

(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2012-501181  
(P2012-501181A)

(43) 公表日 平成24年1月19日(2012.1.19)

(51) Int.Cl.	F I	テーマコード (参考)
<b>C 1 2 Q 1/68 (2006.01)</b>	C 1 2 Q 1/68 Z N A A	2 G O 4 5
<b>G O 1 N 33/53 (2006.01)</b>	G O 1 N 33/53 M	4 B O 6 3
<b>G O 1 N 33/68 (2006.01)</b>	G O 1 N 33/68	

審査請求 未請求 予備審査請求 未請求 (全 96 頁)

(21) 出願番号 特願2011-525187 (P2011-525187)  
 (86) (22) 出願日 平成21年8月27日 (2009. 8. 27)  
 (85) 翻訳文提出日 平成23年4月21日 (2011. 4. 21)  
 (86) 国際出願番号 PCT/US2009/055144  
 (87) 国際公開番号 W02010/025216  
 (87) 国際公開日 平成22年3月4日 (2010. 3. 4)  
 (31) 優先権主張番号 61/092, 270  
 (32) 優先日 平成20年8月27日 (2008. 8. 27)  
 (33) 優先権主張国 米国 (US)

(71) 出願人 591143065  
 ハー・ルンドベック・アクチエゼルスカベ  
 ット  
 デンマーク国、2500 バルビーエーコ  
 ンハーゲン、オッテイリアベエイ、9  
 (74) 代理人 100069556  
 弁理士 江崎 光史  
 (74) 代理人 100111486  
 弁理士 鍛冶澤 貴  
 (74) 代理人 100139527  
 弁理士 上西 克礼  
 (74) 代理人 100164781  
 弁理士 虎山 一郎

最終頁に続く

(54) 【発明の名称】 バイオマーカー・プロファイルを測定するためのシステムおよび方法

(57) 【要約】

本発明は、情動障害の患者を診断するための方法およびシステムに関する。本方法は、対象体の情動障害に対する感受性を予測するのにも有用である。

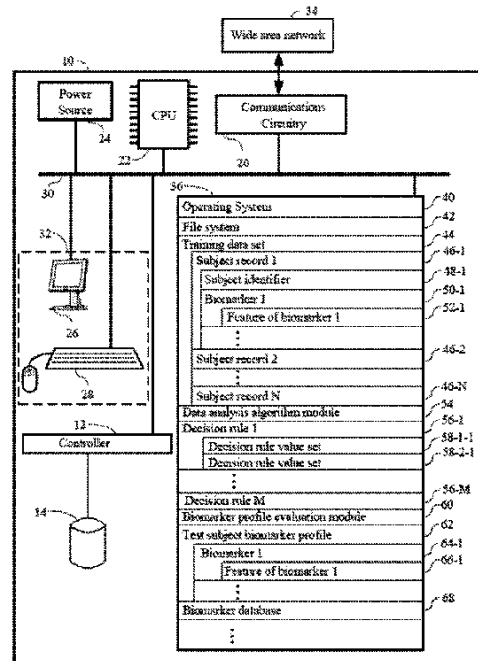


Figure 1.

## 【特許請求の範囲】

## 【請求項 1】

試験対象の情動障害を診断する方法であって、

前記試験対象のバイオマーカー・プロファイル中の複数のバイオマーカーの複数の特徴が、ある値セットを満たすかどうかを評価し、前記値セットを満たすことが、前記情動障害を前記試験対象が罹患していることを予測し、そして前記複数の特徴が、前記複数のバイオマーカーの測定可能な態様であり、前記複数のバイオマーカーが表 1 A に列挙されている少なくとも 2 つのバイオマーカーを含む方法。

## 【請求項 2】

前記試験対象が前記情動障害を罹患しているかどうかの診断を、ユーザー・インターフェース・デバイス、モニター、実在するコンピュータ読み取り可能な記憶媒体、またはローカルもしくはリモート・コンピュータ・システムに出力すること；または前記試験対象が前記情動障害を罹患しているかどうかの診断を、ユーザー読み取り可能な形態で表示することをさらに含む、請求項 1 に記載の方法。

10

## 【請求項 3】

前記複数のバイオマーカーが、表 1 A に列挙されている 2 ~ 29 個のバイオマーカーからなる、請求項 1 に記載の方法。

## 【請求項 4】

前記複数のバイオマーカーが、表 1 A に列挙されている 3 ~ 20 個のバイオマーカーからなる、請求項 1 に記載の方法。

20

## 【請求項 5】

前記複数のバイオマーカーが、表 1 A に列挙されている少なくとも 2 つのバイオマーカーを含む、請求項 1 に記載の方法。

## 【請求項 6】

前記複数のバイオマーカーが、表 1 A に列挙されている少なくとも 3 つのバイオマーカーを含む、請求項 1 に記載の方法。

## 【請求項 7】

前記複数のバイオマーカーが、表 1 A に列挙されている少なくとも 4 つのバイオマーカーを含む、請求項 1 に記載の方法。

## 【請求項 8】

前記複数の特徴が、表 1 A に列挙されている 2 ~ 29 個のバイオマーカーに対応する 2 ~ 29 個の特徴からなる、請求項 1 に記載の方法。

30

## 【請求項 9】

前記複数の特徴が、表 1 A に列挙されている 3 ~ 15 個のバイオマーカーに対応する 3 ~ 15 個の特徴からなる、請求項 1 に記載の方法。

## 【請求項 10】

前記複数の特徴が、表 1 A に列挙されている少なくとも 2 つのバイオマーカーに対応する少なくとも 2 つの特徴を含む、請求項 1 に記載の方法。

## 【請求項 11】

前記複数のバイオマーカーが、ERK1 および MAPK14 を含む、請求項 1 に記載の方法。

40

## 【請求項 12】

前記複数のバイオマーカーが、Gi2 および IL-1b を含む、請求項 1 に記載の方法。

## 【請求項 13】

前記複数のバイオマーカーが、ARRB1 および MAPK14 を含む、請求項 1 に記載の方法。

## 【請求項 14】

前記複数のバイオマーカーが、ERK1 および IL1b を含む、請求項 1 に記載の方法。

50

- 【請求項 15】  
前記複数のバイオマーカーが、ARRB1、IL6およびCD8aを含む、請求項1に記載の方法。
- 【請求項 16】  
前記複数のバイオマーカーが、ARRB1、ODC1およびP2X7を含む、請求項1に記載の方法。
- 【請求項 17】  
前記複数のバイオマーカー中の各バイオマーカーが核酸である、請求項1に記載の方法。
- 【請求項 18】 10  
前記複数のバイオマーカー中の各バイオマーカーが、DNA、cDNA、増幅DNA、RNAまたはmRNAである、請求項1に記載の方法。
- 【請求項 19】  
前記試験対象の前記バイオマーカー・プロファイル中の前記複数の特徴中のある特徴が、前記複数のバイオマーカー中のあるバイオマーカーの測定可能な態様であり、そして前記特徴の特徴値が、前記試験対象から採取された生体試料を使用して決定される、請求項1に記載の方法。
- 【請求項 20】 20  
前記特徴が、前記生体試料中の前記バイオマーカーの含有量であり、前記生体試料が全血である、請求項19に記載の方法。
- 【請求項 21】  
前記評価ステップに先立って、前記第1の値セットを構築することをさらに含む、請求項1に記載の方法。
- 【請求項 22】  
前記構築ステップが、データ分析アルゴリズムを、集団のメンバーから得られる特徴に適用することを含む、請求項21に記載の方法。
- 【請求項 23】 30  
前記集団が、前記情動障害を罹患していない第1の複数の対照対象に由来する第1の複数の生体試料、および前記情動障害を罹患している第2の複数の対象に由来する第2の複数の生体試料を含む、請求項22に記載の方法。
- 【請求項 24】  
前記データ分析アルゴリズムが、決定木、マイクロアレイ予測解析、多重加法型回帰木、ニューラル・ネットワーク、クラスタリング・アルゴリズム、主成分分析法、最近傍解析、線形判別分析、二次判別分析、サポート・ベクター・マシン法、進化法、関連ベクター・マシン法、遺伝的アルゴリズム、射影追跡法または重み付き投票法である、請求項22に記載の方法。
- 【請求項 25】 40  
前記構築ステップが判断ルールを生成し、そして前記評価ステップが、前記判断ルールを前記複数の特徴に適用して、前記複数の特徴が前記第1の値セットを満たすかどうかを決定することを含む、請求項21に記載の方法。
- 【請求項 26】  
前記判断ルールが、前記集団中の対象を(i)前記情動障害を罹患していない対象および(ii)70パーセント以上の正確度で前記情動障害を罹患している対象として分類する、請求項25に記載の方法。
- 【請求項 27】  
前記判断ルールが、前記集団中の対象を(i)前記情動障害を罹患していない対象および(ii)90パーセント以上の正確度で前記情動障害を罹患している対象として分類する、請求項25に記載の方法。
- 【請求項 28】 50  
前記情動障害が、双極性障害I、双極性障害II、気分変調性障害、または抑うつ障害

である、請求項 1 に記載の方法。

【請求項 29】

前記情動障害が、軽度うつ病、中程度のうつ病、重度うつ病、非定型うつ病、メランコリー型うつ病または境界性人格障害である、請求項 1 に記載の方法。

【請求項 30】

コンピュータ読み取り可能な記憶媒体およびそこに内蔵されているコンピュータ・プログラム・メカニズムを含み、前記コンピュータ・プログラム・メカニズムが、請求項 1 に記載の方法を実行するためのインストラクションを含む、コンピュータ・プログラム製品。

【請求項 31】

前記コンピュータ・プログラム・メカニズムが、前記試験対象が前記情動障害を罹患しているかどうかの診断をユーザー・インターフェース・デバイス、モニター、実在するコンピュータ読み取り可能な記憶媒体またはローカルもしくはリモート・コンピュータ・システムに出力するためのインストラクション；または前記試験対象が前記情動障害を罹患しているかどうかの診断を、ユーザー読み取り可能な形態で表示するためのインストラクションをさらに含む、請求項 30 に記載のコンピュータ・プログラム製品。

【請求項 32】

1 つまたは複数のプロセッサ、

前記 1 つまたは複数のプロセッサに接続されたメモリを含み、前記メモリが請求項 1 に記載の方法を実行するためのインストラクションを格納する、コンピュータ。

【請求項 33】

前記メモリが、前記試験対象が前記情動障害を罹患しているかどうかの診断をユーザー・インターフェース・デバイス、モニター、実在するコンピュータ読み取り可能な記憶媒体またはローカルもしくはリモート・コンピュータ・システムに出力するためのインストラクション；または前記試験対象が前記情動障害を罹患しているかどうかの診断をユーザー読み取り可能な形態で表示するためのインストラクションをさらに含む、請求項 32 に記載のコンピュータ。

【請求項 34】

試験対象が情動障害の症状を示す可能性を決定する方法であって、

前記試験対象のバイオマーカー・プロファイル中の複数のバイオマーカーの複数の特徴が、ある値セットを満たすかどうかを評価し、前記値セットを満たすことが、情動障害の症状を前記試験対象が示す可能性を提供し、そして前記複数の特徴が前記複数のバイオマーカーの測定可能な態様であり、前記複数のバイオマーカーが表 1 A に列挙されている少なくとも 2 つのバイオマーカーを含む方法。

【請求項 35】

前記試験対象が情動障害の症状を示す可能性を、ユーザー・インターフェース・デバイス、モニター、実在するコンピュータ読み取り可能な記憶媒体またはローカルもしくはリモート・コンピュータ・システムに出力すること；または前記試験対象が情動障害の症状を示す可能性をユーザー読み取り可能な形態で表示することをさらに含む、請求項 34 に記載の方法。

【発明の詳細な説明】

【技術分野】

【0001】

本出願は、電子的形態で提出された配列表を含有しており、該配列表は、ファイル名が 71021-WO-PC T\_\_SequenceListing\_\_ST25.txt であり、サイズが 148,658 バイトであり、2009 年 8 月 25 日に作成された。該配列表は、参照によりその全体が本明細書に組み込まれる。

【0002】

本発明は、対照対象の遺伝子の mRNA 発現レベルを罹患している対象の mRNA 発現レベルに対してプロファイリングおよび比較することにより、障害を患っている対象の転

10

20

30

40

50

写プロフィールを同定する方法および組成物を提供する。本発明は、そのような対象のバイオマーカーと関連する転写プロフィールを決定することにより、情動障害など対象の障害を予測および診断するための方法および組成物をさらに提供する。

【背景技術】

【0003】

本出願の全体にわたって、種々の刊行物が、括弧内に引用することにより参照されている。これらの刊行物の開示は、本発明が関する技術水準についてより完全に記述するために、それらの全体が参照により本出願に組み込まれる。

【0004】

現行の精神医学診断分類、特に情動障害用の診断分類には、明確な臨床的記述が欠如しており、ある診断実体を別の実体と区別して説明する生物学的特徴が含まれていない。今日の分類は、情動障害、例えば大うつ病性障害の臨床徴候をさらに特定することを可能にするが、その判断基準は依然として大きな論争の的であり、必ずしも生物学的な原理に基づいていない (Parkerら、Am. J. Psychiatry 2000年、157巻(8号) : 1195 ~ 1203頁 (非特許文献1))。

10

【0005】

情動障害には、重度対軽度または中等度うつ病、メランコリー型対非定型うつ病などの精神病性うつ病を含む双極性障害 I および II、気分変調症ならびに大うつ病性障害などの多くの臨床区分が存在する。そのため、これらの区分に関する明確な生物学的マーカーまたはバイオマーカーは、記述されていない。さらに、特定の障害に関する区分が欠如していることは、治療に影響を及ぼす場合がある。さらに、共存症は、2つの障害の存在を区別して説明することができない医師にとって問題となる。

20

【0006】

まとめると、精神医学における臨床評価および非特異的な臨床診断判断基準は、同様の生物学を共有する患者を認識するために生物学的マーカーの必要性を強調している。これは、情動障害に関する特定のジレンマと考えられる。なぜなら臨床的差異および明確な生物学的特徴を示すサブタイプが存在する根拠が出現しつつあるからである (GoldおよびChrousos、Mol. Psychiatry 2002年、7巻(3号) : 254 ~ 275頁 (非特許文献2))。しかしながら、これまでのところ、情動障害に関して患者集団の区分を区別して説明することが一貫して示されている生物学的マーカーは存在しない。

30

【0007】

以前の研究により、デキサメタゾン / 副腎皮質刺激ホルモン放出ホルモン (DEX / CRH) 検査など、対照の対象と比べたうつを罹患している対象の生物学的変化、または治療前後の対象の生物学的変化を測定する検査が研究されている。しかしながら、そのような検査は、少数の患者で検討されており、再現性がなく、および / または生物学的に読み取った情報を特定の表現型と関連させていない。(Ising, M.ら、Biol. Psychiatry、2006年、11月20日、出版前の電子出版 (非特許文献3) ; Kunugi, H.ら、Neuropsychopharm. 2006年、31巻(1号) : 212 ~ 20頁 (非特許文献4))。臨床関連バイオマーカーは、特定の生物学および特定の表現型に関連付けられなければならない、理想的には、治療により正常レベルに復帰すべきであるため、これは妥当である。

40

【0008】

糖尿病、アルツハイマー病、および癌についてのタンパク質バイオマーカーが同定されている。(例えば、米国特許第7,125,663号(特許文献1)、第7,097,989号(特許文献2)、第7,074,576号(特許文献3)、および第6,925,389号(特許文献4)を参照。)しかしながら、質量分析法および抗体に対する特異的結合などのタンパク質バイオマーカーを検出する方法は、再現不能なデータをもたらすことが多く、これらの方法は、ハイスループットでの使用に有利ではない。

【0009】

マイクロアレイを使用するハイスループット発現分析法が、遺伝子発現の変化を評価するために使用されているが、入り混じった結果または無関係な結果をもたらしている (Br

50

enner, S.ら、Nat Biotechnol. 2000年、18巻(6号) : 597~8頁(非特許文献5) ; Schenaら、Science. 1995年、270巻(5235号) : 467~70頁(非特許文献6) ; Velculescu, V.E.ら、Science. 1995年、270巻(5235号) : 484~7頁(非特許文献7) )。対象の数に対して測定される遺伝子発現の比率が大きいいため、抑うつ障害が異質性であることを考慮すると、マイクロアレイ・データでは多数の偽陽性が予測される。(総説は以下を参照、Iwamoto K、およびKato T.、Neuroscientist 2006年、12巻(4号) : 349~61頁(非特許文献8) ; Bunney WEら、Am J Psychiatry 2003年、160巻(4号) : 657~66頁(非特許文献9) ; ならびにIga J、Ueno S、およびOhmori T.、Ann Med 2008年、40巻(5号) : 336~42頁(非特許文献10) )。Sibilleら(Neuropsychopharm. 2004年、29巻(2号) : 351~61頁(非特許文献11) )は大規模ゲノム解析を実施したが、うつ病および自殺に相関する分子的差異を立証する根拠が見つからず、うつ病に関連することが以前に見出されていた遺伝子の発現レベルの変化を再現することができなかった。このような問題のため、一貫したプロファイルは同定されていない。

10

## 【0010】

複数の関連遺伝子のフォーカス・アレイ(focused array)およびqPCRを使用して、ストレス関連遺伝子が同定されているが、これらの研究は、うつ病に関連する診断プロファイルの同定に未だ至っていない(Rokutanら、J. Med. Invest. 2005年、52巻(3~4号) : 137~44頁(非特許文献12) ; Ohmoriら、J. Med. Invest. 2005年、52巻(補巻) : 266~71頁(非特許文献13) )。ラット脳領域では、特定遺伝子の転写における変化が、気分および不安の制御に関連すると考えられているが、これらの変化はヒト血液試料とは相関しない(国際公開第2007106685号A2(特許文献5))。

20

## 【先行技術文献】

## 【特許文献】

## 【0011】

【特許文献1】米国特許第7,125,663号

【特許文献2】米国特許第7,097,989号

【特許文献3】米国特許第7,074,576号

【特許文献4】米国特許第6,925,389号

【特許文献5】国際公開第2007106685号A2

【特許文献6】国際公開第03/061564号

【特許文献7】米国特許第5,981,180号

【特許文献8】国際出願第PCT/US87/00880号

【特許文献9】米国特許出願公開第2003/30134307号A1

30

## 【非特許文献】

## 【0012】

【非特許文献1】Parkerら、Am. J. Psychiatry 2000年、157巻(8号) : 1195~1203頁

【非特許文献2】GoldおよびChrousos、Mol. Psychiatry 2002年、7巻(3号) : 254~275頁

【非特許文献3】Ising, M.ら、Biol. Psychiatry、2006年、11月20日、出版前の電子出版

40

【非特許文献4】Kunugi, H.ら、Neuropsychopharm. 2006年、31巻(1号) : 212~20頁

【非特許文献5】Brenner, S.ら、Nat Biotechnol. 2000年、18巻(6号) : 597~8頁

【非特許文献6】Schenaら、Science. 1995年、270巻(5235号) : 467~70頁

【非特許文献7】Velculescu, V.E.ら、Science. 1995年、270巻(5235号) : 484~7頁

【非特許文献8】Iwamoto K、およびKato T.、Neuroscientist 2006年、12巻(4号) : 349~61頁

【非特許文献9】Bunney WEら、Am J Psychiatry 2003年、160巻(4号) : 657~66頁

【非特許文献10】Iga J、Ueno S、およびOhmori T.、Ann Med 2008年、40巻(5号) : 336~42頁

【非特許文献11】Sibilleら、Neuropsychopharm. 2004年、29巻(2号) : 351~61頁

50

- 【非特許文献 1 2】Rokutanら、J. Med. Invest. 2005年、52巻(3~4号) : 137~44頁
- 【非特許文献 1 3】Ohmoriら、J. Med. Invest. 2005年、52巻(補巻) : 266~71頁
- 【非特許文献 1 4】Atkinson, A.J.ら、Biomarkers and Surrogate Endpoints: Preferred Definitions and Conceptual Framework, Clinical Pharm. & Therapeutics, 2001年3月、69巻(3号) : 89~95頁
- 【非特許文献 1 5】Draghici、2003年、Data Analysis Tools for DNA Microarrays、Chapman & Hall/CRC社
- 【非特許文献 1 6】Lewin、2002年、Genes VII. Oxford University Press Inc.社、ニューヨーク市、ニューヨーク州
- 【非特許文献 1 7】Hastieら、2001年、The Elements of Statistical Learning、Springer-Verlag社、ニューヨーク市 10
- 【非特許文献 1 8】Lenoxら、2002年、American Journal of Medical Genetics (Neuropsychiatric Genetics) 114巻 : 391~406頁
- 【非特許文献 1 9】Paul編、2003年、Fundamental Immunology、第5版、Raven Press社、ニューヨーク市、69~105頁
- 【非特許文献 2 0】Sambrookら、2001年、Molecular Cloning: A Laboratory Manual. 第3版、Cold Spring Harbor Laboratory Press (コールド・スプリング・ハーバー、ニューヨーク)
- 【非特許文献 2 1】Zouら、2002年、Oncogene 21巻 : 4855~4862頁
- 【非特許文献 2 2】Ausubelら、最新版、Current Protocols in Molecular Biology、Greene Publishing and Wiley-Interscience社、ニューヨーク市 20
- 【非特許文献 2 3】Shenaら、1996年、Proc. Natl. Acad. Sci. U.S.A. 93巻 : 10614頁
- 【非特許文献 2 4】Tijessen、1993年、Hybridization With Nucleic Acid Probes、Elsevier Science Publishers B.V.社
- 【非特許文献 2 5】Kricka、1992年、Nonisotopic DNA Probe Techniques、Academic Press社、サンディエゴ市、カリフォルニア州
- 【非特許文献 2 6】Draghici、Data Analysis Tools for DNA Microanalysis、2003年、CRC Press LLC社、ボカラトン、フロリダ州、342~343頁
- 【非特許文献 2 7】MullisおよびFaloona、1987年、Methods Enzymol. 155巻 : 335頁
- 【非特許文献 2 8】Walkerら、1992年、PNAS 89巻 : 382~396頁 30
- 【非特許文献 2 9】Fujaら、2004年、Journal of Biotechnology 108巻 : 193~205頁
- 【非特許文献 3 0】HarlowおよびLane、1988年、Antibodies: A Laboratory Manual、Cold Spring Harbor Laboratory Press社(コールド・スプリング・ハーバー、ニューヨーク市)
- 【非特許文献 3 1】Gnanadesikan、1977年、Methods for Statistical Data Analysis of Multivariate Observations、ニューヨーク市 : Wiley社、1977年
- 【非特許文献 3 2】Breiman、1984年、Classification and Regression Trees、ベルモント、カリフォルニア州 : Wadsworth International Group社
- 【非特許文献 3 3】Tibshirani、1990年、Generalized Additive Models、ロンドン : Chapman and Hall社 40
- 【非特許文献 3 4】Neal、1996年、Bayesian Learning for Neural Networks、ニューヨーク市 : Springer-Verlag社
- 【非特許文献 3 5】Insua、1998年、Feedforward neural networks for nonparametric regression In: Practical Nonparametric and Semiparametric Bayesian Statistics、181~194頁、ニューヨーク市 : Springer社
- 【非特許文献 3 6】Harper、「Pyrolysis and GC in Polymer Analysis」、Dekker社、ニューヨーク市(1985年)
- 【非特許文献 3 7】Wagnerら、2002年、Anal. Chem. 74巻 : 1824~1835頁
- 【非特許文献 3 8】Brightら、2002年、J. Microbiol. Methods 48巻 : 127~38頁
- 【非特許文献 3 9】Dalluge、2000年、Fresenius J. Anal. Chem. 366巻 : 701~711頁 50

