、試験前オッズ1~3を有し、試験のLR-が0.09である場合、試験後のオッズは0.03であり、個体が疾患を有する試験後の確率3%に対応する。このように、多因子形質の尤度比および発生率を用いて、個体が所定の試験結果に基づいて対象多因子形質を有するまたは発症する確率を計算してもよい。

【0064】

予後および診断での利用

予防的測定は、異なる多くの疾患を予防するために成功するが、これらの測定は個体が、疾患の発症前に疾患を発症するリスクを有すると同定されうる場合に限って成功する。多因子疾患の発症は、その発症に影響を及ぼす複雑な組の要因のために、予測することが特に困難である。そのため、個体は自身が多因子疾患を発症するリスクを有することを知らず、これでは疾患の予防には遅すぎる。示された方法は、それらの患者のケアに関して医師が医学的決定を行うための貴重なツールとして役立つ可能性があることは当業者には明確であろう。リスクの決定は、医学的介入が正当化されているか否か、およびどの介入が所定の個体にとって最も適当であるかを決定するために用いられる個体の臨床分析の重要な局面である(Bucher et al.(1994)BMJ 309(6957):761~764; Forrow et al.(1992)Am J. Med 92(2)121~124)。

【0065】

特定の態様において、本発明は、疾患を発症するリスクを有する個体を同定して(予後)、それによって疾患の発症を予防または遅らせるための手段の実行を可能にするための方法を提供する。一つの態様において、個体が所定の疾患を発症するリスクは、疾患関連SNPの組で個々の遺伝子型に基づいてスコアを少なくとも一つの閾値と比較することによって決定してもよい。個々のスコアが閾値を超える場合、予防的手段(例えば、放射線または薬物治療)の開始が正当化される可能性がある。もう一つの態様において、個体が疾患を発症するリスクは、個体の相対リスクを計算すること、および相対リスクに疾患の発生率を掛けることによって決定してもよい。もう一つの態様において、遺伝子試験の感度、特異性、PPV、NPV、および/または精度を用いて、個体が疾患を発症するリスクを計算する。さらにもう一つの態様において、試験に関するLRを用いて、個体が疾患を発症する試験後オッズ/確率を計算する。もう一つの態様において、上記の方法の組み合わせを用いて、個体が疾患を有するまたは発症するリスクを決定する。この情報を医師が用いて、個体に関する適当な治療レジメをよりよく決定してもよい。しばしば、この情報は、疾患、患者、または患者が由来する集団に関する臨床情報と組み合わせて用いられる。いくつかの局面において、本明細書に示した方法は、疾患に対して抵抗性である個体を同定するために用いてもよい。例えば、疾患(例えば、乳癌)の家族歴を有するいくつかの個体は

、疾患を決して発症しない。この知識によって、これらの個体が問題の疾患を発症するリスクをよりよく評価することができ、高リスクではない人に心の平穏を与え、場合によっては激しい予防的治療（例えば、選択的乳房切除）を行わなくてすむであろう。本明細書に示した方法はまた、有害な非疾患状態を発症するリスクが増加した個体を同定して、それによって疾患の発症を予防するためにライフスタイルの変化の動機付けを行うために用いてもよい。例えば、高血圧症に関連したSNPの組を含む多遺伝子試験は、高リスクであることが判明した人に、運動して健康な食事をとるように強い刺激を提供することができるであろう。

【0066】

いくつかの疾患は、患者における明らかな身体症状のみに基づいて診断することが難しい。これらの疾患の診断はしばしば、そのような疾患が異なる個体において多様に発現するために、および/またはその症状が多数の無関係な疾患の症状と類似である可能性があることから混同されている。本発明のさらなる局面において、そのような疾患に関連するSNPの組は、疾患を示す可能性がある表現型を示す個体の診断を助けるために用いてもよい。このように、関連SNPの組に関して個体を遺伝子型同定する段階、および個体が疾患を示すリスクを決定する段階は、身体症状によって示唆される診断に対して支持するまたは異議を唱えることができるであろう。診断が支持される場合、医師はこの情報を用いて、疾患に対する治療レジメの開始のような、個体に関する治療の決定を行うことができる。例えば、セリアック病は、小腸に損傷を与え、食物からの栄養の吸収を妨害する消化管の自己免疫障害である。詳しく述べると、セリアック病は、コムギ、ライムギおよびオオムギにおいて認められるタンパク質であるグルテンに反応して小腸において炎症反応を引き起こし、セリアック病の唯一の治療は、無グルテン食である。個体が異なれば異なる症状を示すことから、セリアック病を診断することは難しい。例えば、主に膨隆腹または下痢のような消化管症状を示す人もあれば、被刺激性または抑うつのみを示す人もある。さらに、病態は、その症状が刺激性腸症候群、クローン病、潰瘍性大腸炎、憩室症、腸感染症、慢性疲労症候群、および抑うつを含む他の多くの病態と類似であることから容易に誤診断されうる。本明細書に示した方法は、セリアック病に関連した遺伝子座の組を同定するために用いてもよく、これらの座位を用いてセリアック病を示す症状を示す個体をスクリーニングしてもよい。その遺伝子組成に基づいてセリアック病を発症するリスクが高いことが判明した個体は、セリアック病を有すると診断される可能性があり、無グルテン食を与えてもよい。

【0067】

他の態様において、本明細書に示した方法を用いて、予防的治療が個体における例えば疾患の発症を確実に防止するか否かを決定するために役立つ可能性がある。例えば、乳癌の予防のために承認された治療が存在し、これらは家族歴、初経期間の開始、子供の数等のような既往に関する臨床情報に依存する。これらの要因は、試験前オッズを計算するために有用であるが、女性が乳癌を発症するか否かをかろうじて予測するに過ぎない。遺伝子試験を試験前オッズと組み合わせて用いると、乳癌を発症するリスクを同定および定量するためのはるかに正確な方法を提供することによって、個体を予防的に（例えば、タモキシフェンによって）治療するか否かの決定に関してかなり優れた手段を提供するであろう。