- 【非特許文献40】Duda、2001年、Pattern Classification、John Wiley & Sons, Inc.社、ニューヨーク市、395~396頁
- 【非特許文献41】Duda、2001年、Pattern Classification、John Wiley & Sons, Inc.社、ニューヨーク市、396~408頁および411~412頁
- 【非特許文献42】Hastieら、2001年、The Elements of Statistical Learning、Springer-Verlag社、ニューヨーク市、9章
- 【非特許文献43】Breiman、1999年、「Random Forests - Random Features」、Technical Report 567、Statistics Department、U.C.Berkeley、1999年9月
- 【非特許文献44】Duda、Pattern Classification、第2版、2001年、John Wiley & Sons, Inc社 10
- 【非特許文献45】Duda、2001年、Pattern Classification、John Wiley & Sons, Inc.社、ニューヨーク市、408~409頁
- 【非特許文献46】Hastieら、2001年、The Elements of Statistical Learning、Springer-Verlag社、ニューヨーク市、283~295頁
- 【非特許文献47】Mount、2001年、Bioinformatics: sequence and genome analysis、Cold Spring Harbor Laboratory Press社、コールド・スプリング・ハーバー、ニューヨーク市
- 【非特許文献48】Abdi、1994年、「A neural network primer」、J. Biol System. 2巻、247~283頁
- 【非特許文献49】DudaおよびHart、Pattern Classification and Scene Analysis、1973年、John Wiley & Sons, Inc.社、ニューヨーク市、211~256頁 20
- 【非特許文献50】Dudaら、Pattern Classification、第2版、John Wiley & Sons, Inc社、ニューヨーク市
- 【非特許文献51】KaufmanおよびRousseeuw、1990年、Finding Groups in Data: An Introduction to Cluster Analysis、Wiley社、ニューヨーク市、ニューヨーク州
- 【非特許文献52】Everitt、1993年、Cluster analysis (第3版)、Wiley社、ニューヨーク市、ニューヨーク州
- 【非特許文献53】Backer、1995年、Computer-Assisted Reasoning in Cluster Analysis、Prentice Hall社、アッパー・サドル・リバー、ニュージャージー州
- 【非特許文献54】CristianiniおよびShawe-Taylor、2000年、An Introduction to Support Vector Machines、Cambridge University Press社、ケンブリッジ市 30
- 【非特許文献55】Boserら、1992年、「A training algorithm for optimal margin classifiers」、in Proceedings of the 5th Annual ACM Workshop on Computational Learning Theory、ACM Press社、ピッツバーグ、ペンシルベニア州、142~152頁
- 【非特許文献56】Vapnik、1998年、Statistical Learning Theory、Wiley社、ニューヨーク市
- 【非特許文献57】Fureyら、2000年、Bioinformatics 16巻、906~914頁
- 【非特許文献58】Tipping, M: Sparse Bayesian Learning and the Relevance Vector Machine、Journal of Machine Learning Research 1巻、2001年、211~244頁
- 【非特許文献59】Deb, K: Multi-Objective Optimization using Evolutionary Algorithms、Wiley社、2001年 40
- 【非特許文献60】Breiman, L.、2001年、Machine Learning 45巻(1号): 5~32頁
- 【非特許文献61】Cortes, C.およびVapnik, V.、1995年、Machine Learning、20巻(3号): 273~97頁
- 【非特許文献62】Ersboll, B.K.およびConradsen, K. (2005年) An Introduction to Statistics. 第7版、IMM
- 【非特許文献63】Draper, N.およびSmith, H. (1981年) Applied Regression Analysis、第2版、ニューヨーク市: John Wiley & Sons, Inc.社
- 【非特許文献64】James K.E.ら、2005年、Statistics in Medicine、24巻(19号): 3019~35頁 50

【非特許文献 6 5】Dettling, M., 2003年、Proceedings of the 3rd International Workshop on Distributed Statistical Computing、3月20~22日、ウィーン、オーストリア、Hornick、Leisch、およびSeilis編

【非特許文献 6 6】Friedman, J., Hastie, T., およびTibshirani, R. 2000年、Annals of Statistics 28巻(2号): 337~407頁

【非特許文献 6 7】V. N. Vapnik、Statistical Learning Theory、Wiley社、ニューヨーク市、1998年

【非特許文献 6 8】Burnham, K. P. およびD. R. Anderson、2002年、Model Selection and Multimodel Inference: A Practical-Theoretic Approach、第2版、Springer-Verlag社

10

【非特許文献 6 9】Devijver, P. A. およびJ. Kittler、1982年、Pattern Recognition: A Statistical Approach、Prentice-Hall社、ロンドン

【非特許文献 7 0】米国精神医学会により出版されたThe Diagnostic and Statistical Manual of Mental Disorders、第4版、改訂版(DSM-IV-TR(登録商標))(1994年10月、2000年5月改訂)

【非特許文献 7 1】Andersen, C. L. ら、Cancer Res、2004年、64巻: 5245~5250頁

【非特許文献 7 2】Jin, P. ら、BMC Genomics、2004年、5巻: 55頁

【非特許文献 7 3】Huggett, J. ら、Genes and Immunity、2005年、6巻: 279~284頁

【非特許文献 7 4】Vandesompele, J. ら、Genome Biol、2002年、3巻(7号): RESEARCH034.1-0034.11(電子出版 2002年6月18日)

20

【非特許文献 7 5】Applied Biosystems Application Note 2006年、publication 127APO 8-01

【非特許文献 7 6】Livak, K. およびSchmittgen, T., Methods、2001年、25巻: 402~408頁

【非特許文献 7 7】Liaw, A., およびWiener, M. 2002年、12月、Classification and regression by randomForest. R News Vol. 2/3: 18~22頁

【発明の概要】

【発明が解決しようとする課題】

【0013】

本発明は、試験対象の情動障害を診断する方法であって、該試験対象のバイオマーカー・プロファイル中の複数のバイオマーカーの複数の特徴が、ある値セットを満たすかどうかを評価し、該値セットを満たすことが、前記情動障害を該試験対象が罹患していることを予測し、そして該複数の特徴が、該複数のバイオマーカーの測定可能な態様であり、該複数のバイオマーカーが、表1Aに列挙されている少なくとも2つのバイオマーカーを含む方法を提供する。

30

【課題を解決するための手段】

【0014】

本発明は、コンピュータ・プログラム製品も提供し、該コンピュータ・プログラム製品は、コンピュータ読み取り可能な記憶媒体およびそこに内蔵されているコンピュータ・プログラム・メカニズムを含み、該コンピュータ・プログラム・メカニズムは、本診断法を実行するためのインストラクションを含む。

40

【0015】

本発明の1つの態様は、1つまたは複数のプロセッサおよび該1つまたは複数のプロセッサに接続されたメモリを含み、該メモリが本診断法を実行するためのインストラクションを格納しているコンピュータを提供する。

【0016】

本発明の別の態様は、試験対象が情動障害の症状を示す可能性を決定する方法であって、該試験対象のバイオマーカー・プロファイル中の複数のバイオマーカーの複数の特徴が、ある値セットを満たすかどうかを評価し、該値セットを満たすことが、情動障害の症状を該試験対象が示す可能性を提供し、そして該複数の特徴が該複数のバイオマーカーの測

50

定可能な態様であり、該複数のバイオマーカーが表 1 A に列挙されている少なくとも 2 つのバイオマーカーを含む方法を提供する。

【0017】

別の態様では、本発明は、複数の対照対象から収集された各生体試料に関する転写解析の尺度である転写プロファイルを提供する。例えば、本発明は、複数のうつ病対象、重度うつ病対象、または双極性対象から収集された各生体試料に関する転写解析の尺度である転写プロファイルを提供する。本発明は、複数の境界性人格障害対象から収集された各生体試料に関する転写解析の尺度である転写プロファイルをさらに提供する。本発明は、複数の P T S D 対象から収集された各生体試料に関する転写解析の尺度である転写プロファイルも提供する。

10

【0018】

本発明は、第 1 の複数の対照対象の集団的尺度を含む転写プロファイルが、例えばデータベースに格納されることも提供する。第 2 の複数の対象、例えば罹患対象の集団的尺度を含む転写プロファイルは、分類アルゴリズムを使用して、第 1 の複数の対照対象の転写プロファイルと比較される。分類アルゴリズムは、対象の各々を分類する出力を提供する。

【0019】

本発明は、患者の転写プロファイルを同定することにより情動障害を診断する方法であって、そのような転写プロファイルを対照対象または対照対象群のプロファイルと比較し、それにより該転写プロファイルにおける変化の存在または非存在に基づいて該患者の情動障害を診断する方法を提供する。

20

【0020】

本発明の 1 つの態様は、情動障害を罹患している対象を診断するための方法であって、  
 ( a ) 複数の対照対象および複数の罹患対象から生体試料を取得すること、  
 ( b ) 該複数の対照対象および該複数の罹患対象の試料中の遺伝子の m R N A 発現レベルを測定し、該遺伝子が、A D A、A R R B 1、A R R B 2、C D 8 a、C D 8 b、C R E B 1、C R E B 2、D P P 4、E R K 1、E R K 2、G i 2、G s、G R、I L 1 b、I L 6、I L 8、I N D O、M A P K 1 4、M A P K 8、M K P 1、M R、O D C 1、P 2 X 7、P B R、P R E P、R G S 2、S 1 0 0 A 1 0、S E R T および V M A T 2 からなる群から選択されること、  
 ( c ) 該複数の対照対象および該複数の罹患対象に由来する各遺伝子の m R N A 発現レベルを収集し、コンピュータ媒体中の m R N A データとして保存すること、  
 ( d ) 分類アルゴリズムにより、そのような m R N A データを処理すること、および  
 ( e ) 該対象を分類する出力データを提供すること、  
 を含み、  
 それにより、情動障害を罹患している対象を診断する方法を提供する。

30

【0021】

本発明は、A D A、A R R B 1、A R R B 2、C D 8 a、C D 8 b、C R E B 1、C R E B 2、D P P 4、E R K 1、E R K 2、G i 2、G s、G R、I L 1 b、I L 6、I L 8、I N D O、M A P K 1 4、M A P K 8、M K P 1、M R、O D C 1、P 2 X 7、P B R、P R E P、R G S 2、S 1 0 0 A 1 0、S E R T および V M A T 2 からなる群から選択される遺伝子の対象の転写プロファイルを、複数の対照対象の遺伝子の転写プロファイルと比較することにより、情動障害に対する該対象の感受性を予測するための方法をさらに提供する。

40

【0022】

本発明の 1 つの態様は、対象が情動障害の症状を示す可能性を予測するための方法であって、  
 ( a ) 複数の対照対象および複数の罹患対象から生体試料を取得すること、  
 ( b ) 該複数の対照対象および該複数の罹患対象の試料中の遺伝子の m R N A 発現レベルを測定し、該遺伝子が、A D A、A R R B 1、A R R B 2、C D 8 a、C D 8 b、C R

50

EB1、CREB2、DPP4、ERK1、ERK2、Gi2、Gs、GR、IL1b、IL6、IL8、INDO、MAPK14、MAPK8、MKP1、MR、ODC1、P2X7、PBR、PREP、RGS2、S100A10、SERTおよびVMAT2からなる群から選択されること、

(c) 該複数の対照対象および該複数の罹患対象に由来する各遺伝子のmRNA発現レベルを収集し、コンピュータ媒体中のmRNAデータとして保存すること、

(d) 分類アルゴリズムにより、そのようなmRNAデータを処理すること、および

(e) 該対象を分類する出力データを提供すること、  
を含み、

それにより、該対象が情動障害の症状を示す可能性を予測する方法を提供する。

10

【図面の簡単な説明】

【0023】

【図1】本発明の実施形態によるコンピュータ・システムの図解である。

【図2】qPCR法により、cDNA1ng当たりのコピー数で測定された、対照対象対うつ病対象におけるARRB1(ベータ-アレスチン1)およびGi2(グアニン・ヌクレオチド結合タンパク質アルファi2)の相対的mRNAレベルをそれぞれ示す散布図である( $p < 0.001$ ; マン・ホイットニー検定)。

【図3】qPCR法により、cDNA1ng当たりのコピー数で測定された、対照対象対うつ病対象におけるMAPK14(p38マイトジェン活性化プロテインキナーゼ14)およびODC1(オルニチンデカルボキシラーゼ1)の相対的mRNAレベルをそれぞれ示す散布図である( $p < 0.001$ ; マン・ホイットニー検定)。

20

【図4】qPCR法により、cDNA1ng当たりのコピー数で測定された、対照対象対重度うつ病対象におけるERK1(細胞外シグナル制御キナーゼ1)、Gi2(グアニン・ヌクレオチド結合タンパク質アルファi2)、およびMAPK14(p38マイトジェン活性化プロテインキナーゼ14)の相対的mRNAレベルをそれぞれ示す散布図である( $p < 0.001$ ; マン・ホイットニー検定)。

【図5】qPCR法により、cDNA1ng当たりのコピー数で測定された、対照対象対重度うつ病/双極性対象におけるGi2(グアニン・ヌクレオチド結合タンパク質アルファi2)、GR(アルファ-グルココルチコイド受容体)、およびMAPK14(p38マイトジェン活性化プロテインキナーゼ14)の相対的mRNAレベルをそれぞれ示す散布図である( $p < 0.001$ ; マン・ホイットニー検定)。

30

【図6】qPCR法により、cDNA1ng当たりのコピー数で測定された、対照対象対境界性人格障害対象におけるGi2(グアニン・ヌクレオチド結合タンパク質アルファi2)、MAPK14(p38マイトジェン活性化プロテインキナーゼ14)、およびMR(ミネラルコルチコイド受容体)の相対的mRNAレベルをそれぞれ示す散布図である( $p < 0.001$ ; マン・ホイットニー検定)。

【図7】qPCR法により、cDNA1ng当たりのコピー数で測定された、196人の対照対象対66人の急性PTSD対象における、ARRB2(ベータ-アレスチン2)、ERK2(細胞外シグナル制御キナーゼ2)、およびRGS2(Gタンパク質シグナル伝達の調節因子2)の相対的mRNAレベルをそれぞれ示す散布図である( $p < 0.001$ ; マン・ホイットニー検定)。

40

【図8】図8Aは、うつ病対象対対照の分類において、遺伝子選択およびトレーニングの両方を実施し、93%の正確度、PPV=93%、およびNPV=94%をスコアするSLRアルゴリズムの性能を図示する。RF遺伝子選択に先行するサポート・ベクター・マシン(SVM)分類器(classifier)は、うつ病対象対対照の分類において、88%の正確度、PPV=89%、およびNPV=88%をスコアする。図8Bは、うつ病対象対対照の分類において、各方法の統計パラメータに基づいて、ランダム・フォレスト法(RF)が表1Aから14個の遺伝子を選択し、段階的ロジスティック回帰法(SLR)が表1Aから17個の遺伝子を選択することを示す。分類プロセスの選択ステップにおいてRF法およびSLR法の両方により選択された重複遺伝子は、灰色で示されている。

50

【図 9】図 9 は、平均発現レベル（転写値）が重度うつ病患者と対照との間で有意に異なっていた（ $p < 0.05$ ）遺伝子を示す。これら遺伝子は、表 5 A で見られるように、算出された  $-\log(p)$  値の大きさによりランク付けされている。

【図 10】図 10 は、各対象の ERK1 および MAPK14 からなる転写プロファイルによる、重度うつ病対象および対照対象の分布を表す。重度うつ病対象は、白抜きの円（ ）により表されており、対照対象は、黒抜きの三角形（ ）により表されている。X 軸および Y 軸は、それぞれ ERK1 および MAPK14 に対する転写値（cDNA 1 ng 当たりのコピー数）を示す。

【図 11】図 11 は、各対象の Gi2 および IL1b からなる転写プロファイルによる、重度うつ病対象および対照対象の分布を表す。重度うつ病対象は、白抜きの円（ ）により表されており、対照対象は、黒抜きの三角形（ ）により表されている。X 軸および Y 軸は、それぞれ Gi2 および IL1b に対する転写値（cDNA 1 ng 当たりのコピー数）を示す。

【図 12】図 12 は、各対象の ERK1 および IL1b からなる転写プロファイルによる、重度うつ病対象および対照対象の分布を表す。重度うつ病対象は、白抜きの円（ ）により表されており、対照対象は、黒抜きの三角形（ ）により表されている。X 軸および Y 軸は、それぞれ ERK1 および IL1b に対する転写値（cDNA 1 ng 当たりのコピー数）を示す。

【図 13】図 13 は、各対象の ARRB1 および MAPK14 からなる転写プロファイルによる、重度うつ病対象および対照対象の分布を表す。重度うつ病対象は、白抜きの円（ ）により表されており、対照対象は、黒抜きの三角形（ ）により表されている。X 軸および Y 軸は、それぞれ ARRB1 および MAPK14 に対する転写値（cDNA 1 ng 当たりのコピー数）を示す。

【発明を実施するための形態】

【0024】

本発明は、バイオマーカー・プロファイル中のバイオマーカー特徴を評価することにより、迅速で正確な情動障害の診断を可能にする。これらバイオマーカー・プロファイルは、対象の生体試料から構築される。

【0025】

#### 5.1 定義

本明細書で使用される場合、「情動障害」は、思考、感情、および挙動に影響を及ぼす、一貫した広範な気分の変化を特徴とする精神障害を意味するものとする。情動障害の例には、これらに限定されないが、抑うつ障害、不安障害、双極性障害、気分変調症、および統合失調感情障害が含まれる。不安障害には、これらに限定されないが、全般性不安障害、パニック障害、強迫性障害、恐怖症、および心的外傷後ストレス障害が含まれる。抑うつ障害には、これらに限定されないが、大うつ病性障害（MDD）、緊張性うつ病、メランコリー型うつ病、非定型うつ病、精神病性うつ病、産後うつ病、双極性うつ病、および軽度、中等度もしくは重度うつ病が含まれる。人格障害には、これらに限定されないが、妄想性人格障害、反社会性人格障害および境界性人格障害が含まれる。

【0026】

「バイオマーカー」は、生体試料中に存在するかまたは生体試料に由来する、タンパク質、ペプチド、プロテオグリカン、糖タンパク質、リポタンパク質、炭水化物、脂質、核酸（例えば、cDNA もしくは増幅 DNA などの DNA、または mRNA などの RNA）、有機もしくは無機化学物質、天然もしくは合成高分子、低分子（例えば、代謝産物）、または先述のもののいずれかの特徴的分子もしくは特徴的断片などの、事実上あらゆる検出可能な化合物、あるいは正常な生物学的プロセス、病原性プロセス、または治療介入に対する薬理的応答もしくはその兆候の指標として客観的に測定および評価される任意の他の特徴である。Atkinson, A.J.ら、Biomarkers and Surrogate Endpoints: Preferred Definitions and Conceptual Framework, Clinical Pharm. & Therapeutics, 2001年3月; 69巻(3号): 89~95頁(非特許文献14)を参照されたい。本状況で使用される場合

10

20

30

40

50

、「～に由来する」は、それが検出される場合、特定の分子が生体試料中に存在することを示す化合物を指す。例えば、特定のcDNAの検出は、生体試料中に特定のRNA転写物が存在することを示すことができる。別の例として、特定の抗体の検出または結合は、生体試料中に特定の抗原（例えば、タンパク質）が存在することを示すことができる。本明細書では、特徴的な分子または断片は、それが検出された場合、上記の特定された化合物の存在または含有量を示す分子または断片である。

#### 【0027】

バイオマーカーは、例えば、生体試料から単離されてもよく、または生体試料中で直接測定されてもよく、または生体試料に存在することを検出もしくは決定されてもよい。バイオマーカーは、例えば、機能的であってもよく、部分的に機能的であってもよく、または非機能的であってもよい。1つの実施形態では、バイオマーカーは単離され、例えば、様々な診断アッセイでのバイオマーカー検出を容易にすることができる特異的結合抗体を誘発するために使用される。任意のイムノアッセイは、バイオマーカー分子に結合可能な任意の抗体、抗体断片またはその誘導體（例えば、Fab、F(ab')<sub>2</sub>、Fv、またはscFv断片）を使用することができる。そのようなイムノアッセイは当技術分野で周知である。加えて、バイオマーカーがタンパク質またはその断片である場合、それを配列決定することができ、確立されている技術を使用して、そのコード遺伝子をクローニングすることができる。

10

#### 【0028】

本明細書で使用される場合、「バイオマーカーの種(a species of a biomarker)」という用語は、本明細書に記述されている特定の遺伝子（例えば、下記の表1Aに列挙されている遺伝子）のスプライス・バリエーションなどの、本明細書に記述されているバイオマーカーの任意の特徴的部分または特徴的断片を指す。本明細書では、特徴的部分または特徴的断片は、それが検出された場合、上記の特定された転写物、cDNA、増幅核酸またはタンパク質の存在または含有量を示す分子の部分または断片である。

20

#### 【0029】

「バイオマーカー・プロファイル」は、バイオマーカーの測定可能な態様（例えば、含有量）などの特徴と共に、1つまたは複数のタイプのバイオマーカー（例えば、mRNA分子、cDNA分子、タンパク質および/もしくは炭水化物、またはその兆候など）を複数含んでいる。バイオマーカー・プロファイルは、そのようなバイオマーカーを少なくとも2つ含み、該バイオマーカーは、例えば、核酸および炭水化物など、同じ種類であってもよく、または異なる種類であってもよい。また、バイオマーカー・プロファイルは、少なくとも2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95または100またはそれ以上のバイオマーカーを含んでいてもよい。1つの実施形態では、バイオマーカー・プロファイルは、数百、または数千ものバイオマーカーを含んでいる。バイオマーカー・プロファイルは、1つまたは複数の対照または内部標準をさらに含むことができる。1つの実施形態では、バイオマーカー・プロファイルは、内部標準として機能する少なくとも1つのバイオマーカーを含む。「兆候」という用語は、本状況において本明細書で使用される場合、バイオマーカー・プロファイルが、核酸、mRNA分子、cDNA分子、タンパク質および/もしくは炭水化物、またはバイオマーカー分子体自体ではないバイオマーカーの任意の他の形態についての記号、データ、略語または他の同様の兆候を含有している状況を単に指す。例えば、本発明の例示的バイオマーカー・プロファイルは、表1Aの遺伝子の名称を含む。