【0068】

本発明の一つの局面において、生体試料における関連SNPの組の存在を検出するように設計されたプローブを含む核酸アレイを含む予後または診断アッセイが提供される。核酸を、試験個体の生体試料から単離して、核酸アレイ上のプローブとハイブリダイズさせる。プローブの強度を分析して、関連するSNP位置のそれぞれで試験個体に関する遺伝子型を提供する。遺伝子型を用いて試験個体に関するスコアを計算して、本明細書に示した方法に従って、個体が疾患を発症するリスクを決定する。

【0069】

関連するSNPの組をさらに、疾患の表現型の発症に関与するゲノムの領域を同定するた

10

20

30

40

50

めに用いてもよい。これらのSNPは、疾患の発現に直接関与してもよく、またはそれらは直接関与する座位と連鎖不平衡であってもよい。例えば、疾患関連SNPは、疾患関連タンパク質の発現もしくは機能に直接影響を及ぼす可能性がある、またはタンパク質の発現もしくは機能に影響を及ぼすもう一つの座位と連鎖不平衡であってもよい。タンパク質の発現または機能に対する直接作用の例には、タンパク質のポリペプチド配列を変化させる多型、およびそれによってタンパク質発現の増加または減少が起こる調節領域（すなわち、プロモーター、エンハンサー等）において起こる多型、が含まれるがそれらに限定されるわけではない。特定の態様において、関連SNPの組を含むゲノム領域を分析して、疾患の生物学的基礎に直接関与する遺伝子（「同定された遺伝子」）を同定する。

【0070】

遺伝子のコード領域に存在する関連SNPを用いて、疾患の診断マーカーとして用いるために、生物学的標本における関連対立遺伝子の発現を検出または定量してもよい。例えば、RNAまたはmRNAをコードする遺伝子が関連対立遺伝子を含むか否かを決定するために、関連SNPを含む核酸をオリゴヌクレオチドプローブとして用いて、試験される生物または組織もしくは臓器のようなその一部におけるRNAまたはmRNAレベルをモニターしてもよい。一つの局面において、生体試料における関連対立遺伝子を検出するために用いられるオリゴヌクレオチドプローブを含む診断または予後キットを提供する。同様に、関連対立遺伝子がコードされるタンパク質のポリペプチド配列の変化を引き起こす場合、遺伝子の対立遺伝子の構成を、免疫学的方法（例えば、ウェスタンブロット、放射免疫沈殿等）または遺伝子産物に関連した活性を測定する活性に基づくアッセイのような、任意の通例の技術を用いてタンパク質レベルでアッセイしてもよい。一つの局面において、生体試料において関連対立遺伝子によってコードされるポリペプチドを検出するためのアッセイを含む診断または予後キットが提供される。細胞が特定のヌクレオチドまたはポリペプチド配列の存在に関してプロービングされる方法は文献において十分に確立されており、本明細書においてさらなる詳述を必要としないが、例えばSambrook et al, 「Molecular Cloning : A Laboratory Manual」 (Cold Spring Harbor Laboratory, New York) (2001) を参照されたい。

【0071】

治療

関連SNPの組は、疾患の予防のための治療薬を開発するために有用となる可能性がある。一つの局面において、同定された遺伝子を遺伝子治療のために用いてもよい。例えば、同定された遺伝子が疾患を示す個体においてダウンレギュレートされることが判明した場合、遺伝子のアップレギュレーションは、試験個体における疾患の発病を予防するために有効な戦略となりうるであろう。同定された遺伝子のアップレギュレーションは、疾患に関連しない遺伝子の対立遺伝子を発現ベクターに組み入れる段階、およびベクターを生物にさらに導入して、それによって生物における遺伝子の発現をアップレギュレートさせる段階によって行ってもよい。そのようなベクターは一般的に、レシピエントゲノムにおける核酸配列の挿入を提供するために、プロモーター配列の付近に存在する簡便な制限部位を有する。転写開始領域、標的遺伝子またはその断片、および転写終了領域を含む転写カセットを調製してもよい。転写カセットを、細胞において一過性または安定に維持される多様なベクター、例えばプラスミド、レトロウイルス、例えばレンチウイルス、アデノウイルス等に導入してもよい。遺伝子またはタンパク質産物を、ウイルス感染、マイクロインジェクション、または小胞の融合を含む任意の数の経路によって組織または宿主細胞に直接導入してもよい。ジェット注入も同様に、Furth et al., Anal.Biochem, 205:365~68(1992)によって記述されるように、筋肉内投与のために用いてもよい。または、DNAを金微粒子上にコーティングして、これを、文献（例えば、Tang, et al., Nature 356:152~54(1992)を参照されたい）に記述されるように粒子衝突装置または「遺伝子銃」によって皮内送達してもよい。

【0072】

同定された遺伝子によってコードされるタンパク質は、疾患に対する素因に関連したタ

10

20

30

40

50

ンパク質の配列のアミノ酸変化が存在する場合、抗体療法のための標的となる可能性がある。例えば、関連対立遺伝子が疾患を引き起こす原因であるタンパク質変種をコードする場合、疾患の発症を阻害するための手段として、疾患関連タンパク質変種に対して特異的な抗体を患者に投与してもよい。特定の態様において、疾患の発症を予防するために、それぞれが異なる疾患関連タンパク質に対して特異的な抗体の組み合わせを患者に投与してもよい。

【0073】

アンチセンス分子を用いて、細胞において同定された遺伝子の関連対立遺伝子の発現をダウンレギュレートしてもよい。アンチセンス分子は遺伝子の対立遺伝子によってコードされるmRNAと二本鎖を形成して、それによってその発現をダウンレギュレートして対応するタンパク質の翻訳を遮断する。例えば、関連対立遺伝子によってコードされるmRNAの配列に基づいて、アンチセンス試薬を開発してもよい。次に、このアンチセンス試薬をヘテロ接合患者（関連対立遺伝子1個と、疾患に関連しない対立遺伝子1個とを有する）に投与して、関連対立遺伝子の発現を減少させ、それによって非関連対立遺伝子の発現を優勢にしてもよい。アンチセンス試薬はアンチセンスオリゴヌクレオチド、特に化学改変を有する合成アンチセンスオリゴヌクレオチド、またはRNAのようなアンチセンス分子を発現する核酸構築物であってもよい。多数の異なる配列を含んでもよいアンチセンス分子の組み合わせを投与してもよい。

10

【0074】

アンチセンス阻害剤の代わりとして、触媒的核酸化合物、例えば、リボザイム、アンチセンス結合体等を用いて関連対立遺伝子の発現を阻害してもよい。リボザイムをインビトロで合成して、患者に投与してもよく、または発現ベクター上でコードさせてもよく、そこからリボザイムを標的細胞において合成させる（例えば、国際特許出願国際公開公報第9523225号およびBeigelman, et al., Nucl. Acids. Res. 23:4434~42(1995)を参照されたい）。触媒活性を有するオリゴヌクレオチドの例は国際公開公報第9506764号に記述されている。アンチセンスオリゴヌクレオチドと、mRNA加水分解を媒介することができる金属錯体、例えばテルピリジルCu(II)との結合体は、Bashkin et al., Appl. Biochem. Biotechnol. 54:43~56(1995)に記述されている。

20

【0075】

同定された遺伝子によってコードされる発現されたタンパク質は、そのタンパク質産物に結合する、調節する、またはその作用を模倣するリガンドまたは基質を同定するために、およびそれによって例えば罹患細胞におけるタンパク質機能の置換もしくは増強を提供する治療物質を提供するため、またはタンパク質機能を調節もしくは無効にする物質を同定するための、薬物スクリーニングアッセイにおいて用いてもよい。標識されたインビトロタンパク質-タンパク質結合アッセイ、タンパク質-DNA結合アッセイ、電気泳動移動度シフトアッセイ、タンパク質結合に関するイムノアッセイ等を含む、多様なアッセイをこの目的のために用いてもよい。本明細書において用いられる「物質」という用語は、同定された遺伝子または遺伝子産物の生理的機能を直接または間接的に変化させる、模倣する、または隠すことができる任意の分子、例えばタンパク質または低分子を記述する。一般的に、様々な濃度に対する異なる反応を得るために、複数のアッセイを物質の異なる濃度によって平行に行う。典型的に、これらの濃度の一つ、例えばゼロ濃度または検出レベル以下を陰性対照として用いる。同様に、精製タンパク質変種の全てまたは断片を、三次元結晶構造の決定のために用いてもよく、これをタンパク質またはその一部の生物学的機能を決定するため、分子間相互作用の形成、膜融合等のために用いることができる。

30

40

【0076】

候補物質は、多数の化学クラスを含むが、典型的にそれらは有機分子または錯体、好ましくは分子量が50ダルトンより大きく約2,500ダルトン未満である低分子有機化合物である。候補物質は、タンパク質との構造的相互作用、特に水素結合によって必要な感応基を含み、典型的にこれには少なくともアミン、カルボニル、ヒドロキシル、およびカルボキシル基が含まれ、しばしば官能化学基少なくとも二つを含む。候補物質はしばしば、一つ

50

またはそれ以上の上記の官能基によって置換された環状炭素または複素環構造および/または芳香族または多芳香族構造を含む。候補物質はまた、以下が含まれるがそれらに限定されるわけではない生体分子において認められる：ペプチド、糖質、脂肪酸、ステロイド、プリン、ピリミジン、その誘導体、構造類似体、または組み合わせ。

【0077】

候補物質は、合成または天然化合物のライブラリを含む多様な起源から得られる。例えば、ランダムオリゴヌクレオチドおよびオリゴペプチドの発現を含む、多様な有機化合物および生体分子のランダムおよび定方向合成に関して多数の手段が利用可能である。または、細菌、真菌、植物および動物抽出物の形での天然化合物のライブラリが利用可能である、または容易に産生される。さらに、天然または合成によって産生されたライブラリおよび化合物は、通常の化学、物理、および生化学手段によって容易に改変されて、コンビナトリアルライブラリを産生するために用いてもよい。構造類似体を生成するために、公知の薬理的物質に、アシル化、アルキル化、エステル化、アミド化等のような定方向またはランダム化学改変を行ってもよい。

10

【0078】

スクリーニングアッセイが結合アッセイである場合、一つまたはそれ以上の分子を、検出可能なシグナルを直接または間接的に提供することができる標識に結合させてもよい。様々な標識には、放射性同位元素、蛍光体、化学発光体、酵素、特異的結合分子、粒子、例えば磁気粒子等が含まれる。特異的結合分子には、ビオチンおよびスト렙トアビジン、ジゴキシンおよびアンチジゴキシン等のような対が含まれる。特異的結合メンバーに関して、相補的メンバーは、通常、公知の技法に従って検出を提供する分子によって標識されるであろう。他の多様な試薬がスクリーニングアッセイに含まれてもよい。これらには、最適なタンパク質-タンパク質結合を促進する、および/または非特異的もしくはバックグラウンド相互作用を減少させるために用いられる、塩、中性タンパク質、例えばアルブミン、洗浄剤等が含まれる。プロテアーゼ阻害剤、ヌクレアーゼ阻害剤、抗菌物質等のようなアッセイの効率を改善する試薬を用いてもよい。

20

【0079】

物質は、任意および全ての溶媒、分散媒体、コーティング、抗酸化剤、等張剤および吸収遅延剤等を含む、薬学的に許容される担体または希釈剤と配合してもよい。物質は、乳糖、マンニトール、コーンスターチ、またはジャガイモデンプンのような通常の添加剤；結晶セルロース、セルロース誘導体、アカシア、コーンスターチ、またはゼラチンのような結合剤；コーンスターチ、ジャガイモデンプン、またはカルボキシメチルセルロースナトリウムのような崩壊剤；タルクまたはステアリン酸マグネシウムのような潤滑剤；ならびに望ましければ緩衝剤、湿潤剤、保存剤、および着色料と配合してもよい。薬学的に許容される物質に関してそのような媒体および物質を用いることは、当技術分野で周知であり、公共に容易に入手可能である。その上、pH調節および緩衝剤、等張性調節剤、安定化剤、湿潤剤等のような薬学的に許容される補助物質も公共に容易に入手可能である。如何なる通常の媒体または物質も活性成分と不適合である場合を除き、本明細書において記述される治療的組成物および方法におけるその使用が企図される。補助活性成分はまた、組成物に組み入れることができる。