30

40

#### 【0030】

バイオマーカー・プロファイル中の各バイオマーカーは、対応する「特徴」を含んでいる。本明細書で使用される場合、「特徴」は、バイオマーカーの測定可能な態様を指す。特徴には、例えば、例示的バイオマーカー・プロファイル1に例示されているような、対

50

象由来の生体試料中のバイオマーカの存在または非存在を含むことができる。

【0031】

【表1】

例示的バイオマーカ・プロファイル1

バイオマーカ	特徴 試料中の存在
遺伝子Aの転写物	存在
遺伝子Bの転写物	非存在

【0032】

例示的バイオマーカ・プロファイル1では、遺伝子Aの転写物についての特徴値は「存在」であり、遺伝子Bの転写物についての特徴値は「非存在」である。

【0033】

特徴は、例えば、例示的バイオマーカ・プロファイル2に例示されているような、対象由来の生体試料中のバイオマーカの含有量を含むことができる。

【0034】

【表2】

例示的バイオマーカ・プロファイル2

バイオマーカ	特徴 相対的単位での試料中の含有量
遺伝子Aの転写物	300
遺伝子Bの転写物	400

【0035】

例示的バイオマーカ・プロファイル2では、遺伝子Aの転写物についての特徴値は300単位であり、遺伝子Bの転写物についての特徴値は400単位である。

【0036】

また、特徴は、例示的バイオマーカ・プロファイル3に例示されているような、バイオマーカの2つまたはそれ以上の測定可能な態様の比率であってもよい。

【0037】

【表3】

例示的バイオマーカ・プロファイル3

バイオマーカ	特徴 遺伝子Aの転写物の含有量/遺伝子Bの転写物の含有量の比率
遺伝子Aの転写物	300/400
遺伝子Bの転写物	

【0038】

例示的バイオマーカ・プロファイル3では、遺伝子Aの転写物についての特徴値および遺伝子Bの転写物についての特徴値は、0.75(300/400)である。

【0039】

いくつかの実施形態では、上記の例示的バイオマーカ・プロファイル1に例示されているように、バイオマーカ・プロファイル中の特徴とバイオマーカとの間には一対一の対応が存在する。いくつかの実施形態では、上記の例示的バイオマーカ・プロファイル3に例示されているように、本発明のバイオマーカ・プロファイル中の特徴とバイオマーカとの間の関係性は、より複雑である。

【0040】

当業者であれば、特徴を計算する他の方法を考案することができ、そのような方法はすべて本発明の範囲内であることを認識するだろう。例えば、特徴は、2回またはそれ以上

10

20

30

40

50

の時点で対象から収集された生体試料にわたるバイオマーカーの含有量の平均を表すことができる。さらに、特徴は、単一の時点で対象から得られた生体試料に由来する2つまたはそれ以上のバイオマーカーの含有量の差異または比率であってもよい。また、バイオマーカー・プロファイルは、少なくとも2、3、4、5、10、20、30またはそれ以上の特徴を含んでいてもよい。1つの実施形態では、バイオマーカー・プロファイルは、数百、または数千もの特徴を含んでいる。

#### 【0041】

いくつかの実施形態では、バイオマーカーの特徴は、定量的PCR (qPCR) を使用して測定される。遺伝子転写物の含有量を測定するqPCRの使用は周知である。いくつかの実施形態では、バイオマーカー特徴は、マイクロアレイを使用して測定される。マイクロアレイの構築および含有量データを取得するためにマイクロアレイを処理するのに用いられる技術は周知であり、例えば、Draghici、2003年、Data Analysis Tools for DNA Microarrays、Chapman & Hall/CRC社 (非特許文献15) および国際公開第03/061564号 (特許文献6) により記述されている。マイクロアレイは、複数のプローブを含む。いくつかの場合では、各プローブは、異なるバイオマーカーを認識、例えば結合する。いくつかの場合では、マイクロアレイの2つまたはそれ以上の異なるプローブは、同じバイオマーカーを認識、例えば結合する。したがって、典型的には、マイクロアレイのプローブ・スポットと対象バイオマーカーとの間の関係性は、2対1対応、3対1対応、または他のいくつかの形態の対応である。しかしながら、マイクロアレイのプローブとバイオマーカーとの間に、固有の1対1の対応が存在する場合があります。

10

20

#### 【0042】

本明細書で使用される場合、「相補的」という用語は、核酸配列 (例えば本明細書に記述されている遺伝子をコードするヌクレオチド配列) の状況において、それらの水素結合特性の結果としての特定の窒素塩基間の化学親和性を指す。例えば、グアニン (G) はシトシン (C) とのみ水素結合を形成し、一方、アデニンは、DNAの場合チミン (T) とのみ水素結合を形成し、RNAの場合はウラシル (U) とのみ水素結合を形成する。これらの反応は、塩基対合と記述されており、対になった塩基 (GとC、またはAとT/U) は、相補的であると言われる。したがって、2つの核酸配列は、それらの窒素含有塩基が水素結合を形成することができる場合、相補的である可能性がある。そのような配列は、互いの「相補体」と呼ばれる。そのような相補体配列は、天然に存在してもよく、または例えばDNA分子またはRNA分子 (例えば、mRNA転写物) のセンス鎖に相補的なアンチセンス核酸分子の場合のように、当業者に公知の任意の方法により化学的に合成することができる。例えば、Lewin、2002年、Genes VII. Oxford University Press Inc. 社、ニューヨーク市、ニューヨーク州 (非特許文献16) を参照されたい。

30

#### 【0043】

本明細書で使用される場合、「データ分析アルゴリズム」とは、トレーニング集団中の対象のバイオマーカー・プロファイルを使用して判断ルールを構築するために使用されるアルゴリズムである。代表的なデータ分析アルゴリズムは、下記に記述されている。「判断ルール」は、データ分析アルゴリズムの最終産物であり、1つまたは複数の値セットを特徴とし、これらの値セットの各々は、情動障害の一態様、情動障害の発症、対象が情動障害になるという予測、または対象が情動障害の症状を示す可能性を示す。1つの特定の例では、値セットは、対象が情動障害を発症するだろうという予測を表す。別の例では、値セットは、対象が情動障害を発症しないだろうという予測を表す。

40

#### 【0044】

「判断ルール」は、バイオマーカー・プロファイルを評価するために使用される方法である。そのような判断ルールは、Hastieら、2001年、The Elements of Statistical Learning、Springer-Verlag社、ニューヨーク市 (非特許文献17) に例示されているように、当技術分野で公知の1つまたは複数の形態を取ることができる。判断ルールを、特徴のデータ・セットに対して実行するために使用し、とりわけ、情動障害の存在、または対象が情動障害の症状を示すか、もしくは有する可能性、または情動障害の発症に感受性を示

50

す可能性を予測することができる。本発明のいくつかの実施形態で使用することができる例示的な判断ルールは、下記でさらに詳細に記述されている。

【0045】

本明細書で使用される場合、「エンドフェノタイプ」という用語は、個体が症状を示しているかどうかに関わらず、疾病に関連するバイオマーカーなどの遺伝可能な特徴が存在することを意味するものとする。(総説は、Lenoxら、2002年、American Journal of Medical Genetics (Neuropsychiatric Genetics) 114巻：391～406頁(非特許文献18)を参照されたい。)

【0046】

本明細書で使用される場合、「遺伝子発現プロファイル」および「転写プロファイル」という用語は、選択された遺伝子のメッセンジャーリボ核酸(mRNA)レベルを相対的に測定することにより決定されるバイオマーカー・プロファイルである。転写プロファイルは、対象または患者の生体試料に由来する遺伝子の転写解析をすることにより測定される。

10

【0047】

本明細書で使用される場合、「健常対照対象」、「健常対照」、および、「対照対象」は、大きな医学的または精神医学的問題を現在有していない対象を意味するものとするが、例えば頭痛を患っていてもよい。対照対象は、好ましくは、低いボディマス指数(BMI、30未満)を有しており、過去3カ月間に薬物を使用しておらず、ストレス・スコア、家族歴スコア、および症状スコアが低いかまたは0である。対照対象は、自己記入質問書により判明するような、精神疾患のあらゆる病歴、物質乱用のあらゆる病歴、精神疾患のあらゆる家族歴、あらゆる若年期ストレス要因、またはあらゆる最近のストレス要因を有していない場合がある。対照対象は、生体試料を取得する前に医師によりさらに評価されてもよいが、評価される必要はない。

20

【0048】

「取得する」および「取得すること」という用語は、本明細書で使用される場合、それぞれ「所有に至る」または「所有に至ること」を意味する。これは、例えば、コンピュータ・システムの保存データからデータを取り出すことにより行うことができる。これはまた、例えば直接測定により行うことができる。

【0049】

本明細書で使用される場合、「表現型」という用語は、対象の遺伝子型および環境の結果である測定可能および/または観察可能な生物学的特徴、臨床特徴または行動特徴を意味するものとする。

30

【0050】

本明細書で使用される場合、「タンパク質」、「ペプチド」および「ポリペプチド」という用語は、別様の指定がない限り同義である。

【0051】

本明細書で使用される場合、「PTSD対照対象」は、極端な心的外傷性ストレス要因にさらされておらず、あらゆる神経精神疾患に罹っていないと医師により評価された対象を意味するものとする。本発明のPTSD対照対象は、一般的に一致した対象であり、例えば、同じ地理的地域からであり、障害を示す対象と同じ性別である。

40

【0052】

本明細書で使用される場合、「特異的に」という用語および類似の用語は、抗体の状況では、抗原または断片に特異的に結合し、他の抗原または他の断片に特異的に結合しないペプチド、ポリペプチドおよび抗体またはその断片を指す。標準的実験技術により、例えば当業者に周知の任意のイムノアッセイにより判明するように、抗原に特異的に結合するペプチドまたはポリペプチドは、より低い親和性で他のペプチドまたはポリペプチドに結合する場合がある。そのようなイムノアッセイには、これらに限定されないが、ラジオイムノアッセイ(RIA)および酵素結合免疫吸着アッセイ(ELISA)が含まれる。抗原に特異的に結合する抗体または断片は、関連抗原と交差反応する場合がある。好ましく

50

は、抗原に特異的に結合する抗体またはそれらの断片は、他の抗原と交差反応しない。抗原 - 抗体相互作用、特異性および交差反応性に関する考察、ならびに上記のすべてを決定するための方法については、例えば、Paul 編、2003年、Fundamental Immunology、第5版、Raven Press社、ニューヨーク市、69 ~ 105頁（非特許文献 19）を参照されたい。

【0053】

本明細書で使用される場合、「対象」は、動物であり、好ましくは哺乳動物であり、より好ましくは非ヒト霊長類であり、最も好ましくはヒトである。「対象」、「個体」、「候補」および「患者」という用語は、本明細書では同義的に使用される。いくつかの実施形態では、対象は動物である。他の実施形態では、対象は哺乳動物である。

【0054】

本明細書で使用される場合、「試験対象」は、典型的には、判断ルールを構築するために使用されるトレーニング集団に存在しない任意の対象である。試験対象には、随意に、情動障害を罹患している疑いまたは情動障害を発症する可能性を有している疑いがあるもよい。

【0055】

本明細書で使用される場合、「トレーニング集団」とは、データ分析アルゴリズムを使用して、情動障害の罹患リスクを有する対象のバイオマーカー・プロファイルを評価するための判断ルールを構築するために使用される対象の集団に由来する一組の試料である。好ましい実施形態では、トレーニング集団は、情動障害を罹患している対象および情動障害を罹患していない対象に由来する試料を含む。

【0056】

本明細書中で使用される場合、「検証集団」とは、判断ルールの正確度または他の性能計量を決定するために使用される対象の集団に由来する一組の試料である。好ましい実施形態では、検証集団は、情動障害を罹患している対象および情動障害を罹患していない対象に由来する試料を含む。好ましい実施形態では、検証集団は、正確度または他の性能計量が求められる判断ルールをトレーニングするために使用されるトレーニング集団の一部である対象を含んでいない。

【0057】

本明細書で使用される場合、「値セット」とは、バイオマーカー・プロファイル中の特徴についての値の組合せまたは値の範囲である。この値セットおよびその中の値の性質は、バイオマーカー・プロファイル中に存在する特徴のタイプ、および値セットを決定付ける判断ルールを構築するために使用されるデータ分析アルゴリズムに依存する。例示のために、例示的バイオマーカー・プロファイル 2 を再び取り上げる。

【0058】

【表 4】

例示的バイオマーカー・プロファイル 2

バイオマーカー	特徴 相対的単位での試料中の含有量
遺伝子 A の転写物	300
遺伝子 B の転写物	400

【0059】

この例では、トレーニング集団の各メンバーのバイオマーカー・プロファイルが取得される。そのような各バイオマーカー・プロファイルは、測定された特徴、ここでは遺伝子 A の転写物の含有量、および測定された特徴、ここでは遺伝子 B の転写物の含有量を含んでいる。これら特徴値、ここでは含有量値は、判断ルールを構築するためにデータ分析アルゴリズムにより使用される。この例では、データ分析アルゴリズムは、下記に記述されている決定木であり、このデータ分析アルゴリズムの最終産物、判断ルールは、決定木である。この判断ルールは、値セットを定義する。1つのそのような値セットは、情動障害を予測する。バイオマーカー特徴値がこの値セットを満たす対象は、情動障害を罹患して

10

20

30

40

50

いる。この種類の例示的値セットは、例示的値セット 1 である。

【 0 0 6 0 】

【 表 5 】

例示的値セット 1

バイオマーカー	値セット構成要素 (相対的単位での試料中の含有量)
遺伝子 A の転写物	< 400
遺伝子 B の転写物	< 600

【 0 0 6 1 】

別のそのような値セットは、情動障害ではない状態を予測する。バイオマーカー特徴値がこの値セットを満たす対象は、情動障害を罹患しているとは診断されない。この種類の例示的値セットは、例示的値セット 2 である。

【 0 0 6 2 】

【 表 6 】

例示的値セット 2

バイオマーカー	値セット構成要素 (相対的単位での試料中の含有量)
遺伝子 A の転写物	> 400
遺伝子 B の転写物	> 600

【 0 0 6 3 】

データ分析アルゴリズムがニューラル・ネットワーク分析であり、このニューラル・ネットワーク分析の最終産物が、適切に加重されたニューラル・ネットワークである場合、1つの値セットは、対象が情動障害を罹患していることを加重ニューラル・ネットワークに示させることになるバイオマーカー・プロファイル特徴値の範囲である。別の値セットは、対象が情動障害を罹患していないことを加重ニューラル・ネットワークに示させることになるバイオマーカー・プロファイル特徴値の範囲である。

【 0 0 6 4 】

本明細書で使用される場合、マイクロアレイの状況での「プローブ・スポット」という用語は、試料中の特定の核酸の含有量を決定するために使用される、本明細書では「プローブ」と呼ばれる一本鎖 DNA 分子（例えば、一本鎖 cDNA 分子または合成 DNA オリゴマー）を指す。例えば、プローブ・スポットは、試験対象に由来する生体試料（例えば、細胞の収集物）中の mRNA レベルを決定するために使用することができる。特定の実施形態では、典型的なマイクロアレイは、グリッドにある既知位置のスライド・ガラス（または、他の基材）上に配置されている複数のプローブ・スポットを含む。各プローブ・スポットの核酸は、遺伝子の配列または目的の遺伝子の連続した一本鎖部分（例えば、10量体、11量体、12量体、13量体、14量体、15量体、16量体、17量体、18量体、19量体、20量体、21量体、22量体、23量体、24量体、25量体以上）であり、特定の遺伝子または目的の遺伝子によりコードされた mRNA に対するプローブである。各プローブ・スポットは単一の核酸配列を特徴とし、その相補的 DNA 鎖または mRNA 分子にのみそれをハイブリダイズさせる条件下でハイブリダイズされる。そのため、基材上に多くのプローブ・スポットが存在していてもよく、各々が固有の遺伝子または目的の配列を表していてもよい。加えて、2つまたはそれ以上のプローブ・スポットが同じ遺伝子配列を表していてもよい。いくつかの実施形態では、標識核酸試料は、プローブ・スポットにハイブリダイズされ、プローブ・スポットへ特異的にハイブリダイズされた標識核酸の量は、特定の生体試料中のその特定の核酸（例えば、特定の遺伝子の mRNA の転写物）のレベルを決定するために定量化することができる。プローブ、プローブ・スポット、およびマイクロアレイは、Draghici、2003年、Data Analysis Tools for DNA Microarrays、Chapman & Hall/CRC社（非特許文献 15）第2章に一般的に記述されてい

10

20

30

40

50

る。

【0065】

5.2 対象をスクリーニングするための方法

本発明は、個体に由来する生体試料中に、情動障害を罹患している疑いのある試験個体のバイオマーカー・プロファイルの2つまたはそれ以上の特徴を検出することにより、情動障害の正確で迅速な予測および/または診断を可能にする。

【0066】

本発明の特定の実施形態では、情動障害を罹患している疑いのある対象は、本発明の方法を使用してスクリーニングされる。これらの実施形態によると、本発明の方法は、例えば、精神科病棟に入院している対象および/またはある種の精神的トラウマを経験した対象をスクリーニングするために使用することができる。

10

【0067】

特定の実施形態では、例えば血液などの生体試料が採取される。いくつかの実施形態では、生体試料は、血液、脳脊髄液、腹水、間質液、赤血球細胞、白血球細胞、または血小板である。白血球細胞（白血球）には、これらに限定されないが、好中球、好塩基球、好酸球、リンパ球、単球、およびマクロファージが含まれる。いくつかの実施形態では、生体試料は、全血のいくつかの成分である。1つの実施形態では、本発明は、RNA安定化剤または保存剤を含有する使用準備済の収集チューブによる全血試料採取を使用する。このプロトコールは、適切な試料取扱い手順に従えば、ばらつきがほとんどないことが証明および保証されている。本発明は、大量の試料セットのハイスループット分析に使用することができる信頼性の高いロバストな転写マーカを提供する。この信頼性の高い方法は、対照と患者とを区別することが示されている。いくつかの実施形態では、細胞画分内または血液の液体内（例えば、血漿または血清画分）のタンパク質、核酸、および/または他の分子（例えば、代謝産物）の混合物のある部分は、バイオマーカー・プロファイルとして分離される。これは、バイオマーカー・プロファイル中のバイオマーカーの特徴を測定することにより達成することができる。いくつかの実施形態では、生体試料は全血であるが、バイオマーカー・プロファイルは、全血から単離される白血球細胞で発現されるか、またはそうでなければ見出されるバイオマーカーから分離される。いくつかの実施形態では、生体試料は全血であるが、バイオマーカー・プロファイルは、全血から単離される赤血球細胞で発現されるか、またはそうでなければ見出されるバイオマーカーから分離される。

20

30

【0068】

バイオマーカー・プロファイルは、少なくとも2つのバイオマーカーを含み、該バイオマーカーは、例えば、核酸および炭水化物など、同じ種類であってもよく、または異なる種類であってもよい。いくつかの実施形態では、バイオマーカー・プロファイルは、少なくとも2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95、96、100、105、110、115、120、125、130、135、140、145、150、155、160、165、170、175、180、185、190、195、または200個またはそれ以上のバイオマーカーを含んでいる。1つの実施形態では、バイオマーカー・プロファイルは、数百、または数千ものバイオマーカーを含んでいる。いくつかの実施形態では、バイオマーカー・プロファイルは、少なくとも2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、25、30、35、40、45、または50個またはそれ以上のバイオマーカーを含んでいる。一例として、いくつかの実施形態では、バイオマーカー・プロファイルは、表1Aから選択される、少なくとも2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20個またはそれ以上のバイオマーカーを含んでいる。

40

【0069】

典型的な実施形態では、バイオマーカー・プロファイル中の各バイオマーカーは、ある

50

特徴により表される。言いかえれば、バイオマーカーと特徴との間には対応性がある。いくつかの実施形態では、バイオマーカーと特徴との間の対応性は1：1であり、各バイオマーカーに対して、1つの特徴があることを意味している。いくつかの実施形態では、各バイオマーカーには複数の特徴がある。いくつかの実施形態では、バイオマーカー・プロファイル中にある1つのバイオマーカーに対応する特徴の数は、バイオマーカー・プロファイル中にある別のバイオマーカーに対応する特徴の数とは異なっている。そのため、いくつかの実施形態では、バイオマーカー・プロファイル中に、少なくとも2、3、4、5、6、または7つまたはそれ以上のバイオマーカーがある場合、バイオマーカー・プロファイルは、少なくとも2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95、96、100、105、110、115、120、125、130、135、140、145、150、155、160、165、170、175、180、185、190、195、または200個またはそれ以上の特徴を含む場合がある。いくつかの実施形態では、バイオマーカー・プロファイルは、少なくとも2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、25、30、35、40、45、または50個またはそれ以上の特徴を含み得る。実施形態に関わらず、これら特徴は、任意の再現性のある測定技術または測定技術の組合せを使用することにより決定することができる。そのような技術には、本明細書中に記述されている任意の技術、または例えば下記の5.4節に開示されている任意の技術を含む、当技術分野で周知の技術が含まれる。典型的には、そのような技術は、単一の時点で対象から採取された生体試料、または複数の時点で採取された複数の試料を使用して、特徴値を測定するために使用される。1つの実施形態では、対象から採取された試料からバイオマーカー・プロファイルを取得するための例示的な技術は、cDNAマイクロアレイである（例えば、下記の5.4.1.2節を参照）。別の実施形態では、対象から採取された試料からバイオマーカー・プロファイルを取得するための例示的な技術は、BDサイトメトリ・ビーズアレイ（CBA）ヒト炎症キットの取扱説明書（BD Biosciences社）などに記載されている、タンパク質に基づくアッセイまたはタンパク質に基づく技術の他の形態または米国特許第5,981,180号（特許文献7）に記述されているビーズアッセイであり、それらの各々は、それらの全体が、特に生体試料中のタンパク質濃度をアッセイするための種々の方法の教示が、参照により本明細書に組み込まれる。さらに別の実施形態では、バイオマーカー・プロファイルは統合されており、これは、バイオマーカー・プロファイルが、核酸であるいくつかのバイオマーカーまたはそれらの徴候、およびタンパク質であるいくつかのバイオマーカーまたはそれらの兆候を含むことを意味する。そのような実施形態では、タンパク質に基づく技術および核酸に基づく技術を両方とも使用して、対象から採取された1つまたは複数の試料からバイオマーカー・プロファイルを取得する。言いかえれば、核酸であるバイオマーカー・プロファイル中のバイオマーカーと関連する特徴の特徴値は、核酸に基づく測定技術（例えば、核酸マイクロアレイ）により取得され、タンパク質であるバイオマーカー・プロファイル中のバイオマーカーと関連する特徴の特徴値は、タンパク質に基づく測定技術により取得される。いくつかの実施形態では、バイオマーカー・プロファイルは、下記の5.3節に記述されているキットなどのキットを使用して取得することができる。

#### 【0070】

##### 5.3 キット

本発明は、対象の情動障害を診断するのに有用なキットも提供する。いくつかの実施形態では、本発明のキットは、少なくとも2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95、96、100、105、110、115、120、125、130、135、140、145、150、155、160、165、170、175、180、185、190、195、または200個またはそれ以上のバイオマーカー、および/またはそのようなバイオマーカーの存在

もしくは含有量を検出するための試薬を含む。他の実施形態では、本発明のキットは、少なくとも2つのバイオマーカー、しかし数百またはそれ以上ものバイオマーカーを含む。いくつかの実施形態では、本発明のキットは、表1Aから選択される少なくとも2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20個またはそれ以上のバイオマーカー、またはそのようなバイオマーカーの存在または含有量を検出するための試薬を含んでいる。5.1節に示されているバイオマーカーの定義によると、いくつかの場合では、バイオマーカーは、実際、例えば、遺伝子、mRNA、またはタンパク質自体ではなく、遺伝子、mRNA、またはタンパク質の特徴的分子である。したがって、バイオマーカーは、特定の遺伝子、mRNA、またはタンパク質自体ではなく、表1Aで識別されている特定の遺伝子、mRNA、もしくはタンパク質、またはそれらの断片の存在または含有量を示す分子であってもよい。いくつかの実施形態では、本発明のキットは、少なくとも2つバイオマーカー、しかし数百またはそれ以上ものバイオマーカーを含む。いくつかの実施形態では、バイオマーカーおよび/またはバイオマーカーの存在または含有量を検出するための試薬の少なくとも25パーセント、少なくとも30パーセント、少なくとも35パーセント、少なくとも40パーセント、少なくとも60パーセント、少なくとも80パーセントは、表1Aからのバイオマーカー、および/または表1Aから選択されるバイオマーカーの存在または含有量を検出するための試薬から選択される。

10

#### 【0071】

本発明のキットのバイオマーカーは、本発明によるバイオマーカー・プロファイルを生成するために使用することができる。キットの化合物の種類の中には、これらに限定されないが、タンパク質およびそれらの断片、ペプチド、プロテオグリカン、糖タンパク質、リポタンパク質、炭水化物、脂質、核酸（例えば、cDNAまたは増幅されたDNAなどのDNA、またはmRNAなどのRNA）、有機または無機の化学薬品、天然または合成のポリマー、低分子（例えば、代謝産物）、または先述のものの中のいずれかの特徴的分子もしくは特徴的断片が含まれる。特定の実施形態では、バイオマーカーは、特定のサイズである（例えば、少なくとも10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95、100、105、110、115、120、125、130、135、140、145、150、155、160、165、170、175、180、185、190、195、200、1000、2000、3000、5000、10k、20k、または100kダルトンまたはそれ以上）。バイオマーカー（複数可）は、アレイの一部であってもよく、またはバイオマーカー（複数可）は、別々におよび/または個々に包装されていてもよい。またキットは、本発明のバイオマーカー・プロファイルを生成するために使用される少なくとも1つの内部標準を含んでいてもよい。同様に、内部標準または標準は、上記に記述されている種類の化合物のいずれであってもよい。

20

30

#### 【0072】

1つの実施形態では、本発明は、例えばマイクロアレイに見い出されるような、基材上のアドレス可能な位置に固定されていてもよくまたは固定されていなくてもよいプローブおよび/またはプライマーを含むキットを提供する。特定の実施形態では、本発明は、そのようなマイクロアレイを提供する。

40

#### 【0073】

本発明のいくつかの実施形態では、キットは、アプタマーなどの特異的バイオマーカー結合成分を含んでいてもよい。バイオマーカーが核酸を含む場合、キットは、バイオマーカーまたはバイオマーカーの相補鎖と二本鎖を形成可能なオリゴヌクレオチド・プローブを提供することができる。オリゴヌクレオチド・プローブは、検出可能に標識されていてもよい。そのような実施形態では、プローブはそれら自体が、本発明の範囲内にあるバイオマーカーである。

#### 【0074】

本発明のキットは、バイオマーカー・プロファイルを構築するのに使用することができ

50

る緩衝液などの追加的な組成物も含んでいてよい。微生物作用の予防は、種々の抗菌性および抗真菌物質、例えばパラベン、クロロブタノール、およびフェノール・ソルビン酸などを含有させることにより保証することができる。糖および塩化ナトリウムなどの等張剤を含むことが望ましい場合もある。

#### 【0075】

本発明のいくつかのキットは、マイクロアレイを含んでいる。1つの実施形態では、このマイクロアレイは、複数のプローブ・スポットを含み、複数のプローブ・スポットにある少なくとも20パーセントのプローブ・スポットは、表1Aのバイオマーカーに対応する。いくつかの実施形態では、複数のプローブ・スポットにあるプローブ・スポットの少なくとも25パーセント、少なくとも30パーセント、少なくとも35パーセント、少なくとも40パーセント、少なくとも60パーセント、または少なくとも80パーセントは、表1Aのバイオマーカーおよび/または表1Aのバイオマーカーの存在または含有量を検出するための試薬に対応する。そのようなプローブ・スポットは、本発明の範囲内のバイオマーカーである。いくつかの実施形態では、マイクロアレイは、基材上に約2～約100個のプローブ・スポットを含む。いくつかの実施形態では、マイクロアレイは、基材上に約2～約100個のプローブ・スポットを含む。本状況で使用される場合、「約」という用語は、記載値の5パーセント以内、記載値の10パーセント内、または記載値の25パーセント以内を意味する。いくつかの実施形態では、そのようなマイクロアレイは、当業者に公知の技術を使用するマイクロアレイ間較正または基準マイクロアレイなどの他のマイクロアレイによる較正のための1つまたは複数のプローブ・スポットを含有している。いくつかの実施形態では、そのようなマイクロアレイは、核酸マイクロアレイである。いくつかの実施形態では、そのようなマイクロアレイは、タンパク質マイクロアレイである。

10

20

#### 【0076】

本発明のいくつかのキットは、コンピュータ読み取り可能な記憶媒体に内蔵されているコンピュータ・プログラム機序を含むコンピュータ・プログラム製品として実装される。さらに、本発明の方法はいずれも、1つまたは複数のコンピュータまたは他の形態の装置に実装することができる。装置の例には、これらに限定されないが、コンピュータおよび分光法測定デバイス（例えば、マイクロアレイ・リーダまたはマイクロアレイ・スキャナー）が含まれる。さらになお、本発明の方法はいずれも、1つまたは複数のコンピュータ・プログラム製品に実装することができる。本発明のいくつかの実施形態は、本明細書で開示されている方法のいずれかまたはすべてをコードするコンピュータ・プログラム製品を提供する。そのような方法は、CD-ROM、DVD、磁気ディスク記憶媒体製品、または他の任意の実在するコンピュータ読み取り可能なデータもしくは実在するプログラム記憶媒体製品に格納することができる。そのような方法は、ROM、1つまたは複数のプログラム可能なチップ、または1つもしくは複数の特定用途向け集積回路（ASIC）などの永久記録媒体に内蔵することができる。そのような永久記録媒体は、サーバ、802.11アクセスポイント、802.11無線ブリッジ/ステーション、中継器、ルーター、携帯電話、または他の電子機器に設置されていてもよい。コンピュータ・プログラム製品にコードされたそのような方法は、インターネットを介してまたは別様に、電子的に配布することもできる。

30

40

#### 【0077】

本発明のいくつかのキットは、本発明の方法のいずれかを個々にまたは集団的に実行する1つまたは複数のプログラムを含有するコンピュータ・プログラム製品を提供する。これらプログラム・モジュールは、CD-ROM、DVD、磁気ディスク記憶媒体製品、または他の任意の実在するコンピュータ読み取り可能なデータもしくは実在するプログラム記憶媒体製品に格納することができる。プログラム・モジュールは、ROM、1つまたは複数のプログラム可能なチップ、または1つもしくは複数の特定用途向け集積回路（ASIC）などの永久記録媒体に内蔵することもできる。そのような永久記録媒体は、サーバ、802.11アクセスポイント、802.11無線ブリッジ/ステーション、中継器、

50

ルーター、携帯電話、または他の電子機器に設置されていてもよい。コンピュータ・プログラム製品のソフトウェア・モジュールは、インターネットを介してまたは別様に、電子的に配布することもできる。

【0078】

本発明のいくつかのキットは、1つまたは複数の処理装置を有するコンピュータ、および1つまたは複数の処理装置に接続されたメモリを含む。メモリは、情動障害罹患リスクのある試験対象のバイオマーカー・プロファイル中の複数の特徴が、値セットを満たしているかどうかを評価するための命令群を格納する。いくつかの実施形態では、値セットを満たすことにより、対象が情動障害を罹患していると診断する。いくつかの実施形態では、値セットを満たすことにより、対象が情動障害を罹患していないと診断される。1つの実施形態では、複数の特徴は、表1Aに列挙されているバイオマーカーに対応する。

10

【0079】

図1では、上記に記述されている機能性を支援する例示的なシステムが詳述されている。本システムは、好ましくは、以下のものを有するコンピュータ・システム10である：

- ・ 中央処理装置22；
- ・ メイン揮発性記憶媒体ユニット14、例えばソフトウェアおよびデータを格納するためのハードディスク・ドライブ、記憶媒体コントローラ12により制御される記憶媒体ユニット14；
- ・ 揮発性記憶媒体ユニット14から読み込んだプログラムおよびデータを含むシステム・メモリ36、好ましくは、システム・コントロール・プログラム、データ、およびアプリケーション・プログラムを格納するための高速ランダム・アクセス・メモリ(RAM)；システム・メモリ36は、読み出し専用メモリ(ROM)を含むこともできる；
- ・ 1つまたは複数の入力デバイス(例えば、キーボード28)およびディスプレイ26または他の出力デバイスを含むユーザー・インターフェース32；
- ・ 任意の有線または無線通信ネットワーク34(例えば、インターネットなどの広域ネットワーク)に接続するためのネットワーク・インターフェース・カード20；
- ・ 本システムの前述の要素を相互接続するための内部バス30；および
- ・ 前述の要素に動力を供給する電源24。

20

【0080】

コンピュータ10の作働は、主として、中央処理装置22により実行されるオペレーティング・システム40により制御される。オペレーティング・システム40は、システム・メモリ36に格納することができる。オペレーティング・システム40に加えて、典型的な実装では、システム・メモリ36は以下のものを含む：

30

- ・ 本発明により使用される種々のファイルおよびデータ構造へのアクセスを制御するためのファイル・システム42；
- ・ 本発明による1つまたは複数の判断ルールの構築に使用されるトレーニング・データ・セット44；
- ・ トレーニング・データを処理し、判断ルールを構築するためのデータ分析アルゴリズム・モジュール54；
- ・ 1つまたは複数の判断ルール56；
- ・ 試験対象のバイオマーカー・プロファイル中の複数の特徴が、第1の値セットまたは第2の値セットを満たすかどうかを決定するためのバイオマーカー・プロファイル評価モジュール60；
- ・ バイオマーカー64およびそのようなバイオマーカーの各々についての特徴66を含む試験対象バイオマーカー・プロファイル62；および
- ・ 本発明の選択バイオマーカー(例えば、表1A)、および/またはそれら選択バイオマーカーの各々についての1つまたは複数の特徴に関するデータベース68。

40

【0081】

トレーニング・データセット46は、複数の対象46のデータを含む。各対象46ごとに、対象識別子48および複数のバイオマーカー50が存在する。各バイオマーカー50

50

ごとに、少なくとも1つの特徴52が存在する。図1には示されていないが、各特徴52ごとに、特徴値が存在する。データ分析アルゴリズムを使用して構築された各判断ルール56ごとに、少なくとも1つの判断ルール値セット58が存在する。

#### 【0082】

図1に例示されているように、コンピュータ10は、ソフトウェア・プログラム・モジュールおよびデータ構造を含む。コンピュータ10に格納されるデータ構造は、トレーニング・データセット44、判断ルール56、試験対象バイオマーカー・プロファイル62、およびバイオマーカー・データベース68を含む。これらデータ構造の各々は、これらに限定されないが、単純なASCIIもしくはバイナリー・ファイル、Excelスプレッドシート、リレーショナル・データベース(SQL)、またはオンライン分析処理(OLAP)データベース(MDXおよび/またはその変法)を含む任意の形態のデータ保存システムを含んでいてもよい。いくつかの特定の実施形態では、そのようなデータ構造は各々、階層構造を含む1つまたは複数のデータベースの形態である(例えば、starスキーマ(star schema))。いくつかの実施形態では、そのようなデータ構造は各々、明示的な階層を有していないデータベースの形態である(例えば、階層的に配置されていない次元表(dimension table))。

10

#### 【0083】

いくつかの実施形態では、システム10に格納されているデータ構造またはシステム10にアクセス可能なデータ構造の各々は、単一のデータ構造である。他の実施形態では、そのようなデータ構造は、実際に、同じコンピュータ10によりホストされていてもよく、またはすべてがホストされていなくてもよい複数のデータ構造(例えば、データベース、ファイル、アーカイブ)を含む。例えば、いくつかの実施形態では、トレーニング・データセット44は、コンピュータ10および/または広域ネットワーク34上にある、コンピュータ10によりアドレス可能なコンピュータのいずれかに格納されている複数のExcelスプレッドシートを含む。別の例では、トレーニング・データセット44は、コンピュータ10に格納されているか、または広域ネットワーク34上にある、コンピュータ10によりアドレス可能な1つまたは複数のコンピュータにわたって分配されているデータベースを含む。

20

#### 【0084】

図1に例示されている多くのモジュールおよびデータ構造は、1つまたは複数のリモート・コンピュータに配置することができることが理解されるだろう。例えば、本出願のいくつかの実施形態は、ウェブ・サービス型の実装である。そのような実施形態では、バイオマーカー・プロファイル評価モジュール60および/または他のモジュールは、ネットワーク34を介してコンピュータ10と通信するクライアント・コンピュータに存在していてもよい。いくつかの実施形態では、例えば、バイオマーカー・プロファイル評価モジュール60は、対話型のウェブ・ページであってもよい。

30

#### 【0085】

いくつかの実施形態では、図1に例示されているトレーニング・データセット44、判断ルール56、および/またはバイオマーカー・データベース68は、単一のコンピュータ(コンピュータ10)に存在しており、他の実施形態では、そのようなデータ構造およびモジュールの1つまたは複数は、1つまたは複数のリモート・コンピュータ(非表示)によりホストされている。1つまたは複数のコンピュータ上の図1に例示されているデータ構造およびソフトウェア・モジュールのあらゆる配置は、これらデータ構造およびソフトウェア・モジュールがネットワーク34を介してまたは他の電子的手段により互いにアドレス可能である限り、本発明の範囲内にある。したがって、本発明は、多岐にわたるコンピュータ・システムを完全に包含する。

40

#### 【0086】

さらに、本発明の別の実施形態は、対象が情動障害を罹患しているかどうかを決定するためのグラフィカル・ユーザー・インターフェースを提供する。グラフィカル・ユーザー・インターフェースは、リモート・コンピュータから受信した搬送波に埋め込まれている

50

デジタル信号にコードされている結果を表示するための表示領域を含む。複数の特徴は、複数のバイオマーカーの測定可能な態様である。複数のバイオマーカーは、表 1 A に列挙されている少なくとも 2 つのバイオマーカーを含む。試験対象のバイオマーカー・プロファイル中の複数の特徴が、第 1 の値セットを満たす場合、結果は第 1 の値を有する。試験対象のバイオマーカー・プロファイル中の複数の特徴が、第 2 の値セットを満たす場合、結果は第 2 の値を有する。

#### 【 0 0 8 7 】

##### 5 . 4 バイオマーカー・プロファイルの生成

1 つの実施形態によると、本発明の方法は、対象から採取した生体試料からバイオマーカー・プロファイルを生成することを含む。生体試料は、例えば、末梢組織、全血、脳脊髄液、腹水、間質液、赤血球細胞、白血球細胞、または血小板であってもよい。

10

#### 【 0 0 8 8 】

##### 5 . 4 . 1 核酸バイオマーカーを検出する方法

本発明の特定の実施形態では、バイオマーカー・プロファイル中のバイオマーカーは核酸である。そのようなバイオマーカーおよびバイオマーカー・プロファイルの対応する特徴は、例えば、本明細書に記述されている 1 つまたは複数の遺伝子（例えば、表 1 A に列挙されている遺伝子）の発現産物（例えば、ポリヌクレオチドまたはポリペプチド）を検出することにより生成することができる。特定の実施形態では、バイオマーカー・プロファイル中のバイオマーカーおよび対応する特徴は、本明細書に開示されている遺伝子（例えば、表 1 A に列挙されている遺伝子）から発現される 1 つまたは複数の核酸を、決して限定ではないが、ハイブリダイゼーション、マイクロアレイ分析、RT - PCR、ヌクレアーゼ保護アッセイ、およびノーザンブロット分析を含む当業者に周知である任意の方法を使用して、検出および / または分析することにより取得される。

20

#### 【 0 0 8 9 】

ある実施形態では、本発明の方法および組成物により検出および / または分析される核酸には、例えば、メッセンジャー RNA ( mRNA ) 分子、mRNA スプライス・バリエーションを含む発現された RNA 分子、および調節 RNA、cRNA 分子（例えば、*in vitro* で転写される cDNA 分子から調製される RNA 分子）などの RNA 分子、ならびにそれらの特徴的断片が含まれる。本発明の方法および組成物により検出および / または分析される核酸には、例えば、ゲノム DNA 分子、cDNA 分子などの DNA 分子およびそれらの特徴的断片（例えば、オリゴヌクレオチド、EST、STS など）が含まれていてもよい。

30

#### 【 0 0 9 0 】

本発明の方法および組成物により検出および / または分析される核酸分子は、試料から単離されるゲノム DNA 分子またはゲノム外 DNA 分子などの天然に存在する核酸分子、または生体試料に存在するか、生体試料から分離されるか、もしくは生体試料に由来する mRNA 分子などの RNA 分子であってもよい。本発明の方法および組成物により検出および / または分析される核酸の試料は、例えば、DNA、RNA、または DNA および RNA のコポリマーの分子を含む。一般的に、これら核酸は、特定の遺伝子または遺伝子の対立遺伝子に、または特定の遺伝子転写物に（例えば、特定の細胞タイプで発現される特定の mRNA 配列に、またはそのような mRNA 配列に由来する特定の cDNA 配列に）対応する。本発明の方法および組成物により検出および / または分析される核酸は、同じ遺伝子の異なるエキソンに対応してもよく、例えば、その結果、その遺伝子の異なるスプライス・バリエーションを検出および / または分析することができる。

40

#### 【 0 0 9 1 】

特定の実施形態では、核酸は、生物学的試料に存在するか、または生物学的試料から分離されるか、または生物学的試料から部分的に分離される核酸から *in vitro* で調製される。例えば、1 つの実施形態では、RNA は、試料（例えば、全細胞 RNA、ポリ ( A ) + メッセンジャー RNA、それらの画分）から抽出され、メッセンジャー RNA は、全抽出 RNA から精製される。全 RNA およびポリ ( A ) + RNA を調製するための方

50

法は、当技術分野で周知であり、例えば、Sambrookら、2001年、Molecular Cloning: A Laboratory Manual. 第3版、Cold Spring Harbor Laboratory Press (コールド・スプリング・ハーバー、ニューヨーク) (非特許文献20) で一般的に記述されている。

【0092】

5.4.1.1 核酸アレイ

本発明のある実施形態では、核酸アレイを使用して、本明細書に記述されている遺伝子 (例えば、表1Aに列挙されている遺伝子) の任意の1つまたは複数の発現を検出することにより、バイオマーカー・プロファイル中のバイオマーカーの特徴を生成する。本発明の1つの実施形態では、cDNAマイクロアレイなどのマイクロアレイを使用して、バイオマーカー・プロファイル中のバイオマーカーの特徴値を決定する。cDNAアレイの診断使用は、当技術分野で周知である。(例えば、Zouら、2002年、Oncogene 21巻:4855~4862頁(非特許文献21); ならびにDraghici、2003年、Data Analysis Tools for DNA Microarrays、Chapman & Hall/CRC社(非特許文献15)を参照)。cDNAマイクロアレイ分析の例示的な方法は、下記に記述されている。

【0093】

ある実施形態では、バイオマーカー・プロファイル中のバイオマーカーの特徴値は、生体試料に存在するmRNA転写物の核酸配列を表すかまたは対応する、1つまたは複数のプローブ・スポットを含むマイクロアレイのアレイ検出可能標識核酸(例えば、試料から合成された蛍光標識cDNA)にハイブリダイズすることにより取得される。

【0094】

核酸アレイ、例えばマイクロアレイは、多くの方法で製作することができ、それらのいくつかは本明細書で下記に記述されている。好ましくは、アレイは再現性が高く、所与のアレイの複数の複製を製作することが可能であり、前記マイクロアレイからの結果は互いに比較可能である。好ましくは、アレイは、結合(例えば、核酸ハイブリダイゼーション)条件下で安定的な物質から製作される。当業者であれば、アレイ上のプローブ・スポットに試験プローブをハイブリダイズさせるのに好適な支持物、基材、または担体について知るだろうし、または日常的な実験作業を使用することによりそれを確認することができるだろう。

【0095】

使用されるアレイ、例えばマイクロアレイは、1つまたは複数の試験プローブを含むことができる。いくつかの実施形態では、そのような試験プローブの各々は、検出されるRNAまたはDNAのサブ配列に相補的な核酸配列を含む。各プローブは、典型的には異なる核酸配列を有しており、アレイの固体表面上の各プローブの位置は、通常は既知であるか、または決定することができる。本発明により有用なアレイには、例えば、オリゴヌクレオチド・マイクロアレイ、cDNAに基づくアレイ、SNPアレイ、スプライス・バリエーション・アレイ、および本明細書に記述されている遺伝子(例えば、表1Aに列挙されている遺伝子)発現の定性的、定量的、または半定量的測定を提供することができる任意の他のアレイが含まれていてもよい。いくつかのタイプのマイクロアレイは、アドレス可能なアレイである。より具体的には、いくつかのマイクロアレイは、位置的にアドレス可能なアレイである。いくつかの実施形態では、アレイの各プローブは、固体支持体上の既知の所定位置に位置しており、そのため、各プローブの同一性(例えば配列)は、アレイ上(例えば、支持体上または表面上)のその位置から決定することができる。いくつかの実施形態では、アレイは、特注アレイである。マイクロアレイは、Draghici、2003年、Data Analysis Tools for DNA Microarrays、Chapman & Hall/CRC社(非特許文献15)で一般的に記述されている。