30

40

【0080】

以下の方法および賦形剤は単なる例であって、決して制限的ではない。本発明の同定された方法を、治療的投与のための多様な製剤に組み入れることができる。より詳しく述べると、先に考察した適当な薬学的に許容される担体または希釈剤を配合することによって複合体を薬学的組成物に調製することができ、錠剤、カプセル剤、粉剤、顆粒剤、軟膏、溶液、ゲル、ミクロスフェア、およびエアロゾルのような固体、半固体、液体、またはガス様剤形に調製してもよい。さらに、物質は、植物油または他の類似の油、合成脂肪酸グリセリド、高等脂肪酸のエステルまたはプロピレングリコールのような水性または非水性溶媒に；および望ましければ溶解剤、等張剤、懸濁剤、乳化剤、安定化剤、および保存剤のような通常の添加剤と共にそれらを溶解、懸濁、または乳化させることによって注射用

50

製剤に調製してもよい。さらに、物質は、吸入によって投与されるエアロゾル製剤において利用してもよい。本明細書に記載の方法によって同定された物質は、ジクロロジフルオロメタン、プロパン、窒素等のような加圧された許容される噴射剤に調製することができる。または物質を、乳化基剤または水溶性基剤のような多様な基剤と混合することによって直腸投与のための坐剤に調製してもよく、これらには体内で融解するが、室温では固体であるココアバター、カーボワックスおよびポリエチレングリコールのような媒体が含まれる。

【0081】

徐放製剤のためのインプラントは当技術分野で周知である。インプラントは、生体分解性または非生体分解性のポリマーと共にミクロスフェア、スラブ等として調製される。例えば、乳酸および/またはグリコール酸のポリマーは、宿主によって認容性が良好な侵食性のポリマーを形成する。同定された物質を含むインプラントは、活性物質の局所濃度が体の残りと比較して増加するように、作用部位の近位に配置してもよい。単位投与剤形、例えば小さじ1杯、大さじ1杯、ゲルカプセル、錠剤、または坐剤が本発明の組成物の既定量を含む、シロップ剤、エリキシル剤、および懸濁剤のような、経口または直腸内投与のための単位投与剤形を提供してもよい。同様に、注射または静脈内投与のための単位投与剤形は、滅菌水、通常の生理食塩液、またはもう一つの薬学的に許容される担体における溶液としての組成物において本発明の化合物を含んでもよい。新規単位投与剤形に関する明細書は、用いる特定の化合物および得られる作用、ならびに宿主におけるそれぞれの活性成分に関連した薬力学に依存する。

【0082】

物質の投与は様々な方法で行うことができる。製剤は、経口、吸入によって投与してもよく、または注射、例えば血管内、腫瘍内、皮下、腹腔内、筋肉内等の注射を行ってもよい。物質は局所、全身投与であってもよく、または活性用量を埋め込み部位で保持するように作用するインプラントを用いることによって局所投与であってもよい。治療的製剤の用量は、副作用を最小限にしながら疾患またはその症状に対処するために十分であるように、利用する特定の物質および製剤、疾患の性質、投与回数、投与方法、宿主からの物質の排泄等に応じて変化するであろう。いくつかの場合において、経口投与は、静脈内投与する場合とは異なる用量を必要とするであろう。化合物は適した期間のあいだに、疾患の進行が実質的に停止するように、有効量で投与されるであろう。初回用量はより高く、その後より低い維持用量であってもよい。用量は、有効な用量レベルを維持するために、1回投与、毎週1回、2週間に1回投与してもよく、または低用量に分割して、毎日、1週間に2回等投与してもよい。治療は、例えば心室細動後のように短期間であってもよく、または例えば心室細動のさらなる事例を予防するために長期間であってもよい。組成物はインビボで用いるために医師の指針に従って得て用いられるであろうと企図される。

【0083】

薬物ゲノム学

他の態様において、本発明の方法によって同定された関連SNPの組を薬学ゲノム学および医薬品開発に用いる。共通の多因子疾患に関して利用できる治療選択肢はかなりの数であることから、どの群の治療選択肢が所定の患者にとって最も有効であるかを決定することはしばしば難しい。典型的に、いくつかの治療選択肢を試みて初めて安全かつ有効である治療を発見するであろう。そのあいだに患者は疾患の作用を受け続け、おそらく調べた治療選択肢の一つまたはそれ以上に反応して有害事象を経験するであろう。本明細書に示した方法は、治療レジメの開始前に患者集団を階層化するために有用である。薬物または他の医学治療に対する患者の反応に関連した多型座位が同定される。反応は有害事象であってもよく、または治療の有効性に関連してもよい。関連座位を用いて患者集団をスクリーニングして、患者の関連座位に関連する遺伝子プロフィールを作製して、これはどの個体が薬物または医学治療を受けるべきであって、どの個体がそうでないかを医師が決定するために役立つであろう。例えば、有害事象を示す素因を有する個体および薬物に対して有効な反応を有する可能性が低い個体を、その薬物による治療から除外してもよく、その

代わりに代わりの手段（異なる薬物または他の医学治療）によって治療してもよい。

【0084】

そのような一つの態様において、個体を、特定の薬物治療に対する有害反応の公知のリスクを付与する疾患に関連したSNPの組に関してスクリーニングする。疾患を発症するリスクが高い個体を治療レジメから除外する。例えば、LQTS（QT延長症候群）を有する個体は、抗不整脈薬を投与した場合に心室細動のリスクが高い。そのような薬物を投与する前に、LQTSに関連した座位の組に関して患者集団をスクリーニングすること、およびLQTSを発症するリスクが高いそれらの個体を除外することが有用であろう。疾患に関連したSNPの組は、関連試験を行うことによって決定され、疾患を発症する患者のリスクは上記のように行われる。疾患を発症する高リスクは、抗不整脈薬に反応した有害事象の危険因子であると見なされ、この情報を医師が用いて、個体に関する適当な治療選択肢を決定してもよい。例えば、個体が疾患を発症するリスクが高い場合、薬物の投与を行わなくてもよい。個体が疾患を発症するリスクが低い場合、薬物の投与は実行可能な治療選択肢である可能性がある。

【0085】

本発明のもう一つの態様において、薬物治療レジメの有効性は薬物の有効性に関連したSNPの組で個体の遺伝子型に基づいて個体に関して予測される。この情報は、薬物が個体に関する有効な治療である確率を決定するために、または他の薬物もしくは治療選択肢をその代わりに検討すべきであるか否かを決定するために用いられる。例えば、薬物に対して有効な反応を示さない（「非反応者」）個体の症例群と、有効な反応を示す個体（「反応者」）の対照群とを用いて、関連試験を行ってもよい。症例および対照群のメンバーを複数のSNP位置で遺伝子型を同定して、相対的対立遺伝子頻度をそれぞれのSNPに関して計算し、有効な反応に関連したSNPの組が、症例群および対照群のあいだで有意差を示す対立遺伝子頻度の差を有するSNPであると同定される。関連するSNPでその遺伝子型に基づいて症例および対照群のそれぞれのメンバーに関してスコアを計算して、これらのスコアを用いて、薬物に対する有効な反応を有しない個体のリスクを予想する遺伝子試験の一つまたはそれ以上の適当な閾値を決定する。適当な閾値の決定にはまた、以下の一つまたはそれ以上が含まれてもよい：薬物に関する臨床知識、治療される適応、および患者集団、ならびに遺伝子試験の感度、特異性、PPV、NPV、精度、LR+およびLR-の計算。薬物を投与するための候補である個体を、関連するSNP位置のそれぞれで遺伝子型を同定して、関連するSNPの組でその遺伝子型に基づいて個体に関してスコアを計算する。個体が閾値より大きいスコアを有する場合、個体は、非反応者である可能性があるとして分類され、代わりの治療が検討されるであろう。個体が閾値と等しいまたはそれより低いスコアを有する場合、個体は、反応者である可能性があるとして分類され、薬物の投与が推奨されるであろう。もう一つの態様において、個体が非反応者であるリスクは、個体に関する相対的リスクを計算して、相対リスクに薬物の公知の有効性に基づく非反応者の発生率を掛けることによって決定してもよい。もう一つの態様において、個体が反応者である尤度は、多遺伝子試験の精度、LR+、LR-、PPV、および/またはNPVを用いて計算される。この情報は、個体の適当な治療を決定するために医師が用いることができる。

【0086】

関連する態様において、医師が患者の治療をよりよく個別化することができるように、診断を治療領域に関して開発してもよい。単一の薬物に集中するのではなく、治療領域診断は、患者が単治療領域に関連する一連の薬物の反応者である尤度に関する情報を提供するであろう。例えば、SSRI（選択的セロトニン再取り込み阻害剤）、TCA（三環性抗うつ剤）、MAOI（モノアミンオキシダーゼ阻害剤）、およびトリアゾロピリジンを含む、抑うつを治療するために市販されている多数の薬物がある。これらのタイプの薬物のそれぞれの有効性に関連した多型座位を同定するために関連試験を行ってもよく、次にそれらの座位を用いて、どのクラスの薬物が所定の個体に関して最も有効であるかを決定するために患者集団をスクリーニングすることができるであろう。それぞれの薬物に関して、症例群は、薬物に対して有効な反応を示す抑うつを有する個体を含み、対照群は、薬物に対して

有効な反応を示さない個体を含む。関連SNPは、対照における場合より症例において有意に異なる対立遺伝子頻度を有するSNPであると同定される。それぞれのクラスの薬物に関して、有効な反応を有する見込みが高い（例えば、 $>80\%$ 、 $>90\%$ 、 $>95\%$ 、または $>98\%$ ）個体を同定する閾値を決定する。抗うつ剤治療を必要とする個体を、それぞれの薬物タイプに関連するSNPに関してスクリーニングして、医師は個体の遺伝子型情報およびそれぞれのクラスの薬物に関して決定された閾値に基づいて個体に関して適当な治療選択を決定する。

【0087】

さらに関連する態様において、薬物の有効性に関連したSNPを用いて可能性がある非反応者を治療から除外するために患者集団を階層化することによって薬物の有効性を改善してもよい。一つの例において、薬物に曝露された患者の $\sim 32\%$ が反応者であると分類される。反応者の症例群と非反応者の対照群に関して関連試験を行い、SNP 25個が反応者の表現型に関連することが判明した。症例および対照に関して計算したスコアに基づいて、反応者の 81% および非反応者の 40% がスコア >19 を有することが判明した。したがって、薬物を投与する前に患者集団を階層化するために、閾値として19を用いることによって、薬物の全体的な有効性は $\sim 32\%$ から $\sim 50\%$ へと改善する。それを行うあいだに、薬物に曝露された非反応者の数は実質的に減少し、除外された人を近いうちにもう一つの治療によって治療してもよい。この程度の有効性の変化は新薬の承認を得るために役立ちうる、または既に承認された薬物に関してより広い使用を促進しうるであろう。