【0096】

本発明のいくつかの実施形態では、発現された転写物(例えば、本明細書に記述されている遺伝子の転写物)は、核酸アレイで表される。そのような実施形態では、一組の結合部位は、発現された転写物の様々な配列セグメントに相補的である様々な核酸を有するプローブを含むことができる。この種類に入る例示的な核酸は、長さが、15~200個の

塩基、20～100個の塩基、25～50個の塩基、40～60個の塩基、または他のいくつかの範囲の塩基であってもよい。各プローブ配列は、その標的配列に相補的な配列に加えて、1つまたは複数のリンカー配列も含むことができる。本明細書で使用される場合、リンカー配列とは、その標的配列に相補的な配列と支持体表面との間の配列である。例えば、本発明の核酸アレイは、各標的遺伝子またはエキソンに特異的な1つのプローブを含むことができる。しかしながら、必要に応じて、核酸アレイは、いくつかの発現された転写物（例えば、本明細書、例えば表1Aに記述されている遺伝子の転写物）に特異的な、少なくとも2、5、10、100、または1000個またはそれ以上のプローブを含有することができる。例えば、アレイは、遺伝子の最長mRNAアイソフォームの配列全体にわたって結合するプローブを含有していてもよい。

10

**【0097】**

細胞、例えば生体試料中の細胞のRNAに相補的なcDNAが製作され、好適なハイブリダイゼーション条件下でマイクロアレイにハイブリダイズされる場合、本明細書で記述されている遺伝子（例えば、表1Aに列挙されている遺伝子）に対応するアレイ部位に対するハイブリダイゼーションのレベルは、その遺伝子から転写された1つまたは複数のmRNAの、細胞における出現率を反映することになることが理解されるだろう。あるいは、特定の遺伝子により産生される多数のアイソフォームまたは選択的スプライス・バリエーションを識別しようとする場合、全細胞mRNAに相補的な検出可能標識（例えば、フルオロフォア）cDNAは、マイクロアレイにハイブリダイズすることができ、細胞において転写されないかまたはRNAスプライシング中に除去される遺伝子のエキソンに対応するアレイ上の部位は、シグナル（例えば、蛍光シグナル）をほとんどまたはまったく示さず、エキソンを発現するコードされたmRNAが出現する遺伝子のエキソンに対応する部位は比較的強いシグナルを示す。その後、選択的スプライシングにより同一遺伝子から産生される様々なmRNAの相対的存在量は、その遺伝子についてモニターされるエキソンのセット全体にわたる信号シグナルのパターンにより決定される。

20

**【0098】**

1つの実施形態では、異なるハイブリダイゼーション時間でのハイブリダイゼーション・レベルが、別々の同一マイクロアレイで別々に測定される。そのような測定の各々について、ハイブリダイゼーション・レベルが測定されるハイブリダイゼーション時間で、好ましくは、結合またはハイブリダイズされた核酸はすべて保持されるが、未結合核酸はすべて除去される条件化で、高濃度から中程度濃度の塩濃度（例えば、0.5～3Mの塩濃度）の水溶液中、室温で、マイクロアレイを短時間洗浄する。その後、各プローブ上の残留ハイブリダイズ核酸分子の検出可能な標識を、使用されている特定の標識法に適切な方法により測定する。その後、その結果生じたハイブリダイゼーション・レベルを組み合わせ、ハイブリダイゼーション曲線を形成する。別の実施形態では、ハイブリダイゼーション・レベルは、単一のマイクロアレイを使用してリアルタイムで測定される。この実施形態では、マイクロアレイを継続的に試料とハイブリダイズさせ、非侵襲性の様式でマイクロアレイを各ハイブリダイゼーション時間で検査する。さらに別の実施形態では、1つのアレイを使用し、短時間ハイブリダイズさせ、洗浄し、ハイブリダイゼーション・レベルを測定し、同じ試料に戻し、さらにある期間ハイブリダイズさせ、洗浄し、再び測定して、ハイブリダイゼーション時間曲線を得ることができる。

30

40

**【0099】**

いくつかの実施形態では、核酸ハイブリダイゼーション条件および洗浄条件は、分析される核酸バイオマーカーが、典型的にはその相補的DNAが位置している特定のアレイ部位で、アレイの相補的核酸配列と特異的に結合または特異的にハイブリダイズするように選択される。

**【0100】**

そこに位置している二本鎖プローブDNAを含有するアレイを、標的核酸分子と接触させる前に、DNAを一本鎖にする変性条件に供してもよい。一本鎖プローブDNA（例えば、合成オリゴデオキシリボ核酸）を含有するアレイは、標的核酸分子と接触させる前に

50

、例えば自己相補的配列により形成されるヘアピンまたは二量体を除去するために、変性させることが必要な場合がある。

#### 【0101】

最適なハイブリダイゼーション条件は、プローブおよび標的核酸の長さ（例えば、200塩基を超えるオリゴマー対ポリヌクレオチド）およびタイプ（例えば、RNAまたはDNA）に依存するだろう。核酸に対する特異的な（つまり、ストリンジェントな）ハイブリダイゼーション条件の一般的なパラメーターは、Sambrookら（上記）（非特許文献20）およびAusubelら、最新版、Current Protocols in Molecular Biology、Greene Publishing and Wiley-Interscience社、ニューヨーク市（非特許文献22）に記述されている。ShenaらのcDNAマイクロアレイが使用される場合、典型的なハイブリダイゼーション条件は、0.2% SDSを加えた5×SSC中65℃で4時間のハイブリダイゼーション、その後低ストリンジェンシー洗浄緩衝液（0.2% SDSを加えた1×SSC）中25℃での洗浄、その後より高ストリンジェンシー洗浄緩衝液（0.2% SDSを加えた0.1×SSC）中25℃で10分間である（Shenaら、1996年、Proc. Natl. Acad. Sci. U.S.A. 93巻：10614頁（非特許文献23））。有用なハイブリダイゼーション条件は、例えば、Tijessen、1993年、Hybridization With Nucleic Acid Probes、Elsevier Science Publishers B.V.社（非特許文献24）；Kricka、1992年、Nonisotopic DNA Probe Techniques、Academic Press社、サンディエゴ市、カリフォルニア州（非特許文献25）；およびZouら、2002年、Oncogene 21巻：4855～4862頁（非特許文献21）；およびDraghici、Data Analysis Tools for DNA Microanalysis、2003年、CRC Press LLC社、ボカラトーン、フロリダ州、342～343頁（非特許文献26）にも提供されている。

10

20

#### 【0102】

特定の実施形態では、マイクロアレイは、例えば、下記の5.4.1.2節に記述されている方法により生成されたRT-PCR産物を分類するために使用することができる。

#### 【0103】

##### 5.4.1.2 RT-PCR

ある実施形態では、本発明のバイオマーカー・プロファイル中のバイオマーカーの特徴値を決定するために、本明細書に記述されている遺伝子（例えば、表1Aに列挙されている遺伝子）の1つまたは複数の発現レベルを、逆転写（RT）をポリメラーゼ連鎖反応（PCR）と組み合わせて使用して、試料からRNAを増幅することにより測定する。この実施形態によると、逆転写は、定量的であってもよく、または半定量的であってもよい。本明細書で教示されているRT-PCR法は、上記で、例えば5.4.1.1節に記述されているマイクロアレイ法と共に使用することができる。例えば、バルクPCR反応を実施し、PCR産物を分離し、マイクロアレイのプローブ・スポットとして使用することができる。

30

#### 【0104】

試料に由来する全RNAまたはmRNAをテンプレートとして使用し、遺伝子（複数可）の転写部分に特異的なプライマーを、逆転写を開始するために使用する。RNAをcDNAに逆転写する方法は周知であり、Sambrookら、2001年、上記（非特許文献20）に記述されている。プライマー設計は、公開されているか、またはGenBankなどの任意の公的に利用可能な配列データベースから入手可能である公知のヌクレオチド配列に基づいて達成することができる。例えば、プライマーは、本明細書に記述されている遺伝子（例えば、表1A中を参照）のいずれのためにも設計することができる。さらに、プライマー設計は、市販のソフトウェア（例えば、Primer Designer 1.0、Scientific Software社など）を使用することにより達成することができる。逆転写産物は、その後PCR用のテンプレートとして使用される。

40

#### 【0105】

PCRは、目的の標的配列を増幅する耐熱性のDNA依存性DNAポリメラーゼにより触媒される複数サイクルのDNA複製を使用することによって、迅速に特定の核酸配列を増幅するための方法を提供する。PCRには、増幅しようとする核酸、増幅しようとする

50

配列を隣接する2つの一本鎖オリゴヌクレオチド・プライマー、DNAポリメラーゼ、デオキシリボヌクレオシド三リン酸、緩衝液、および塩の存在が必要である。PCR法は当技術分野で周知である。PCR、例えば、MullisおよびFaloon、1987年、Methods Enzymol. 155巻：335頁（非特許文献27）に記述されているように実施される。

#### 【0106】

PCRは、テンプレートDNAまたはcDNA（少なくとも1fg；1～1000ngがより有用である）および少なくとも25pmolのオリゴヌクレオチド・プライマーを使用して実施することができる。典型的な反応混合物は、以下のものを含んでいる：2μlのDNA、25pmolのオリゴヌクレオチド・プライマー、2.5μlの10M PCR緩衝液1（Perkin Elmer社製、フォスターシティ、カリフォルニア州）、0.4μlの1.25M dNTP、0.15μl（または2.5ユニット）のTaq DNAポリメラーゼ（Perkin Elmer社製、フォスターシティ、カリフォルニア州）、および全容積を25μlにするための脱イオン水。鉱油で被覆し、プログラム可能なサーマル・サイクラーを使用してPCRを実施する。

#### 【0107】

PCRサイクルの各ステップの長さおよび温度ならびにサイクル数は、実際に必要とされるストリンジェンシーに従って調整される。アニーリング温度およびタイミングは、プライマーがテンプレートにアニーリングすると予測される効率、および許容することができる mismatches の程度の両方により決定される。プライマー・アニーリング条件のストリンジェンシーを最適化する能力は、十分に当業者の知識内にある。30～72の間のアニーリング温度が使用される。テンプレート分子の初期変性は、通常、92～99、4分間で生じ、その後、変性（94～99で15秒～1分間）、アニーリング（上記で考察されているようにして決定された温度；1～2分間）、および伸張（72で1分間）からなる20～40サイクルが続く。最終伸張ステップは、一般的に72で4分間実行され、その後無期限（0～24時間）4のステップを続けてもよい。

#### 【0108】

定量RT-PCR（「QRT-PCR」）は、本来定量的であり、遺伝子発現レベルの定量的測定を提供するために実施することもできる。QRT-PCRでは、逆転写およびPCRを2段階で実施してもよく、PCRと組み合わせた逆転写を同時に実施してもよい。Taqman（Perkin Elmer社製、フォスターシティ、カリフォルニア州）またはApplied Biosystems社（フォスターシティ、カリフォルニア州）により提供されているものなどの市販キットの技術のうちの1つが、転写物特異的アンチセンス・プローブを用いて実施される。このプローブは、PCR産物（例えば、遺伝子由来の核酸断片）に特異的であり、消光性および蛍光性リポーター・プローブをオリゴヌクレオチドの5'末端に複合体化させて調製される。異なる蛍光マーカ―は、異なるリポーターに結合されており、1つの反応で2つの産物の測定を可能にする。Taq DNAポリメラーゼは、活性化されると、その5'-3'エキソヌクレアーゼ活性により、テンプレートに結合したプローブの蛍光リポーターを引き剥がす。消光体は今や存在せず、リポーターは蛍光を発する。リポーターの色変化は、各特異的産物の量に比例しており、蛍光光度計で測定される。このように、各色の量が測定され、PCR産物が定量化される。PCR反応は、多数の個体に由来する試料が同時に処理および測定されるように、96ウェルプレートで実施される。Taqman系は、ゲル電気泳動を必要としないという追加的な利点を有しており、標準曲線と共に使用される場合、定量化を可能にする。

#### 【0109】

PCR産物を定量的に検出するのに有用な第2の技術は、市販のQuantitect SYBR Green PCR（Qiagen社、バレンシア、カリフォルニア州）などの挿入染料（intercalating dye）を使用することである。RT-PCRは、PCR段階中にPCR産物に組み込まれ、PCR産物の量に比例した蛍光をもたらすSYBR Greenを蛍光標識として使用して実施される。

#### 【0110】

10

20

30

40

50

TaqmanおよびQuantitect SYBR系は両方とも、RNAの逆転写後に使用してもよい。逆転写は、PCRステップと同じ反応混合物中で実施されてもよく（一段階プロトコール）、または逆転写は、PCRを使用して増幅する前に最初に実施（二段階プロトコール）してもいずれでもよい。

【0111】

加えて、蛍光分子および消光分子を有するプローブを使用するMOLECULAR BEACONS（登録商標）を含む、mRNA発現産物を定量的に測定する他の系が知られており、このプローブは、ヘアピン構造を形成可能であり、ヘアピン形態時には蛍光分子が消光され、ハイブリダイズ時には蛍光増加が、遺伝子発現の定量的測定をもたらす。

【0112】

RNA発現を定量的に測定する追加的な技術には、これらに限定されないが、ポリメラーゼ連鎖反応、リガーゼ連鎖反応、Qベータレプリカーゼ（例えば、国際出願第PCT/US87/00880号（特許文献8）を参照）、等温増幅法（例えば、Walkerら、1992年、PNAS 89巻：382～396頁（非特許文献28）を参照）、鎖置換増幅（SDA）、修復連鎖反応（repair chain reaction）、非対称定量PCR（例えば、米国特許出願公開第2003/30134307号A1（特許文献9））、およびFujaら、2004年、Journal of Biotechnology 108巻：193～205頁（非特許文献29）に記述されている多重マイクロスフェアビーズアッセイ（multiplex microsphere bead assay）が含まれる。

【0113】

5.4.2 タンパク質を検出する方法

本発明の特定の実施形態では、バイオマーカー・プロファイル中のバイオマーカーの特徴値は、タンパク質を検出することにより、例えば、1つまたは複数の本明細書に記述されている遺伝子（例えば表1Aに列挙されている遺伝子）の発現産物（例えば、核酸またはタンパク質）、またはそのようなタンパク質の翻訳後修飾された形態もしくは別様に修飾またはプロセッシングされた形態を検出することにより取得することができる。特定の実施形態では、バイオマーカー・プロファイルは、これらに限定されないが、タンパク質マイクロアレイ分析法、免疫組織化学法、質量分析法を含む当業者に公知である任意のタンパク質検出法を使用して、本明細書で開示されている遺伝子（例えば、表1Aに列挙されている遺伝子）から発現される1つまたは複数のタンパク質および/またはその特徴的断片を検出および/または分析することにより生成される。

【0114】

標準的技術を使用して、試料中に存在する1つまたは複数の目的タンパク質（例えば、表1Aに列挙されている遺伝子から発現されるタンパク質）の量を決定することができる。例えば、例えばウエスタンブロット、免疫沈降後のドデシル硫酸ナトリウムポリアクリルアミドゲル電気泳動法（SDS-PAGE）、および免疫細胞化学法などの、例えばイムノアッセイを使用する標準的技術を使用して、試料中に存在する1つまたは複数の目的タンパク質の量を決定することができる。目的タンパク質を検出するための1つの例示的な作用剤は、目的タンパク質に特異的に結合可能な抗体、好ましくは直接的または間接的のいずれかで検出可能に標識された抗体である。

【0115】

そのような検出法の場合、必要に応じて、分析される試料由来のタンパク質は、当業者に周知の技術を使用して容易に単離することができる。タンパク質単離法は、例えば、HarlowおよびLane、1988年、Antibodies: A Laboratory Manual、Cold Spring Harbor Laboratory Press社（コールド・スプリング・ハーバー、ニューヨーク市）（非特許文献30）に記述されているものなどであり得る。

【0116】

5.5 データ分析アルゴリズム

その対応する特徴値が情動障害を診断可能であるバイオマーカーが、本発明で特定される。これらバイオマーカーの同一性およびそれらの対応する特徴（例えば、発現レベル）を使用して、情動障害を罹患している対象と罹患していない対象とを区別する1つの判断

10

20

30

40

50

ルールまたは複数の判断ルールを開発することができる。これら例示的なデータ分析アルゴリズムまたは当技術分野で公知の他の技術を使用して判断ルールが構築されたならば、その判断ルールを使用して、試験対象を複数の表現型クラス（例えば、情動障害を罹患している、情動障害を罹患していない）のうちの1つに分類することができる。これは、試験対象から得られたバイオマーカー・プロファイルに、判断ルールを適用することにより達成される。したがって、そのような判断ルールは、診断指標として非常に大きな価値を有する。

**【0117】**

本発明は、1つの態様の中では、トレーニング集団から得られるバイオマーカー・プロファイルに対して、試験対象からのバイオマーカー・プロファイルを評価することを提供 10  
する。いくつかの実施形態では、トレーニング集団中の対象ならびに試験対象から得られる各バイオマーカー・プロファイルは、複数の異なるバイオマーカーの各々の特徴を含む。いくつかの実施形態では、この比較は、(i) トレーニング集団からのバイオマーカー・プロファイルを使用して判断ルールを開発し、(ii) 判断ルールを試験対象からのバイオマーカー・プロファイルに適用することにより達成される。このように、本発明のいくつかの実施形態で適用される判断ルールを使用して、試験対象が情動障害を罹患しているかどうかを決定する。

**【0118】**

本発明のいくつかの実施形態では、判断ルールを適用した結果が、対象が情動障害を罹患していることを示す場合、対象は「情動障害」対象と診断される。判断ルールを適用した結果が、対象が情動障害を罹患していないことを示す場合、対象は「非情動障害」対象と診断される。したがって、いくつかの実施形態では、上述の二分決定状況における結果は、4つの考え得る結果を示す： 20

(i) 真に情動障害を罹患している。この場合、判断ルールは、対象が情動障害を罹患していることを示し、対象は実際に情動障害を罹患している（真陽性、TP）；

(ii) 偽に情動障害を罹患している。この場合、判断ルールは、対象が情動障害を罹患していることを示しているが、実際には対象は情動障害を罹患していない（偽陽性、FP）；

(iii) 真に情動障害を罹患していない。この場合、判断ルールは、対象が情動障害を罹患していないことを示し、対象は実際に情動障害を罹患していない（真陰性、TN） 30  
；または

(iv) 偽に情動障害を罹患していない。この場合、判断ルールは、対象が情動障害を罹患していないことを示し、対象は実際に情動障害を罹患している（偽陰性、FN）。

**【0119】**

TP、FP、TN、FNは、別様に定義することができることが理解されるだろう。そのような代替的定義はすべて本発明の範囲内にあるが、本発明の理解を容易にするために、別段の指示がない限り、上記(i)～(iv)の定義により与えられたTP、FP、TN、FNの定義が、本明細書で使用されるだろう。

**【0120】**

当業者であれば理解するように、多くの定量的判断基準を使用して、試験バイオマーカー・プロファイルと基準バイオマーカー・プロファイルとの間でなされる比較（例えば、試験対象からのバイオマーカー・プロファイルに判断ルールを適用すること）の性能を関連させることができる。これらには、陽性予測値（PPV）、陰性予測値（NPV）、特異性、感受性、正確度、および確実性が含まれる。加えて、受信者動作曲線（ROC）などの他の構成概念を使用して、判断ルールの性能を評価することができる。本明細書で使用される場合： 40

**【0121】**

## 【数 1】

$$PPV = \frac{TP}{TP+FP}$$

$$NPV = \frac{TN}{TN+FN}$$

$$\text{特異性} = \frac{TN}{TN+FP}$$

$$\text{感受性} = \frac{TP}{TP+FN}$$

$$\text{正確度} = \text{確実性} = \frac{TP+TN}{N}$$

10

数式中、Nは、比較された試料の数（例えば、試験試料の数）である。例えば、情動障害の分類が求められている対象が10人である場合を想定する。10人の試験対象の各々のバイオマーカー・プロファイルが構築される。その後、バイオマーカー・プロファイルの各々は、判断ルールを適用することにより評価され、この判断ルールは、トレーニング集団から得られたバイオマーカー・プロファイルに基づいて開発された。この例では、上記の数式のNは10に等しい。典型的には、Nは試料の数であり、各試料は集団の異なるメンバーから収集された。この集団は、実際には、2つの異なるタイプであり得る。1つのタイプでは、集団は、判断ルールを構築または改善するために、その試料および表現型データ（例えば、バイオマーカーの特徴値および対象が情動障害を罹患しているかどうかの兆候）が使用された対象を含む。そのような集団は、本明細書ではトレーニング集団と呼ばれる。別のタイプでは、集団は、判断ルールを構築するためには使用されなかった対象を含む。そのような集団は、本明細書では検証集団と呼ばれる。別段の指示がない限り、Nにより表される集団は、2つの集団タイプの混合ではなく、排他的にトレーニング集団または排他的に検証集団のいずれかである。正確度などのスコアは、それらが検証集団ではなくトレーニング集団に基づく場合、より高いだろう（1により近い）と理解されよう。にも関わらず、本明細書では、別段の指示が明示的でない限り、確実性（正確度）を含む判断ルール（または試験対象からのバイオマーカー・プロファイルを評価する他の形態）の性能を評価するために使用される判断基準はすべて、トレーニング集団または検証集団のいずれかに、その判断基準に対応する判断ルールを適用することにより測定される判断基準を指す。さらに、上記で定義されているPPV、NPV、特異性、感受性、および正確度の定義は、Draghici、Data Analysis Tools for DNA Microanalysis、2003年、CRC Press LLC社、ボークラトーン、フロリダ州、342～343頁（非特許文献26）にも見出すことができる。

20

30

## 【0122】

いくつかの実施形態では、Nは、1人を超えているか、5人を超えているか、10人を超えているか、20人を超えているか、10から100人の間であるか、100人を超えているか、または1000人未満の対象である。いくつかの実施形態では、判断ルール（または他の形態の比較）は、トレーニング集団または検証集団に対して、少なくとも約99%、またはそれさえもを超える確実性を有することができる。他の実施形態では、確実性は、トレーニング集団または検証集団に対して（およびしたがって、臨床患者などの、トレーニング集団の一部でない単一の対象に対して）、少なくとも約97%、少なくとも約95%、少なくとも約90%、少なくとも約85%、少なくとも約80%、少なくとも約75%、少なくとも約70%、少なくとも約65%、または少なくとも約60%である。有用な確実性の度合いは、本発明の特定の方法に依存して変動する場合がある。本明細書中で使用される場合、「確実性」は「正確度」を意味する。1つの実施形態では、感受性および/または特異性は、トレーニング集団または検証集団に対して、少なくとも約9

40

50

7%、少なくとも約95%、少なくとも約90%、少なくとも約85%、少なくとも約80%、少なくとも約75%、少なくとも約70%である。いくつかの実施形態では、そのような判断ルールを使用して、定められた正確度で、対象が情動障害を罹患しているかどうかを予測する。いくつかの実施形態では、そのような判断ルールを使用して、定められた正確度で情動障害を診断する。いくつかの実施形態では、そのような判断ルールを使用して、定められた正確度で、対象が情動障害の症状を示す可能性を決定する。

#### 【0123】

適切な確実性で試験対象を分類するために判断ルールに使用することができる特徴の数は、2またはそれ以上である。いくつかの実施形態では、特徴の数は、3またはそれ以上、4またはそれ以上、10またはそれ以上、または10~200の間である。しかしながら、判断ルールに使用される特徴の数は、求められている確実性の度合いに応じて、より多くともまたはより少なくともよいが、すべての場合で少なくとも2である。1つの実施形態では、試験対象を分類するために判断ルールに使用することができる特徴の数は、高い確実性で試験対象の分類を可能にするために最適化される。

10

#### 【0124】

判断ルールを開発するための関連データ分析アルゴリズムには、これらに限定されないが、以下のものを含む判別分析が含まれる：線形、ロジスティック、およびより柔軟性のある判別分析（例えば、Gnanadesikan、1977年、Methods for Statistical Data Analysis of Multivariate Observations、ニューヨーク市：Wiley社、1977年（非特許文献31）を参照）；分類回帰樹木（CART、classification and regression tree）および変法などの樹木に基づくアルゴリズム（例えば、Breiman、1984年、Classification and Regression Trees、ベルモント、カリフォルニア州：Wadsworth International Group社（非特許文献32）を参照）；一般化加法モデル；（例えば、Tibshirani、1990年、Generalized Additive Models、ロンドン：Chapman and Hall社（非特許文献33）を参照）；ならびにニューラル・ネットワーク（例えば、Neal、1996年、Bayesian Learning for Neural Networks、ニューヨーク市：Springer-Verlag社（非特許文献34））；およびInsua、1998年、Feedforward neural networks for nonparametric regression In: Practical Nonparametric and Semiparametric Bayesian Statistics、181~194頁、ニューヨーク市：Springer社（非特許文献35）、ならびに下記の5.5.2節を参照）。

20

30

#### 【0125】

1つの実施形態では、試験対象のバイオマーカー・プロファイルと、トレーニング集団から取得されたバイオマーカー・プロファイルとの比較が実施され、比較は判断ルールを適用することを含む。判断ルールは、コンピュータ・パターン認識アルゴリズムなどのデータ分析アルゴリズムを使用して構築される。判断ルールの構築に好適な他のデータ分析アルゴリズムには、これらに限定されないが、特徴値の分布における相異を検出するロジスティック回帰アルゴリズムまたはノンパラメトリック・アルゴリズムが含まれる（例えば、ウィルコクソン符号付順位検定（未補正および補正））。判断ルールは、1、2、3、4、5、10、または20個またはそれ以上バイオマーカーからの測定された観察可能値に対応する2、3、4、5、10、または20個またはそれ以上の特徴に基づいてよい。1つの実施形態では、判断ルールは、数百個またはそれ以上の特徴に基づく。判断ルールは、分類木アルゴリズムを使用して構築することもできる。例えば、トレーニング集団からの各バイオマーカー・プロファイルは、少なくとも3つの特徴を含むことができ、特徴は分類木アルゴリズムの予測因子である（下記の5.5.1節を参照）。判断ルールは、少なくとも約70%、少なくとも約75%、少なくとも約80%、少なくとも約85%、少なくとも約90%、少なくとも約95%、少なくとも約97%、少なくとも約98%、少なくとも約99%、または約100%の正確度で、集団（または種類）内のメンバーシップを予測する。

40

#### 【0126】

好適なデータ分析アルゴリズムは、当技術分野で公知であり、それらのいくつかは、Hastieら、上記（非特許文献17）で概説されている。特定の実施形態では、本発明のデー

50

タ分析アルゴリズムは、分類回帰樹木 (CART; 下記の5.5.1節)、多重加法型回帰木 (MART, Multiple Additive Regression Tree)、マイクロアレイ予測解析 (PAM, Prediction Analysis for Microarrays)、またはランダム・フォレスト分析 (下記の5.5.1節)を含む。そのようなアルゴリズムは、血液試料などの生体物質に由来する複雑なスペクトルを分類して、対象が、正常であるか、または特定の疾患状態に特徴的なバイオマーカー発現レベルを示すかを識別する。他の実施形態では、本発明のデータ分析アルゴリズムは、ANOVAおよびノンパラメトリックの等価的方法、線形判別分析、ロジスティック回帰分析、最近傍分類器解析 (nearest neighbor classifier analysis)、ニューラル・ネットワーク (下記の5.5.2節)、主成分分析、二次判別分析、回帰分類器 (regression classifier) およびサポート・ベクター・マシン (下記の5.5.4節)、関連ベクター・マシンおよび遺伝アルゴリズム (下記の5.5.5節)を含む。判断ルールを構築するために、および/または判断ルール適用の速度および効率を増加させるために、および研究者のバイアスを回避するために、そのようなアルゴリズムを使用することができるが、当業者であれば、本発明の方法を実行するためには、コンピュータに基づくアルゴリズムは必要ではないことを理解するだろう。

#### 【0127】

バイオマーカー・プロファイルを生成するために使用された方法に関わらず、判断ルールを使用して、バイオマーカー・プロファイルを評価することができる。例えば、バイオマーカー・プロファイルを評価するために使用することができる好適な判断ルールは、Harper、「Pyrolysis and GC in Polymer Analysis」、Dekker社、ニューヨーク市 (1985年) (非特許文献36)に考察されているように、ガスクロマトグラフィーを使用して生成される。さらに、Wagnerら、2002年、Anal. Chem. 74巻:1824~1835頁 (非特許文献37)では、静的飛行時間型二次イオン質量分析法 (TOF-SIMS)により取得されたスペクトルに基づいて対象を分類する能力を向上させる判断ルールが開示されている。加えて、Brightら、2002年、J. Microbiol. Methods 48巻:127~38頁 (非特許文献38)では、MALDI-TOF-MSスペクトルを解析することにより、高い確実性 (79~89%の正確な分類率)で細菌株を識別する方法が開示されている。Dalluge、2000年、Fresenius J. Anal. Chem. 366巻:701~711頁 (非特許文献39)では、複雑な生体試料中のバイオマーカー・プロファイルを分類するための、MALDI-TOF-MSおよび液体クロマトグラフィー・エレクトロスプレー・イオン化質量分析法 (LC/ESI-MS)の使用について考察されている。

#### 【0128】

##### 5.5.1 決定木

本発明で特定されるバイオマーカーの特徴値を使用して構築することができる1つのタイプの判断ルールは、決定木である。本明細書では、最終「決定木」は判断ルールであるが、「データ分析アルゴリズム」は、決定木を構築することができる任意の技術である。決定木は、トレーニング集団および特定のデータ分析アルゴリズムを使用して構築される。決定木は、Duda、2001年、Pattern Classification、John Wiley & Sons, Inc.社、ニューヨーク市、395~396頁 (非特許文献40)により一般的に記述されている。樹木に基づく方法は、特徴空間を一組の長方形に分割し、その後モデル (定数のような)をその各々にフィッティングする。

#### 【0129】

トレーニング集団データは、トレーニング・セット集団にわたって、本発明のバイオマーカーの特徴 (例えば、発現値、またはいくつかの他の観察可能値)を含んでいる。決定木を構築するために使用することができる1つの特定のアルゴリズムは、分類回帰樹木 (CART)である。他の特定の決定木アルゴリズムには、これらに限定されないが、ID3、C4.5、MART、およびランダム・フォレストが含まれる。CART、ID3、およびC4.5は、Duda、2001年、Pattern Classification、John Wiley & Sons, Inc.社、ニューヨーク市、396~408頁および411~412頁 (非特許文献41)に記述されている。CART、MART、およびC4.5は、Hastieら、2001年、The Elements of Statis

10

20

30

40

50

tical Learning、Springer-Verlag社、ニューヨーク市、9章（非特許文献42）に記述されている。ランダム・フォレストは、Breiman、1999年、「Random Forests - Random Features」、Technical Report 567、Statistics Department、U.C.Berkeley、1999年9月（非特許文献43）に記述されている。

【0130】

本発明のいくつかの実施形態では、決定木は、本発明のバイオマーカーの組合せの特徴を使用して、対象を分類するために使用される。決定木アルゴリズムは、教師あり学習アルゴリズムの種類に属する。決定木の目的は、現実世界の例データから分類器（木）を導出することである。この木は、決定木を導出するためには使用されていない未知の例を分類するために使用することができる。そのため、決定木は、トレーニング・データから導出される。例示的なトレーニング・データは、複数の対象（トレーニング集団）のデータを含んでいる。それぞれの対象各々について、それぞれの対象のクラス（例えば、情動障害を罹患している／情動障害を罹患していない）の複数の特徴がある。本発明の1つの実施形態では、トレーニング・データは、トレーニング集団わたるバイオマーカーの組合せの発現データである。

10

【0131】

一般的に、多くの異なる決定木アルゴリズムが存在し、それらの多くは、Duda、Pattern Classification、第2版、2001年、John Wiley & Sons, Inc社（非特許文献44）に記述されている。決定木アルゴリズムには、特徴処理、不純物測定、停止判断基準、および枝刈り（pruning）に関する考慮が必要とされることが多い。特定の決定木アルゴリズムには、これらに限定されないが、分類回帰樹木（CART）、多変量決定木、ID3、およびC4.5が含まれる。

20

【0132】

1つの手法では、決定木が使用される場合、本発明に記述されている遺伝子の選択組合せの、トレーニング集団にわたる遺伝子発現データは、平均0および単位分散（mean zero and unit variance）に対して正規化される。トレーニング集団のメンバーは、トレーニング・セットおよび試験セットに無作為に分割される。例えば、1つの実施形態では、トレーニング集団のメンバーの3分の2は、トレーニング・セットに配置され、トレーニング集団のメンバーの3分の1は、試験セットに配置される。本発明に記述されているバイオマーカーの選択組合せの発現値を使用して、決定木を構築する。その後、決定木が試験セットのメンバーを正確に分類する能力を決定する。いくつかの実施形態では、この計算は、バイオマーカーの所与の組合せごとに数回実施される。各計算を反復する際には、トレーニング集団のメンバーは、トレーニング・セットおよび試験セットに無作為に割当てられる。その後、バイオマーカーの組合せの品質は、決定木計算のそのような反復の各々の平均として得られる。

30

【0133】

各分岐が、本発明の一組のバイオマーカーの中の対応するバイオマーカーの特徴値、または2つのそのようなバイオマーカーの相対的な特徴値に基づく単変量決定木に加えて、多変量決定木を、判断ルールとして実装することができる。そのような多変量決定木では、決定のいくつかまたはすべては、実際に本発明の複数のバイオマーカーの特徴値の線形結合を含む。そのような線形結合は、分類に対する勾配降下法または誤差平方和判断基準の使用などの公知の技術を使用してトレーニングすることができる。そのような決定木を例示するために、以下の数式を想定する：

40

$$0.04x_1 + 0.16x_2 < 500$$

数式中、 $x_1$  および  $x_2$  は、本発明のバイオマーカーの中からの2つの異なるバイオマーカーの2つの異なる特徴を指す。判断ルールを選択するために、特徴  $x_1$  および  $x_2$  の値を、未分類の対象から得られる測定から取得する。その後、これらの値を数式に挿入する。500未満の値が算出される場合、決定木における第1の分岐が得られる。そうでなければ、決定木における第2の分岐が得られる。多変量決定木は、Duda、2001年、Pattern Classification、John Wiley & Sons, Inc.社、ニューヨーク市、408~409頁（非特許

50

文献 45) に記述されている。

【0134】

本発明で 사용할 ことができる別の手法は、多変量適応的回帰スプライン (MARS) である。MARS は、回帰のための適応的手順であり、本発明が取り組んでいる高次元の問題によく適している。MARS は、段階的線形回帰の一般化、または回帰設定において CART の性能を向上する CART 法の変法と見ることができる。MARS は、Hastie ら、2001年、The Elements of Statistical Learning、Springer-Verlag社、ニューヨーク市、283~295頁 (非特許文献 46) に記述されている。

【0135】

5.5.2 ニューラル・ネットワーク

いくつかの実施形態では、本発明の選択バイオマーカーについて測定された特徴データ (例えば、RT-PCR データ、質量分析データ、マイクロアレイ・データ) を使用して、ニューラル・ネットワークをトレーニングすることができる。ニューラル・ネットワークは、2段階回帰または分類判断ルールである。ニューラル・ネットワークは、重み層により出力ユニット層に接続されている入力ユニット (およびバイアス) 層を含む階層構造を有する。回帰の場合、出力ユニット層は、典型的には1つの出力ユニットのみを含んでいる。しかしながら、ニューラル・ネットワークは、複数の定量的応答をシームレスな様式で取り扱うことができる。

10

【0136】

多層ニューラル・ネットワークには、入力ユニット (入力層)、隠れユニット (隠れ層)、および出力ユニット (出力層) がある。さらに、入力ユニットではない各ユニットに接続されている単一のバイアスユニットがある。ニューラル・ネットワークは、Duda ら、2001年、Pattern Classification、第2版、John Wiley & Sons, Inc.社、ニューヨーク市 (非特許文献 45) ; および Hastie ら、2001年、The Elements of Statistical Learning、Springer-Verlag社、ニューヨーク市 (非特許文献 17) に記述されている。ニューラル・ネットワークは、Draghici、2003年、Data Analysis Tools for DNA Microarrays、Chapman & Hall/CRC社 (非特許文献 15) ; および Mount、2001年、Bioinformatics: sequence and genome analysis、Cold Spring Harbor Laboratory Press社、コールド・スプリング・ハーバー、ニューヨーク市 (非特許文献 47) にも記述されている。下記に開示されるものは、ニューラル・ネットワークのいくつかの例示的な形態である。

20

30

【0137】

ニューラル・ネットワークの使用に対する基本的な手法は、トレーニングされていないネットワークから出発して、トレーニング・パターンを入力層に提示し、シグナルをネットに通し、出力層での出力を決定することである。その後、これら出力を標的値と比較し、いかなる相異も誤差に相当する。この誤差または判断基準関数は、重みの何らかのスカラ関数であり、ネットワーク出力が所望の出力と一致する場合に最小化される。したがって、重みは、この誤差の測定を低減するように調整される。回帰の場合、この誤差は、誤差平方和であってもよい。分類の場合、この誤差は、平方誤差またはクロス・エントロピー (偏差) のいずれであってもよい。例えば、Hastie ら、2001年、The Elements of Statistical Learning、Springer-Verlag社、ニューヨーク市 (非特許文献 17) を参照されたい。

40

【0138】

3つの一般的に使用されるトレーニング・プロトコールは、確率的、バッチ、およびオンラインである。確率的トレーニングでは、パターンは、トレーニング・セットから無作為に選択され、ネットワークの重みは、各パターン提示のために更新される。確率的逆伝播 (stochastic back-propagation) などの勾配降下法によりトレーニングされた多層非線形ネットワークは、ネットワーク・トポロジーにより定義される分類器の重み値の最尤推定を実施する。バッチ・トレーニング (batch training) では、学習が起る前にパターンはすべてネットワークに提示される。典型的には、バッチ・トレーニングでは、トレーニング・データを通していくつかの受け渡しがなされる。オンライン・トレーニング (on

50

line training) では、各パターンは、一度だけ (once and only once) でネットに提示される。

#### 【0139】

いくつかの実施形態では、重みの出発値を考慮する。重みが 0 に近い場合、ニューラル・ネットワークの隠れ層において一般的に使用されるシグモイドの操作部分 (operative part of the sigmoid) (例えば、Hastieら、2001年、The Elements of Statistical Learning、Springer-Verlag社、ニューヨーク市 (非特許文献 17) を参照) は、おおよそ線形であり、したがって、ニューラル・ネットワークは、ほぼ線形分類器へと崩壊する。いくつかの実施形態では、重みの出発値は、0 に近い無作為の値であるように選択される。したがって、分類器は、ほぼ線形で始まり、重みが増加するにつれて非線形になる。個別ユニットは向きに限局化し、必要であれば非線形性を導入する。正確に 0 の重みを使用することは、0 の導関数および完全な相称性に結びつき、アルゴリズムは決して移動しない。あるいは、大きな重みを用いて出発することは、不良な解に結びつくことが多い。

10

#### 【0140】

入力のスケーリングは、底部層における効果的な重みのスケーリングを決定するため、最終解の品質に大きな影響を及ぼす場合がある。したがって、いくつかの実施形態では、最初は、出発値はすべて平均 0 および標準偏差 1 を有するように標準化される。これにより、正則化プロセスにおいて入力がすべて等しく処理されることが保証され、無作為出発重みの意味のある範囲を選択することが可能になる。標準化入力と共に、[ - 0.7, + 0.7 ] の範囲にわたる無作為の一定した重みを取ることが典型的である。

20

#### 【0141】

3層ネットワークの使用における反復性問題は、ネットワークで使用する隠れユニットの最適数である。3層ネットワークの入力および出力の数は、解決しようとする問題により決定される。本発明では、所与のニューラル・ネットワークへの入力数は、トレーニング集団から選択されるバイオマーカーの数と等しいだろう。ニューラル・ネットワークの出力数は、典型的にはちょうど 1 だろう。しかしながら、いくつかの実施形態では、ちょうど 2 を超える状態がネットワークにより定義され得るように、1 を超える出力が使用される。例えば、多重出力ニューラル・ネットワークを使用して、健常表現型と情動障害の種々の段階とを区別することができる。あまりにも多くの隠れユニットがニューラル・ネットワークで使用されると、ネットワークはあまりにも多くの自由度を有することになり、トレーニングが長くなり過ぎ、ネットワークがデータを過剰フィッティングすることになるというリスクがある。隠れユニットがあまりにも少数である場合、トレーニング・セットは学習することができない。しかしながら、一般的に言えば、少数過ぎるよりは多過ぎる隠れユニットを有するほうが良好である。隠れユニットが少数過ぎると、分類器は、データの非線形性をとらえるのに十分な柔軟性を有することができず、隠れユニットが多過ぎると、下記に記述されているような適切な正則化または枝刈りが使用される場合、余分な重みが 0 に向かって収縮する場合がある。典型的な実施形態では、隠れユニットの数は、5 ~ 100 の範囲のどこかにあり、この数は、入力の数およびトレーニング事例の数が増加するにつれて増加する。

30

#### 【0142】

使用する隠れユニットの数を決定する 1 つの一般的な手法は、正則化手法を適用することである。正則化手法では、古典的トレーニング誤差だけでなく分類器の複雑さにも依存する新しい判断基準関数が構築される。具体的には、新しい判断基準関数は、高度に複雑な分類器にペナルティを課し、この判断基準中での最小値を探索することは、トレーニング・セットの誤差とトレーニング・セット + 正則化項の誤差とのバランスをとることであり、それは、解の制限または所望の特性を表す：

40

$$J = J_{\text{pat}} + J_{\text{reg}}$$

パラメーター は、正則化を課す強度の多少を調整する。言いかえれば、より大きな値の は、重みを 0 へと収縮させる傾向があるだろう。典型的には、検証セットとの相互検証を使用して を推定する。この検証セットは、トレーニング集団の無作為サブセットを

50

確保することにより取得することができる。ペナルティの他の形態、例えば重み推定ペナルティ (weight elimination penalty) が提案されている (例えば、Hastieら、2001年、The Elements of Statistical Learning、Springer-Verlag社、ニューヨーク市 (非特許文献17) を参照)。

【0143】

使用する隠れユニットの数を決定する別の手法は、最も必要でない重みを除去、つまり枝刈りすることである。1つの手法では、最も小さな重さを有する重みが除去される (0に設定される)。そのような重さに基づく枝刈りは機能し得るが、最適ではない。小さな重さの重みが学習およびトレーニング・データにとって重要であることがある。いくつかの実施形態では、重さに基づく枝刈り手法を使用するのではなく、Wald統計量が計算される。Wald統計量の基本観念は、それらを使用して、分類器における隠れユニット (重み) の重要性を推定することができるということである。その後、最小の重要性を有する隠れユニットが除去される (それらの入力および出力重みを0に設定することにより)。この点に関する2つのアルゴリズムは、最適脳損傷 (OBD、Optimal Brain Damage) および最適脳外科医 (OBS、Optimal Brain Surgeon) アルゴリズムであり、これらは、二次近似を用いてトレーニング誤差がどのように重みに依存するかを予測し、最小のトレーニング誤差増加をもたらす重みを除去する。

10

【0144】

最適脳損傷および最適脳外科医は、重み  $w$  における極小誤差に対してネットワークをトレーニングし、その後最小トレーニング誤差増加をもたらす重みを枝刈りするという同一の基礎的手法を共有する。全重みベクトル  $w$  の変化に対する誤差の関数的増加の予測は、以下の通りである：

20

【0145】

【数2】

$$\delta J = \left( \frac{\partial J}{\partial w} \right)^t \cdot \delta w + \frac{1}{2} \delta w^t \cdot \frac{\partial^2 J}{\partial w^2} \cdot \delta w + O(\|\delta w\|^3)$$

数式中、

【0146】

【数3】

$$\frac{\partial^2 J}{\partial w^2}$$

30

はヘッシアン行列である。第1項は、誤差の極小値で消滅し、第3項以上の高次項は無視される。1つの重みを欠失するという制限が与えられれば、この関数を最小化するための一般解は以下の通りである：

【0147】

【数4】

$$\delta w = - \frac{w_q}{[\mathbf{H}^{-1}]_{qq}} \mathbf{H}^{-1} \cdot u_q \text{ および } L_q = \frac{1}{2} - \frac{w_q^2}{[\mathbf{H}^{-1}]_{qq}}$$

40

数式中、 $u_q$  は、重み空間の  $q$  次方向に沿った単位ベクトルであり、 $L_q$  は、重み  $q$  が枝刈りされ、他の重みが  $w$  に更新された場合の、重み  $q$  の突出に対する近似、つまりトレーニング誤差の増加である。これらの数式は、 $\mathbf{H}$  の逆行列を必要とする。この逆行列を計算する1つの方法は、小さな値を用いて出発することである。 $\mathbf{H}_0^{-1} = \mathbf{I}$ 、ここでは、小さなパラメータであり、事実上、重み定数である。次に、この行列を、以下の数式に従って各パターンを用いて更新する：