【0088】

なおもう一つの態様において、本明細書に示した方法を用いて、どの商標の薬物を用いるべきか、またはより安価な後発品をその代わりに用いてもよいか否かを評価してもよい。例えば、関連試験は、一般的代用薬に対する臨床反応陽性に関連する遺伝子座位を同定するために行われるであろう。治療を必要とする患者を、これらの関連座位で遺伝子型を同定して、スコアを計算する。次に、個々のスコアを用いて個体における後発品の有効性を予測して、医師はこの情報を用いて個体に関する治療の決定を行うであろう。開示された方法のこの応用はまた、医療費償還の決定のためにも用いることができるであろう。例えば、後発品が個体Aにおいて有効でない可能性があることが判明した場合、商標医薬品をAに投与して、商標医薬品の費用をAに償還することができるであろう；個体Bが後発品に対して有効な反応を有する可能性がある場合、個体Bにはより高価な商標医薬品を投与せず、後発品の費用のみが償還されるであろう。

【0089】

本発明のもう一つの態様において、薬物の投与に反応して個体が有害事象を経験するリスクは、薬物に関連した有害事象の発生率に関連したSNPの組での個体の遺伝子型に基づいて決定される。個体が、治療レジメに反応して有害事象を経験するリスクが高いことが判明した場合、治療レジメを行わずに、他の治療選択肢を検討してもよい。例えば、薬物に反応して有害事象を示した個体の症例群と、有害事象を示さなかった個体の対照群とを用いて、関連試験を行ってもよい。症例群および対照群のメンバーを、複数のSNP位置で遺伝子型を同定して、それぞれのSNPに関して相対的対立遺伝子頻度を計算して、有害事象に関連するSNPを、症例群と対照群のあいだで有意差を示す対立遺伝子頻度の差を有するSNPであると同定する。関連SNPにおけるその遺伝子型に基づいて症例群および対照群のそれぞれのメンバーに関してスコアを計算して、これらのスコアを用いて、適当なレベルの感度、特異性、PPV、NPV、LR+、LR-および/または精度によって、個体が薬物に反応して有害事象を経験するリスクを予測する多遺伝子試験に関して一つまたはそれ以上の適当な閾値を決定する。先に考察したように、閾値の選択は、有害事象の重症度、治療される疾患または障害、および治療される個体の既往のような、臨床要因に基づいてもよい。例えば、有害事象が死亡である場合、薬物を投与した場合に死亡の高い確率を有する個体を同定するためには、高い感度が必須である。薬物を投与する前に、個体を関連するSNP位置のそれぞれで遺伝子型を同定して、関連するSNPの組でその遺伝子型に基づいて個体に関してスコアを計算する。例えば、個体が、症例および対照群のスコアから決定した閾値

より大きいスコアを有する場合、個体は、薬物を投与した場合に有害事象を経験する可能性があるとは分類され、薬物の使用は行われまいであろう。個体が、閾値と等しいまたはそれより低いスコアを有する場合、個体は有害事象を有しない可能性があるとは分類され、薬物の投与が推奨される。個体が一つの閾値より低いまたは等しいが、もう一つの閾値より大きいスコアを有する場合、個体は、有害事象を経験する中間の尤度を有すると分類され、代わりの薬物治療を用いてもよく、または厳密なモニタリングを行った場合に限って薬物が投与され、または有害事象を相殺するためにもう一つの治療と併用して投与してもよい。有害事象を経験するリスクが中間である個体の最善の治療レジメの決定は、非常に高いまたは低いリスクを有する個体に関して最善の治療レジメを決定する場合より大きく他の情報（例えば、臨床データ、FDAまたは患者の入力等）に依存する可能性がある。もう一つの態様において、個体が有害事象を経験するリスクは、個体に関する相対リスクを計算して、有害事象を経験する個体の公知の発生率を相対的リスクに掛けることによって決定してもよい。次に、この情報を医師が用いて、個体に関する適当な治療を決定することができる。薬物の投与に反応した有害事象には、アレルギー反応、心不整脈、卒中、気管支痙攣、消化管障害、失神、性的不能、発疹、発熱、筋肉痛、頭痛、悪心、先天性欠損、ホットフラッシュ、気分の変化、眩暈、激昂、嘔吐、睡眠障害、嗜眠、不眠、薬物中毒、および死亡が含まれるがそれらに限定されるわけではない。

10

20

30

40

50

【0090】

関連する態様において、薬物の安全性に関連したSNPを用いて、薬物の投与に反応して有害事象を示す可能性がある個体を治療から除外するために、患者集団を階層化することによって、薬物の安全性を改善してもよい。一つの例において、新薬が、優れた有効性、認容性、および簡便性を有することが判明したが、薬物によって治療した個体の4%が重度の有害事象を経験し、有害事象のこの発生によって、薬物の使用は、例えば他の治療が失敗した個体に限定された。しかし、管理機関は、有害事象の発生率が少なくとも50%低下すれば、薬物をより広い使用に関して承認することができるであろうと明記した。このことは有害事象を経験する可能性がある個体を治療前に同定することができれば達成することができ、有害事象を経験した個体の症例群と経験しなかった個体の対照群とに関して関連試験を行って、有害事象を経験したSNP 20個の組を同定する。関連試験からの結果を表3に示し、リスクカットオフ値を第一の欄に示し、対応するリスクカットオフ値より大きいスコアを有する症例の%を第二の欄に示し、対応するリスクカットオフ値より大きい対照の%を第三の欄に示し、相対リスクを第四の欄に示し、%感度を第五の欄に示し、%特異性を第六の欄に示し、PPV（%として）を第七の欄に示し、およびNPV（%として）を第八の欄に示す。

【0091】

【表3】

リスク カットオフ値	%症例	%対照	相対リスク	感度	特異性	PPV	NPV
20	40.0%	2.8%	14.2	40.0%	97.2%	37.3%	97.5%
19	51.6%	5.6%	9.2	51.6%	94.4%	27.7%	97.9%
18	58.0%	9.9%	5.4	58.0%	90.1%	19.6%	98.1%
16	75.0%	28.5%	2.5	75.0%	71.5%	9.9%	98.6%
15	91.2%	39.8%	2.3	91.2%	60.2%	8.7%	99.4%

【0092】

これらの値を用いて、閾値として19を用いることは、薬物から利益を受けることができる個体5.6%のみを消失させながら、有害事象に関して最高のリスクを有する患者約51.6%を排除するであろう。したがって、閾値として19を用いて、その4%が有害事象を経験するリスクが高いと仮定して、被験者1000人をスクリーニングする場合、 $74[(1000)(0.04)$

$(.516) + (1000)(.96)(.056)$ 例が除外され、残りの926例が治療されるであろう。したがって、治療した個体に対する有害事象のリスクは、 $[(1000)(0.04)(1-.516)/926 = 0.02]$ または2%であろう。このように、薬物の投与前に、患者集団を階層化するために診断における閾値として19を用いれば、有害事象の発生率を4%から2%に減少させ、それによってより広い用途に関して薬物を資格を与えるであろう。同様に、18も同様に閾値として用いることができ、この場合は、個体23/1000を除外して、治療した個体における有害事象の予測発生率は1.9%となるであろう。しかし、有害事象の発生率のこのような減少は、試験に関する特異性およびPPVの双方における減少に連結している。適当なリスク/利益診断閾値の選択は、試験そのものに関する情報（特異性、感度、PPV、NPV等）のみならず、本明細書に示した方法の実践者と管理機関（例えば、FDA）との協力作用、および臨床有用性に基く判断を必要とする可能性がある。そのような薬学ゲノム学の目標は、PPVとバランスをとりながら（薬物から利益を受けることができる患者の除外を最小限にする）、NPVを最大限にすることであろう（治療した個体における有害事象の発生率を減少させる）。有害事象の発生率を減少させるために本明細書に記述の方法を用いることは、新薬の承認を助け、既に承認された薬物のより広い使用を促進することができるであろう。例えば、そのような診断を薬物に連結させることによって、有害事象の発生率を一般的に許容されるレベルまで低下させて、結果的にそうでなければ承認されないであろう薬物を救済することが可能となるであろう。

10

【0093】

薬物に結合させた診断の承認に関する適当な閾値は、薬物の提供元（例えば、製薬会社）、および管理機関（例えば、F.D.A）のあいだでの交渉に大きく依存することは当業者に明らかであろう。これは、診断が薬物の有効性または安全性を改善するためである場合である。例えば、有害事象の発生率が先の例において2%まで低下した場合、管理機関はより厳密な安全性レベルを要求する可能性があり、したがって、薬物の治療から除外するための個体を同定するためにより低い閾値を要求する可能性があり、それによってより高いNPVのためにPPVを犠牲にする可能性がある。

20

【0094】

特定の局面において、本発明は、多因子形質を発症または示す個体のリスクを決定するための大きく改善された方法を提供する。特定の局面において、方法はさらに、多因子疾患の予後、診断、または治療を開発するために用いられる。他の局面において、方法は、治療レジメを投与する前に個体における薬物反応を予測するために用いられる。本明細書に示した方法はさらに、貴重な時間およびお金が限られた価値の治療に浪費されないように、個体に関して正しい医学的介入（最も有効な、最も安全な、最も安価な等）を迅速に発見するための手段を提供することによって、医療の全体的な費用を減少させるために役立つであろう。上記の説明は説明的であって、制限的ではないと理解される。様々な態様および改変を本明細書に開示の本発明に行ってもよく、それらも本発明の範囲および趣旨に含まれることは当業者に容易に明らかとなるはずである。したがって、本発明の範囲は、上記の説明を参照して決定されてはならず、その代わりに、そのような請求の範囲が資格を有する同等物の全ての範囲と共に添付の請求の範囲を参照して決定すべきである。本明細書において言及した刊行物は全て、本発明に関連して用いてもよい試薬、方法論、および概念を説明および開示する目的で引用されている。これらの参考文献は本明細書に記述の本発明に関連して先行技術であると認めたわけではないと解釈すべきである。開示を通して、様々な特許、特許出願、および刊行物を参照する。特に明記していなければ、それぞれはその全ての目的に関してその全内容物が参照により本明細書に組み入れられる。

30

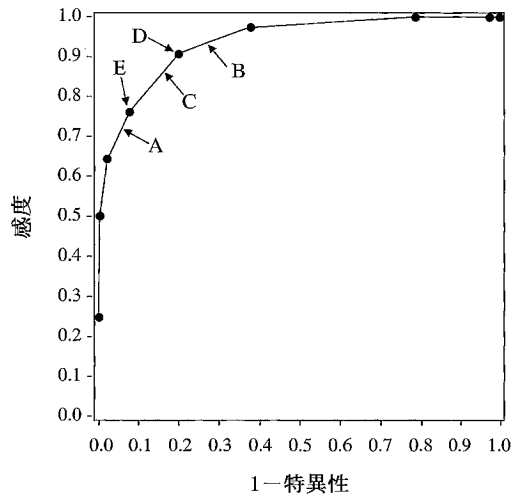
40

【図面の簡単な説明】

【0095】

【図1】多遺伝子試験の閾値を確立するための例としての受信者動作特性曲線を示す。

【図 1】



【国際調査報告】

INTERNATIONAL SEARCH REPORT		International application No. PCT/US05/07375
A. CLASSIFICATION OF SUBJECT MATTER IPC: G06F 19/00(2006.01); C12Q 1/68(2006.01) USPC: 702/20; 435/6 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) U.S. : 702/20; 435/6 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 practicable, search terms used) Please See Continuation Sheet		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	ZEE et al. Multi-locus interactions predict risk for post-PTCA restenosis: an approach to the genetic analysis of common complex disease. The Pharmacogenomic Journal. 2002, Vol. 2, pages 197-201, especially page 197.	1-64
☐ Further documents are listed in the continuation of Box C. ☐ 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 application or patent 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 12 May 2006 (12.05.2006)	Date of mailing of the international search report 15 JUN 2006	
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US Commissioner for Patents P.O. Box 1450 Alexandria, Virginia 22313-1450 Facsimile No. (571) 273-3201	Authorized officer John S. Brusca Telephone No. 571 272-1600	

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US05/07375

Continuation of B. FIELDS SEARCHED Item 3:

Medline, Biosis, US Patent issued and publications, Derwent WPI

search terms: snp, single nucleotide polymorphism, multifactorial, polygenic, diagnostic, predict, score, threshold

フロントページの続き

(51)Int.Cl.	F I	テーマコード (参考)
C 1 2 N 15/09 (2006.01)	C 1 2 N 15/00	F
	C 1 2 N 15/00	A

(31)優先権主張番号 10/956,224

(32)優先日 平成16年9月30日(2004.9.30)

(33)優先権主張国 米国(US)

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

(72)発明者 マクカミッシュ マーク

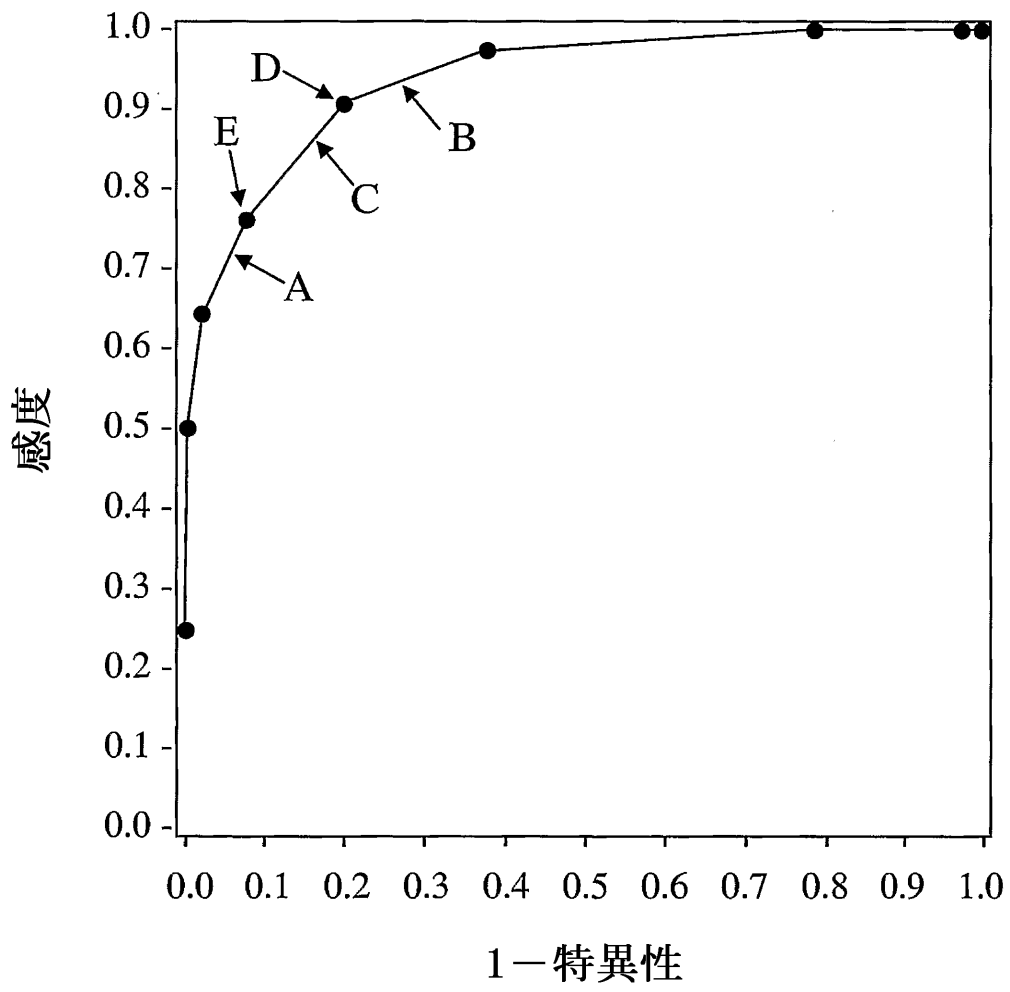
アメリカ合衆国 カリフォルニア州 クパチーノ ルニャール ロード 2 2 3 6 2

Fターム(参考) 2G045 AA25 AA40 JA01 JA03

4B024 AA11 CA01 CA09 CA11 HA11

4B063 QA01 QA13 QA19 QR32 QR35 QR55 QS32 QX01

【要約の続き】



专利名称(译)	<无法获取翻译>		
公开(公告)号	JP2007529714A5	公开(公告)日	2008-04-17
申请号	JP2007502088	申请日	2005-03-03
[标]申请(专利权)人(译)	每摄政科学墨水		
申请(专利权)人(译)	Parejen科学墨水.		
[标]发明人	コックスデビッドアール マクカミッシュマーク		
发明人	コックス デビッド アール. マクカミッシュ マーク		
IPC分类号	G01N33/53 C12Q1/68 C12Q1/02 G01N33/15 G01N33/50 C12N15/09		
CPC分类号	G16B20/00 G16B40/00 G16H20/00 G16H50/30		
FI分类号	G01N33/53.M C12Q1/68.A C12Q1/02 G01N33/15.Z G01N33/50.Z C12N15/00.F C12N15/00.A		
F-TERM分类号	2G045/AA25 2G045/AA40 2G045/JA01 2G045/JA03 4B024/AA11 4B024/CA01 4B024/CA09 4B024/CA11 4B024/HA11 4B063/QA01 4B063/QA13 4B063/QA19 4B063/QR32 4B063/QR35 4B063/QR55 4B063/QS32 4B063/QX01		
代理人(译)	清水初衷		
优先权	60/550662 2004-03-05 US 60/566302 2004-04-28 US 60/590534 2004-07-22 US 10/956224 2004-09-30 US		
其他公开文献	JP2007529714A		

摘要(译)

描述了用于评估个体发展或表现出多因子性状的可能性以及用于预测个体中药物治疗方案的有效性的方法。该方法包括在多个双等位基因多态性基因座处确定个体的多个基因型，使用基因型计算个体的分数，并将该分数与至少一个阈值进行比较。还描述了基因测试，用于评估个体发展或表现出多因素性状的可能性。