【0148】

【数 5】

$$\mathbf{H}_{m+1}^{-1} = \mathbf{H}_m^{-1} - \frac{\mathbf{H}_m^{-1} \mathbf{X}_{m+1} \mathbf{X}_{m+1}^T \mathbf{H}_m^{-1}}{\frac{n}{a_m} + \mathbf{X}_{m+1}^T \mathbf{H}_m^{-1} \mathbf{X}_{m+1}} \quad \text{数式1}$$

数式中、下付き文字は、提示されているパターンに対応し、 $a_m$  は、 $m$  と共に減少する。十分なトレーニング・セットが提示された後、逆ヘッシアン行列は、 $\mathbf{H}^{-1} = \mathbf{H}_n^{-1}$  で与えられる。アルゴリズムの形態の最適脳外科医法は以下の通りである：

```
begin initialize  $n_H, w$ , (開始 $n_H, w$ 、 を初期化する)
  train a reasonably large network to minimum error (適度に大きなネットワークをトレーニングして、誤差を最小化する)
  do compute  $\mathbf{H}^{-1}$  by Eqn. 1 ( $\mathbf{H}^{-1}$  を数式 1 で計算する)
```

【0149】

【数 6】

$$q^* \leftarrow \arg \min_q w_q^2 / (2[H^{-1}]_{qq}) \quad (\text{突出}L_q)$$

$$w \leftarrow w - \frac{w_{q^*}}{[H^{-1}]_{q^*q^*}} H^{-1} e_{q^*} \quad (\text{突出}L_q)$$

until  $J(w) > \theta$  ( $J(w) > \theta$  まで継続する)

return  $w$  ( $w$  を返す)

end (終了)

20

【0150】

最適脳損傷法は計算としてはより単純である。なぜならライン 6 の逆ヘッシアン行列の計算は、対角行列が特に単純であるからである。上記のアルゴリズムは、誤差が初期化された判断基準より大きい場合に終了する。別の手法は、重みの除去による  $J(w)$  の変化が、ある判断基準値より大きい場合に終了するように、ライン 6 を変更することである。いくつかの実施形態では、逆伝播ニューラル・ネットワーク、例えば、Abdi、1994年、「A neural network primer」、J. Biol System. 2巻、247~283頁（非特許文献 48）を参照されたい。

【0151】

30

### 5.5.3 クラスタリング

いくつかの実施形態では、本発明の選択バイオマーカ－の特徴を使用して、トレーニング・セットをクラスタリングする。例えば、本発明に記述されている 10 個の特徴（10 個のバイオマーカ－に対応する）が使用される場合を想定する。トレーニング集団の各メンバー  $m$  は、10 個のバイオマーカ－の各々の特徴値（例えば、発現値）を有することになる。トレーニング集団中のメンバー  $m$  に由来するそのような値は、以下のベクトルを定義する：

$$X_{1m} \quad X_{2m} \quad X_{3m} \quad X_{4m} \quad X_{5m} \quad X_{6m} \quad X_{7m} \quad X_{8m} \quad X_{9m} \quad X_{10m}$$

ここで、 $X_{im}$  は、生物  $m$  の  $i$  番目のバイオマーカ－の発現レベルである。トレーニング・セットに  $m$  生物が存在する場合、 $i$  バイオマーカ－の選択は、 $m$  ベクトルを定義することになる。本発明の方法は、ベクトルに使用されるすべてのあらゆるバイオマーカ－の各発現値が、すべてのあらゆるベクトル  $m$  に表されている必要がないことに留意されたい。言い換えれば、 $i$  番目のバイオマーカ－の 1 つが見出されていない対象からのデータでも、クラスタリングに使用することができる。そのような場合では、見当らない発現値には、「0」または他の何らかの正規化値のいずれかが割当てられる。いくつかの実施形態では、特徴値は、クラスタリングに先立って、0 の平均値および単位分散を有するように正規化される。

40

【0152】

トレーニング・グループにわたって同様の発現パターンを示すトレーニング集団のメン

50

バーは、一緒にクラスタリングされる傾向があるだろう。本発明の遺伝子の特定の組合せは、ベクトルがトレーニング集団に見出される形質グループにクラスタリングされる場合、本発明のこの態様の良好な分類器であるとみなされる。例えば、トレーニング集団が、クラス a : 情動障害を罹患していない研究中的対象、およびクラス b : 情動障害を罹患している研究中的対象を含んでいる場合、理想的なクラスタリング分類器は、集団を2つのグループにクラスタリングし、一方のクラスター・グループはクラス a を固有に表し、他方のクラスター・グループはクラス b を固有に表すだろう。

【0153】

クラスタリングは、DudaおよびHart、Pattern Classification and Scene Analysis、1973年、John Wiley & Sons, Inc.社、ニューヨーク市、211~256頁（非特許文献49）（以下「Duda 1973年」と称する）に記述されている。Duda 1973年の6.7節に記述されているように、クラスタリング問題は、データセットの自然なグループ化を見出す問題として記述されている。自然なグループ化を特定するために、2つの課題に取り組む。第1に、2つの試料間の類似性（または相違性）を測定する方法を決定する。この計量（類似性尺度）を使用して、あるクラスター中の試料が、他のクラスター中の試料と類似するよりも、互いにより類似していることを保証する。第2に、類似性尺度を使用して、データをクラスターに分割するための機序を決定する。

10

【0154】

類似性尺度は、Duda 1973年の6.7節で考察されており、クラスタリング調査を始める1つの方法は、データセット中の全試料対間の距離関数を定義し、距離のマトリックスを計算することであると述べられている。距離が類似性の良好な尺度である場合、同じクラスター中の試料間の距離は、異なるクラスター中の試料間の距離より著しく短いだろう。しかしながら、Duda 1973年の215頁に述べられるように、クラスタリングは、距離計量の使用を必要としない。例えば、非計量的類似性関数  $s(x, x')$  を使用して、2つのベクトル  $x$  および  $x'$  を比較することができる。慣例的に、 $s(x, x')$  は、 $x$  および  $x'$  がある程度「類似」している場合にその値が大きい対称的関数である。非計量的類似性関数  $s(x, x')$  の一例は、Duda 1973年の216頁に提供される。

20

【0155】

データセット中のポイント間の「類似性」または「相違性」を測定するための方法が選択されると、クラスタリングは、データの任意の区分のクラスタリング品質を測定する判断基準関数を必要とする。判断基準関数を極値化するデータセットの区分を使用して、データをクラスタリングする。Duda 1973年の217頁を参照されたい。判断基準関数は、Duda 1973年の6.8節で考察されている。

30

【0156】

より最近になって、Dudaら、Pattern Classification、第2版、John Wiley & Sons, Inc社、ニューヨーク市（非特許文献50）が出版された。537~563頁には、クラスタリングについて詳細に記述されている。クラスタリング技術については、以下の文献により多くの情報を見出すことができる：KaufmanおよびRousseeuw、1990年、Finding Groups in Data: An Introduction to Cluster Analysis、Wiley社、ニューヨーク市、ニューヨーク州（非特許文献51）；Everitt、1993年、Cluster analysis（第3版）、Wiley社、ニューヨーク市、ニューヨーク州（非特許文献52）；およびBacker、1995年、Computer-Assisted Reasoning in Cluster Analysis、Prentice Hall社、アッパー・サドル・リバー、ニュージャージー州（非特許文献53）。本発明で使用することができる特定の例示的なクラスタリング技術には、これらに限定されないが、階層的クラスタリング（最近傍アルゴリズム、最長法アルゴリズム（farthest-neighbor algorithm）、平均連結法アルゴリズム（average linkage algorithm）、重心アルゴリズム（centroid algorithm）、または平方和アルゴリズムを使用する凝集クラスタリング）、 $k$ -平均クラスタリング（ $k$ -means clustering）、曖昧  $k$ -平均クラスタリング・アルゴリズム（fuzzy  $k$ -means clustering algorithm）、およびJarvis-Patrickクラスタリングが含まれる。

40

50

## 【 0 1 5 7 】

## 5 . 5 . 4 サポート・ベクター・マシン

本発明のいくつかの実施形態では、サポート・ベクター・マシン (SVM) を使用して、本発明に記述されている遺伝子の特徴値を使用して、対象を分類する。SVMは、比較的新しいタイプの学習アルゴリズムである。例えば、CristianiniおよびShawe-Taylor、2000年、An Introduction to Support Vector Machines、Cambridge University Press社、ケンブリッジ市 (非特許文献54) ; Boserら、1992年、「A training algorithm for optimal margin classifiers」、in Proceedings of the 5<sup>th</sup> Annual ACM Workshop on Computational Learning Theory、ACM Press社、ピッツバーグ、ペンシルベニア州、142~152頁 (非特許文献55) ; Vapnik、1998年、Statistical Learning Theory、Wiley社、  
10  
ニューヨーク市 (非特許文献56) ; Mount、2001年、Bioinformatics: sequence and genome analysis、Cold Spring Harbor Laboratory Press社、コールド・スプリング・ハーバー、ニューヨーク市 (非特許文献47) ; Duda、Pattern Classification、第2版、2001年、John Wiley & Sons, Inc.社 (非特許文献45) ; およびHastie、2001年、The Elements of Statistical Learning、Springer社、ニューヨーク市 (非特許文献17) ; およびFureyら、2000年、Bioinformatics 16巻、906~914頁 (非特許文献57) を参照されたい。分類に使用される場合、SVMは、それらから最大距離にある超平面を有する所与のセットの二成分標識データ・トレーニング・データを分離する。線形分離が可能でない場合、SVMは、特徴空間に対する非線形マッピングを自動的に実現する「カーネル (kernel)」の技術と組み合わせて作働することができる。SVMにより特徴空間に見出される  
20  
超平面は、入力空間の非線形決定境界に対応する。

## 【 0 1 5 8 】

1つの手法では、SVMが使用される場合、特徴データは、平均0および単位分散に対して正規化され、トレーニング集団のメンバーは、トレーニング・セットおよび試験セットに無作為に分割される。例えば、1つの実施形態では、トレーニング集団のメンバーの3分の2は、トレーニング・セットに配置され、トレーニング集団のメンバーの3分の1は、試験セットに配置される。本発明に記述されている遺伝子の組合せの発現値を使用して、SVMをトレーニングする。その後、トレーニングされたSVMが試験セットのメンバーを正確に分類する能力を決定する。いくつかの実施形態では、この計算は、分子マーカーの所与の組合せごとに数回実施される。各計算を反復する際には、トレーニング集団  
30  
のメンバーは、トレーニング・セットおよび試験セットに無作為に割当てられる。その後、バイオマーカーの組合せの品質は、SVM計算のそのような反復の各々の平均として得られる。

## 【 0 1 5 9 】

## 5 . 5 . 5 関連ベクター・マシンおよび遺伝的アルゴリズム

関連ベクター・マシン (RVM) は、回帰ならびに教師ありのマルチクラス分類問題に使用可能なカーネルに基づくベイジアン統計モデルである。(Tipping, M: Sparse Bayesian Learning and the Relevance Vector Machine、Journal of Machine Learning Research 1巻、2001年、211~244頁 (非特許文献58))。分類ツールとして使用される場合、  
40  
トレーニングされたRVMは、新しいデータポイントのクラス・メンバーシップに関する確率論的予測を行う。RVMモデルでは、所定のセットの説明変数 (つまり、遺伝子またはバイオマーカー) が、ロジスティクス・リンク関数によりクラス・メンバーシップ確率に影響を及ぼすことが仮定されている。多数の候補変数から選択される説明変数の最適セットを決定するために、RVMモデルは、異なるサブセットの候補変数に対してトレーニングおよび試験される多数のRVMを評価する遺伝的最適化アルゴリズム内部で作働している (Deb, K: Multi-Objective Optimization using Evolutionary Algorithms、Wiley社、2001年 (非特許文献59))。各変数サブセットの性能は、相互検証により評価される。

## 【 0 1 6 0 】

## 5 . 5 . 6 他のデータ分析アルゴリズム

10

20

30

40

50

上述のデータ分析アルゴリズムは、陽転者と非陽転者とを区別するために判断ルールを構築するために使用することができるタイプの方法の例に過ぎない。さらに、上述の技術の組合せを使用することができる。決定木およびブースティングの組合せの使用などのいくつかの組合せが記述されている。しかしながら、他の多くの組合せが可能である。加えて、射影追跡 (Projection Pursuit) および加重投票 (Weighted Voting) などの当技術分野の他の技術を使用して、判断ルールを構築することができる。

#### 【0161】

##### 5.6 バイオマーカー

特定の実施形態では、バイオマーカー・プロファイルは、表1Aに列挙されている少なくとも2つの異なるバイオマーカーを含む。バイオマーカー・プロファイルは、少なくとも2つのバイオマーカーのそれぞれの対応する特徴をさらに含む。そのようなバイオマーカーは、例えば、mRNA転写物、cDNAもしくは他の何らかの核酸、例えば増幅核酸、またはタンパク質であってもよい。一般的に、少なくとも2つのバイオマーカーは、少なくとも2つの異なる遺伝子に由来している。少なくとも2つの異なるバイオマーカー中のバイオマーカーが表1Aに列挙されている場合、バイオマーカーは、例えば、列挙されている遺伝子により産生される転写物、その相補体、またはその特徴的断片もしくは相補体、またはそのcDNA、そのcDNAの特徴的断片、または転写物もしくはその相補体のすべてもしくは部分に対応する特徴的増幅核酸分子、または該遺伝子によりコードされるタンパク質、またはそのタンパク質の特徴的断片、または上記のいずれかの兆候であってもよい。そのような実施形態によると、本発明のバイオマーカー・プロファイルは、当業者に公知である任意の標準的アッセイを使用して、または本明細書に記述されているアッセイで、バイオマーカーを検出して、取得することができる。そのようなアッセイは、例えば、特定の遺伝子、または目的遺伝子 (例えば表1Aに開示されている遺伝子) の対立遺伝子の発現産物 (例えば、核酸および/またはタンパク質) の検出が可能である。1つの実施形態では、そのようなアッセイでは、核酸マイクロアレイが使用される。

#### 【0162】

いくつかの実施形態では、バイオマーカー・プロファイルは、表1Aに列挙されている2~29個のバイオマーカーを有する。いくつかの実施形態では、バイオマーカー・プロファイルは、表1Aに列挙されている3~20個のバイオマーカーを有する。いくつかの実施形態では、バイオマーカー・プロファイルは、表1Aに列挙されている4~15個のバイオマーカーを有する。いくつかの実施形態では、バイオマーカー・プロファイルは、表1Aに列挙されている少なくとも2個のバイオマーカーを有する。いくつかの実施形態では、バイオマーカー・プロファイルは、表1Aに列挙されている少なくとも3個のバイオマーカーを有する。いくつかの実施形態では、バイオマーカー・プロファイルは、表1Aに列挙されている少なくとも4個のバイオマーカーを有する。いくつかの実施形態では、バイオマーカー・プロファイルは、表1Aに列挙されている少なくとも2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、または25個またはそれ以上のバイオマーカーを含む。いくつかの実施形態では、そのようなバイオマーカーの各々は核酸である。いくつかの実施形態では、そのようなバイオマーカーの各々はタンパク質である。いくつかの実施形態では、バイオマーカー・プロファイル中のバイオマーカーのいくつかは核酸であり、バイオマーカー・プロファイル中のバイオマーカーのいくつかはタンパク質である。

#### 【0163】

##### 5.7 具体的な実施形態

本発明の1つの態様は、情動障害の症状を示す可能性が高い対象の遺伝子転写プロファイルを識別する方法に関する。そのような遺伝子転写プロファイルは、表1Aから選択される遺伝子などの、対象の生体試料に由来する選択された遺伝子の転写解析に基づく。

#### 【0164】

本発明を使用して、単一のプロファイルへと凝集させることができる個々のバイオマーカーの含有量 (例えば、発現レベル) を識別および分析することが可能である。そのよう

な含有量プロファイルは、疾患を分類するための特徴として使用される。下記で考察されているように、転写解析は、対照対象および罹患対象の全血試料中の遺伝子発現プロファイルを決断するために行われる。表 1 A から選択される遺伝子の含有量は、表 4、表 5、および表 6 に例示されている。表 4、表 5、および表 6 の各々は、それぞれうつ病対象、重度うつ病対象、および双極性対象の、対照と比較した遺伝子転写プロファイルの代表的な例である。1つの実施形態では、表 4 に示されているようなうつ病遺伝子転写プロファイルを有する対象は、うつ病を罹患していると診断される。別の実施形態では、表 5 に示されているような重度うつ病遺伝子転写プロファイルを有する対象は、重度うつ病を罹患していると診断される。別の実施形態では、表 6 に示されているような双極性遺伝子転写プロファイルを有する対象は、双極性障害を罹患していると診断される。遺伝子転写プロファイルのさらなる代表的な例は、表 4 A および 5 B に示されている。

10

**【 0 1 6 5 】**

1つの例では、遺伝子発現プロファイルを決断するために使用されるバイオマーカーは、表 1 A に記述されている遺伝子から選択された。代表的な転写バイオマーカー・プローブ・セットも、表 1 A に記述されている。そのプローブ・セットを使用して、周知の方法により定量 P C R ( q P C R ) を実施した。

**【 0 1 6 6 】**

本発明の態様は、表 1 A から選択される遺伝子の転写解析により決定されるような各対象に転写プロファイルを供給する。

**【 0 1 6 7 】**

転写解析は、当技術分野で周知の方法により実施することができる。例として、メッセンジャー R N A ( m R N A ) を含む R N A は、動物身体、特にヒト身体の細胞物質、または細胞物質を含有する液体から単離することができる。細胞物質は、m R N A を含む細胞内容物を含有することが理解される。本発明で使用される生体試料は、例えば、末梢組織、全血、脳脊髄液、腹水、および間質液から選択されてもよい。

20

**【 0 1 6 8 】**

本発明の他の実施形態では、生体試料は、全血、脳脊髄液、および末梢組織からなる群から選択される。本発明は、赤血球細胞 ( R B C )、白血球細胞、および血小板からなる群から選択される全血の画分を使用して実施することもできる。白血球細胞 ( 白血球 ) には、これらに限定されないが、好中球、好塩基球、好酸球、リンパ球、マクロファージ、および単球が含まれる。

30

**【 0 1 6 9 】**

試料中の遺伝子発現を測定するために、その試料中の R N A または m R N A を逆転写にかけて複製 D N A を生成し、その後その D N A 配列に基づくプローブまたはプライマー配列を使用して標準的方法により分析できる。個々の遺伝子の各々は、ポリメラーゼ連鎖反応 ( P C R )、定量 P C R、in situ ハイブリダイゼーション、ノーザンブロット解析、ビーズに基づくアッセイまたは遺伝子アレイなどの固体支持体固定化アッセイ、および当技術分野で周知の他の方法により解析することができる。

**【 0 1 7 0 】**

本明細書に記述されている本発明の態様によると、定量 P C R ( q P C R ) を使用して、m R N A レベルを測定する。1つまたは複数の核酸プローブを使用して、生体試料に由来する m R N A レベルを測定した。プローブまたはプライマーは、目的の遺伝子に相補的なヌクレオチド ( n t ) 配列であり、そのようなプローブ/プライマーの選択および合成は、当業者に周知の方法により行われる。本発明のプローブ/プライマーは、表 1 A に記述されているヌクレオチド配列に限定されない。

40

**【 0 1 7 1 】**

本発明は、対象から取得した生体試料から分析されるような対象の転写プロファイルを決断することにより、対照対象と比較して罹患対象を分類する方法をさらに提供する。

**【 0 1 7 2 】**

本発明は、表 1 A から選択される遺伝子の転写解析により決定される個別的転写プロフ

50

ファイルを提供する。そのような転写プロファイルは、既知の健常対照対象または既知の罹患対象の転写プロファイルに類似していると決定される場合、対象に個別的であると決定される。既知の健常対照対象または既知の罹患対象の転写プロファイルとの類似性は、本明細書に記述されているように、分類アルゴリズムなどの分類法により決定される。

#### 【0173】

いくつかの実施形態では、転写データは、本明細書に記述されているように、複数の対照対象から収集される。転写データは、本明細書中に記述されているように、情動障害などの疾患または障害を罹患している複数の対象から収集される。データ分析アルゴリズムは、各セットの転写データを入力として、各転写データ・セットに含有されている分類遺伝子を区別または識別するために使用される。そのようなアルゴリズムは、典型的には、「分類器」としても知られている分類アルゴリズムとして記述されている。この課題を実施するために使用されるデータ分析アルゴリズムは、当業者に周知であり、以下の例を使用することができる：ランダム・フォレスト法 (Breiman, L., 2001年、Machine Learning 45巻 (1号) : 5~32頁 (非特許文献60))、サポート・ベクター・マシン (SVM) (Cortes, C. および Vapnik, V., 1995年、Machine Learning, 20巻 (3号) : 273~97頁 (非特許文献61))、段階的ロジスティック回帰法 (SLR) (Ersboll, B.K. および Conradsen, K. (2005年) An Introduction to Statistics, 第7版、IMM (非特許文献62) ; Drapper, N. および Smith, H. (1981年) Applied Regression Analysis, 第2版、ニューヨーク市 : John Wiley & Sons, Inc. 社 (非特許文献63))、再帰分割法 (RPART) (James K.E. ら、2005年、Statistics in Medicine, 24巻 (19号) : 3019~35頁 (非特許文献64))、罰則付ロジスティック回帰分析 (Penalized Logistic Regression Analysis) (PELORA) (Dettling, M., 2003年、Proceedings of the 3<sup>rd</sup> International Workshop on Distributed Statistical Computing, 3月20~22日、ウィーン、オーストリア、Hornik, L., Leisch, および Seilis 編 (非特許文献65))、ニューラル・ネットワーク、関連ベクター・マシン (RVM)、ロジットブースト (LogitBoost) (Friedman, J., Hastie, T., および Tibshirani, R. 2000年、Annals of Statistics 28巻 (2号) : 337~407頁 (非特許文献66))、マイクロアレイ予測解析 (PAM)、およびその他 (V. N. Vapnik, Statistical Learning Theory, Wiley社、ニューヨーク市、1998年 (非特許文献67) 参照)。そのような分類アルゴリズムまたは「分類器」は、それらの転写データに基づく患者の分類に関する出力を提供するために調整およびトレーニングされる。

#### 【0174】

トレーニングされた分類アルゴリズムにより選択される遺伝子またはバイオマーカーを分類することにより、特定のデータセットが属するクラス、例えば、対照データに関連するクラスまたは疾患データと関連するクラスのいずれかに関連する転写データの予測尺度がもたらされる。

#### 【0175】

任意の特定の理論により束縛されることは望まないが、ランダム・フォレスト・アルゴリズムは、多数の決定木からの出力に基づいて対象を分類するアンサンブル学習法とみなされている。各決定木は、利用可能なデータのブートストラップ試料に対してトレーニングされ、決定木の各ノードは、最も良好な説明変数 (つまり、遺伝子またはバイオマーカー) により分割される。ランダム・フォレスト法は、自動的な変数選択を提供することも、選択された変数間の非線形相互作用を記述することも、両方とも可能である。

#### 【0176】

段階的ロジスティック回帰法 (SLR) は、データ入力をロジスティック曲線にフィッティングすることにより、事象発生の確率を予測する統計モデルとみなされている。ロジスティック・モデルでは、所定のセットの説明変数 (つまり、遺伝子またはバイオマーカー) が、ロジスティクス・リンク関数により、確率に影響を及ぼすことが仮定されている。多数の候補変数から選択される説明変数の最適セットを決定するために、多数のロジスティック回帰モデルを、初期モデルから段階的な様式で構築し、最も正確なモデルを決定するために赤池の情報量基準 (AIC) を評価することにより比較する (Burnham, K.

P.およびD. R. Anderson、2002年、Model Selection and Multimodel Inference: A Practical-Theoretic Approach、第2版、Springer-Verlag社（非特許文献68））。

【0177】

サポート・ベクター・マシン（SVM）は、一般化された線形分類器のファミリーに属するとみなされている。SVMは、2グループ分類における入力データを、n次元空間における2セットのベクトルとしてとらえ、2セットのベクトル間のマージンを最大化する超平面によりデータを分離する。ベクトルは、最大化する超平面に対して最小距離をとり、サポート・ベクターと呼ばれる。SVMは、自動的な変数（つまり、遺伝子またはバイオマーカー）選択を提供しない。

【0178】

関連ベクター・マシン（RVM）では、所定のセットの説明変数（つまり、遺伝子またはバイオマーカー）が、ロジスティックス・リンク関数によりクラス・メンバーシップ確率に影響を及ぼすことが仮定されている。RVMは、多数の候補変数から選択される説明変数の最適セットを決定しようと試みる。RVMは、多数のRVMを評価および相互検証し、候補変数（つまり、遺伝子またはバイオマーカー）の最適セットを選択する遺伝的最適化アルゴリズムと共に作働させることができる。

【0179】

分類アルゴリズムを用いて構築された転写プロファイルは、前述のデータ分析アルゴリズムのうちの一つを使用して、さらにトレーニングされる。分類誤差は、トレーニングされた分類アルゴリズムがクラス内のメンバーシップを予測する、正確度の尺度である。分類誤差は、リーブ・ワン・アウト相互検証法（LOOCV、leave-one-out cross validation）、K分割検証法（K-fold validation）、または10分割検証法（ten-fold validation）などの相互検証法により決定することができる（Devijver, P. A.およびJ. Kittler, 1982年、Pattern Recognition: A Statistical Approach、Prentice-Hall社、ロンドン（非特許文献69））。

【0180】

所定の転写プロファイルを用いたアルゴリズムの正確度は、トレーニング中にそのアルゴリズムにより予測された真陽性（TP）、真陰性（TN）、偽陽性（FP）、および偽陰性（FN）の数を決定することにより測定することができる。正確度は、以下のように測定される：

$$\text{正確度} = (TP + TN) / TP + TN + FP + FN$$

【0181】

陽性的中率（PPV）、またはアルゴリズムにより陽性とスコアリングされた罹患対象の割合は、以下のように測定される：

$$PPV = TP / TP + FP$$

【0182】

陰性的中率（NPV）、またはアルゴリズムにより陰性とスコアリングされた対照対象（疾患を罹患していない）の割合は、以下のように測定される

$$NPV = TN / TN + FN$$

【0183】

分類アルゴリズムの性能は、その分類が正確な変数（つまり、遺伝子）を良好に識別した程度を評価するジャカード類似性係数（Jaccard similarity coefficient）（ジャカード指標）によっても決定される。トレーニングされた分類アルゴリズムの正確度は、約60%、65%、70%、75%、80%、85%、90%、または95%を上回ることができる。トレーニングされた分類アルゴリズムのジャカード指標は、約60%、65%、70%、75%、80%、85%、90%、または95%を上回ることができる。トレーニングされた分類アルゴリズムのPPVおよびNPVは、約60%、65%、70%、75%、80%、85%、90%、または95%を上回ることができる。

【0184】

対象の分類は、情動障害を罹患しているか、または情動障害の症状を示す可能性の高い

10

20

30

40

50

対象の診断に有用であってもよい。対象を分類するための遺伝子転写プロファイルは、表 1 A の遺伝子の転写解析に基づく。本明細書に記述されている方法により分析されるような対象の転写プロファイルは、対象が罹患対象のクラスに属するかどうかを示すだろう。

【0185】

いくつかの実施形態では、本発明は、試験対象の情動障害を診断する方法であって、該試験対象のバイオマーカー・プロファイル中の複数のバイオマーカーの複数の特徴が、ある値セットを満たすかどうかを評価し、該値セットを満たすことが、前記情動障害を該試験対象が罹患していることを予測し、該複数の特徴が、該複数のバイオマーカーの測定可能な態様であり、該複数のバイオマーカーが、表 1 A に列挙されている少なくとも 2 つのバイオマーカーを含む方法を提供する。本方法は、試験対象が情動障害を罹患しているかどうかの診断を、ユーザー・インターフェース・デバイス、モニター、実在するコンピュータ読み取り可能な記憶媒体、またはローカルもしくはリモート・コンピュータ・システムに出力すること；または試験対象が情動障害を罹患しているかどうかの診断を、ユーザー読み取り可能な形態で表示することをさらに含む。

10

【0186】

本発明のいくつかの実施形態では、複数のバイオマーカーは、表 1 A に列挙されている 2 ~ 29 個のバイオマーカーからなる。他の実施形態では、複数のバイオマーカーは、表 1 A に列挙されている 3 ~ 20 個のバイオマーカーからなる。さらに他の実施形態では、複数のバイオマーカーは、表 1 A に列挙されている少なくとも 2、3、4、または 5 個のバイオマーカーを含む。

20

【0187】

いくつかの実施形態では、複数の特徴は、表 1 A に列挙されている 2 ~ 29 個のバイオマーカーに対応する 2 ~ 29 個の特徴からなる。他の実施形態では、複数の特徴は、表 1 A に列挙されている 3 ~ 15 個のバイオマーカーに対応する 3 ~ 15 個の特徴からなる。さらに他の実施形態では、複数の特徴は、表 1 A に列挙されている少なくとも 2 個のバイオマーカーに対応する少なくとも 2 個の特徴を含む。

【0188】

他の実施形態では、複数のバイオマーカーは、ERK1 および MAPK14 を含む。他の実施形態では、複数のバイオマーカーは、Gi2 および IL-1b を含む。他の実施形態では、複数のバイオマーカーは、ARRB1 および MAPK14 を含む。他の実施形態では、複数のバイオマーカーは、ERK1 および IL1b を含む。

30

【0189】

本発明のいくつかの態様では、前記複数のバイオマーカー中の各バイオマーカーは、核酸である。他の態様では、前記複数のバイオマーカー中の各バイオマーカーは、DNA、cDNA、増幅 DNA、RNA、または mRNA である。さらに他の態様では、前記複数のバイオマーカー中の各バイオマーカーは、タンパク質である。

【0190】

他の実施形態では、試験対象のバイオマーカー・プロファイル中の前記複数の特徴中のある特徴は、複数のバイオマーカー中のバイオマーカーのある測定可能な態様であり、前記特徴の特徴値は、前記試験対象から採取された生体試料を使用して決定される。他の実施形態では、特徴は、生体試料中の前記バイオマーカーの含有量である。さらに他の実施形態では、生体試料は、末梢組織、全血、脳脊髄液、腹水、間質液、赤血球細胞、白血球細胞、または血小板である。

40

【0191】

別の実施形態では、前記複数の特徴中の特徴は、前記バイオマーカー・プロファイル中のバイオマーカーの測定可能な態様であり、前記特徴の特徴値は、前記試験対象から採取された試料を使用して決定される。いくつかの実施形態では、バイオマーカー・プロファイル中のバイオマーカーは、核酸の兆候またはタンパク質の兆候である。他の実施形態では、バイオマーカー・プロファイル中のバイオマーカーは、mRNA 分子の兆候または cDNA 分子の兆候である。いくつかの実施形態では、mRNA 分子または cDNA 分子の

50

兆候は、cDNA 1 ng 当たりのコピー数などの転写物値である。いくつかの実施形態では、バイオマーカー・プロファイル中の第1のバイオマーカーは、核酸の兆候であり、バイオマーカー・プロファイル中の第2のバイオマーカーは、タンパク質の兆候である。

【0192】

本発明のいくつかの態様では、値セットは、表4に示されているようなバイオマーカーの含有量を含み、表4の値セットを満たすことは、対象がうつ病を罹患していることを予測する。他の態様では、値セットは、表5に示されているようなバイオマーカーの含有量を含み、表5の値セットを満たすことは、対象が重度うつ病を罹患していることを予測する。他の態様では、値セットは、表6に示されているようなバイオマーカーの含有量を含み、表6の値セットを満たすことは、対象が双極性うつ病を罹患していることを予測する。さらに、本発明は、表4Aにあるようなうつ病診断用の値セット、および表5Bにあるような重度うつ病診断用の値セットを提供する。

10

【0193】

表4、5、および6に示されている値セットは、cDNA、つまりバイオマーカー遺伝子の転写物の1 ng 当たりのコピー数で表されるバイオマーカーの含有量により表される。例えば、表4にあるバイオマーカーARRB1についての、うつ病対象の転写物値の範囲は、 $189062 \pm 62727$  コピー/ng cDNAであり、それは、 $126335 \sim 251789$  コピー/ng cDNAの範囲と等しい。表4にあるバイオマーカーCD8aについての、うつ病対象の転写物値の範囲は、 $8304 \pm 5825$  コピー/ng cDNAであり、それは、 $2479 \sim 14129$  コピー/ng cDNAの範囲と等しい。本発明のいくつかの態様では、値セットを満たすことは、各バイオマーカーの所与の範囲内の値を有することを意味する。

20

【0194】

いくつかの実施形態では、1 ng のcDNA当たり $15148 \sim 35504$  コピーの範囲内のERK1含有量、および1 ng のcDNA当たり $39241 \sim 107071$  コピーの範囲内のMAPK14含有量を含む値セットは、対象がうつ病を罹患していることを予測する。他の実施形態では、1 ng のcDNA当たり $61734 \sim 168500$  コピーの範囲内のGi2含有量、および1 ng のcDNA当たり $15939 \sim 43323$  コピーの範囲内のIL1b含有量を含む値セットは、対象がうつ病を罹患していることを予測する。他の実施形態では、1 ng のcDNA当たり $126335 \sim 251789$  コピーの範囲内のARRB1含有量、および1 ng のcDNA当たり $39241 \sim 107071$  コピーの範囲内のMAPK14含有量を含む値セットは、対象がうつ病を罹患していることを予測する。他の実施形態では、1 ng のcDNA当たり $15148 \sim 35504$  コピーの範囲内のERK1含有量、および1 ng のcDNA当たり $15939 \sim 43323$  コピーの範囲内のIL1b含有量を含む値セットは、対象がうつ病を罹患していることを予測する。

30

【0195】

他の実施形態では、 $0.25 \sim 0.45$  の範囲内にある、MAPK14含有量で除算されたERK1含有量の比率を含む値セットは、対象がうつ病を罹患していることを予測する。他の実施形態では、 $0.16 \sim 0.36$  の範囲内にある、IL1b含有量で除算されたGi2含有量の比率を含む値セットは、対象がうつ病を罹患していることを予測する。他の実施形態では、 $0.29 \sim 0.49$  の範囲内にある、ARRB1含有量で除算されたMAPK14含有量の比率を含む値セットは、対象がうつ病を罹患していることを予測する。他の実施形態では、 $0.075 \sim 0.95$  の範囲内にある、IL1b含有量で除算されたERK1含有量の比率を含む値セットは、対象がうつ病を罹患していることを予測する。

40

【0196】

他の実施形態では、 $0.19 \sim 0.39$  の範囲内にある、MAPK14含有量で除算されたERK1含有量の比率を含む値セットは、対象が重度うつ病を罹患していることを予測する。他の実施形態では、 $0.18 \sim 0.38$  の範囲内にある、IL1b含有量で除算

50

された G i 2 含有量の比率を含む値セットは、対象が重度うつ病を罹患していることを予測する。他の実施形態では、0.32 ~ 0.52 の範囲内にある、A R R B 1 含有量で除算された M A P K 1 4 含有量の比率を含む値セットは、対象が重度うつ病を罹患していることを予測する。他の実施形態では、0.60 ~ 0.80 の範囲内にある、I L 1 b 含有量で除算された E R K 1 含有量の比率を含む値セットは、対象が重度うつ病を罹患していることを予測する。

【0197】

上記の方法の他の態様では、本方法は、評価ステップに先立って、前記バイオマーカー・プロファイルを構築することをさらに含む。他の実施形態では、構築ステップは、前記試験対象の生体試料から前記複数の特徴を取得することを含む。いくつかの態様では、バイオマーカー・プロファイルは、第2のバイオマーカーの特徴値で第1のバイオマーカーの特徴値を除算することによるバイオマーカー含有量の比率を決定することにより構築される。そのようなバイオマーカー・プロファイルは、表4、表5、または表6に示される値を使用して構築することができる。

10

【0198】

他の実施形態では、試料は、末梢組織、全血、脳脊髄液、腹水、間質液、赤血球細胞、白血球細胞、または血小板である。

【0199】

上記の方法のさらに他の態様では、本方法は、評価ステップに先立って、前記第1の値セットを構築することをさらに含む。他の実施形態では、構築ステップは、データ分析アルゴリズムを、集団のメンバーから得られた特徴に適用することを含む。

20

【0200】

いくつかの態様では、特徴は、E R K 1 および M A P K 1 4 を含むバイオマーカーの測定可能な態様であり、特徴値は、前記試験対象から採取された血液試料を使用して決定される。

【0201】

他の実施形態では、集団は、情動障害を罹患していない第1の複数の対照対象に由来する第1の複数の生体試料、および情動障害を罹患している第2の複数の対象に由来する第2の複数の生体試料を含む。さらに他の実施形態では、データ分析アルゴリズムは、決定木、マイクロアレイ予測解析、多重加法型回帰木、ニューラル・ネットワーク、クラスタリング・アルゴリズム、主成分分析法、最近傍解析、線形判別分析、二次判別分析、サポート・ベクター・マシン法、進化法 (evolutionary method)、関連ベクター・マシン法、遺伝的アルゴリズム、射影追跡法、または加重投票法である。

30

【0202】

別の実施形態では、構築ステップは、判断ルールを生成し、そこでは、前記評価ステップが、前記判断ルールを複数の特徴に適用して、それらが第1の値セットを満たすかどうかを決定することを含む。いくつかの実施形態では、判断ルールは、前記集団中の対象を、(i) 情動障害を罹患していない対象、および(ii) 70パーセント以上の正確度で情動障害を罹患している対象として分類する。いくつかの実施形態では、判断ルールは、前記集団中の対象を、(i) 情動障害を罹患していない対象、および(ii) 90パーセント以上の正確度で情動障害を罹患している対象として分類する。

40

【0203】

本発明のある態様では、情動障害は、双極性障害 I、双極性障害 II、気分変調性障害、または抑うつ障害である。他の態様では、情動障害は、軽度うつ病、中程度のうつ病、重度うつ病、非定型うつ病、メランコリー型うつ病、または境界性人格障害である。さらに他の態様では、情動障害は、(i) 心的外傷後ストレス障害、または(ii) 心的外傷後ストレス障害を伴わない心的外傷。いくつかの態様では、情動障害は、急性心的外傷後ストレス障害、または軽減された心的外傷後ストレス障害である。

【0204】

本発明は、試験対象の情動障害を診断するために使用されるキットであって、該試験対

50

象のバイオマーカー・プロファイル中の複数のバイオマーカーの複数の特徴が、ある値セットを満たすかどうかを評価するための試薬および説明書を含み、該値セットを満たすことが、前記情動障害を該試験対象が罹患していることを予測し、該複数の特徴が、該複数のバイオマーカーの測定可能な態様であり、該複数のバイオマーカーが、表 1 A に列挙されている少なくとも 2 つのバイオマーカーを含むキットを提供する。いくつかの態様では、試薬は、表 1 A から選択されるバイオマーカーのヌクレオチド配列を認識するプローブおよび/またはプライマーを含む。本発明のキットは、本発明によるバイオマーカー・プロファイルを生成するために使用される。いくつかの態様では、本発明のキットは、表 1 A に列挙されている少なくとも 2 つのバイオマーカーを含む複数のバイオマーカーから、試験対象のバイオマーカー・プロファイルを試験および評価するための説明書を提供する。他の態様では、本発明のキットは、試験対象のバイオマーカー・プロファイルがそのような値セットを満たすかどうかを決定するために、値セットを含む説明書を提供する。

10

20

30

40

50

**【0205】**

本発明は、コンピュータ・プログラム製品も提供し、該コンピュータ・プログラム製品は、コンピュータ読み取り可能な記憶媒体およびそこに内蔵されているコンピュータ・プログラム・メカニズムを含み、該コンピュータ・プログラム・メカニズムは、上記の方法のいずれかを実行するためのインストラクションを含む。いくつかの態様では、コンピュータ・プログラム・メカニズムは、試験対象が情動障害を罹患しているかどうかの診断を、ユーザー・インターフェース・デバイス、モニター、実在するコンピュータ読み取り可能な記憶媒体、またはローカルもしくはリモート・コンピュータ・システムに出力するためのインストラクション；または試験対象が情動障害を罹患しているかどうかの診断を、ユーザー読み取り可能な形態で表示するためのインストラクションをさらに含む。

**【0206】**

本発明は、1つまたは複数のプロセッサ、および該1つまたは複数のプロセッサに接続されたメモリを含み、該メモリが上記の方法のいずれかを実行するためのインストラクションを格納するコンピュータも提供する。本発明のいくつかの態様では、メモリは、試験対象が情動障害を罹患しているかどうかの診断を、ユーザー・インターフェース・デバイス、モニター、実在するコンピュータ読み取り可能な記憶媒体、またはローカルもしくはリモート・コンピュータ・システムに出力するためのインストラクション；または試験対象が情動障害を罹患しているかどうかの診断を、ユーザー読み取り可能な形態で表示するためのインストラクションをさらに含む。

**【0207】**

本発明は、試験対象が情動障害の症状を示す可能性を決定する方法であって、該試験対象のバイオマーカー・プロファイル中の複数のバイオマーカーの複数の特徴が、ある値セットを満たすかどうかを評価し、該値セットを満たすことが、情動障害の症状を該試験対象が示す可能性を提供し、該複数の特徴が、該複数のバイオマーカーの測定可能な態様であり、該複数のバイオマーカーが、表 1 A に列挙されている少なくとも 2 つのバイオマーカーを含む方法をさらに提供する。

**【0208】**

いくつかの実施形態では、複数のバイオマーカーは、ERK1 および MAPK14 を含む。他の実施形態では、複数のバイオマーカーは、Gi2 および IL-1b を含む。他の実施形態では、複数のバイオマーカーは、ARRB1 および MAPK14 を含む。他の実施形態では、複数のバイオマーカーは、ERK1 および IL1b を含む。

**【0209】**

本発明のいくつかの実施形態では、複数のバイオマーカーは、ERK1、PBR、および MAPK14 を含む。別の実施形態では、複数のバイオマーカーは、PBR、Gi2、および IL1b を含む。他の実施形態では、複数のバイオマーカーは、ERK1、ARRB1、および MAPK14 を含む。いくつかの実施形態では、複数のバイオマーカーは、MAPK14、ERK1、および CD8b を含む。他の実施形態では、複数のバイオマーカーは、MAPK14、ERK1、および P2X7 を含む。さらに他の実施形態では、複

数のバイオマーカーは、ARRB1、IL6、およびCD8aを含む。ある実施形態では、複数のバイオマーカーは、ARRB1、ODC1、およびP2X7を含む。

【0210】

さらに他の実施形態では、本方法は、試験対象が情動障害の症状を示す可能性を、ユーザー・インターフェース・デバイス、モニター、実在するコンピュータ読み取り可能な記憶媒体、またはローカルもしくはリモート・コンピュータ・システムに出力すること；または試験対象が情動障害の症状を示す可能性を、ユーザー読み取り可能な形態で表示することをさらに含む。

【0211】

本発明は、複数の対照対象から収集された各生体試料に関する転写解析の尺度である転写プロファイルを提供する。本発明は、複数のうつ病対象、重度うつ病対象、または双極性対象から収集された各生体試料に関する転写解析の尺度である転写プロファイルを提供する。本発明は、複数の境界性人格障害対象から収集された各生体試料に関する転写解析の尺度である転写プロファイルをさらに提供する。本発明は、複数のPTSD対象から収集された各生体試料に関する転写解析の尺度である転写プロファイルを提供する。

10

【0212】

本発明は、第1の複数対照対象の集団的尺度を含む転写プロファイルが、例えばデータベースに格納されることも提供する。第2の複数対象、例えば罹患対象の集合的尺度を含む転写プロファイルは、データ分析アルゴリズム、特にトレーニングされた分類アルゴリズムを使用して、第1の複数対照対象の転写プロファイルと比較される。トレーニングされた分類アルゴリズムは、各セットの対象を分類する。トレーニングされた分類アルゴリズムは、分類を診断および割当てするのに有用な予測値を提供する。トレーニングされた分類アルゴリズムは、対象が障害の症状を示す可能性を予測するのに有用な予測値を提供する。

20

【0213】

本発明の別の実施形態は、健常対照対象および罹患対象の転写プロファイルと比較した対象の個別的な転写プロファイルに基づいて、疾患または障害に対する対象の感受性を診断または予測すること、または障害の症状を示す可能性を予測することに関する。診断に使用するための遺伝子転写プロファイルは、表1Aから選択される遺伝子転写解析に基づく。

30

【0214】

本発明の1つの態様は、異なるタイプの情動障害、特に大うつ病性障害、双極性障害、境界性人格障害、および心的外傷後ストレス障害の診断に関する。

【0215】

本発明の別の態様は、転写プロファイルを識別することにより患者集団を区別することに関する。例えば、通常、大うつ病であると診断されるだろう患者は、転写プロファイルにより、例えばメランコリー型うつ病および非定型うつ病などのうつ病のサブタイプに区別することができる。これらのうつ病サブタイプの鑑別治療応答に関する根拠は存在する。共存症を示す患者、つまり複数の障害に関するDSM-IV（登録商標）判断基準を満たす患者は、転写プロファイルの識別から利益を得るだろう。転写プロファイルは、1つの障害に関する共通の生物学的基盤を識別することができる。

40

【0216】

1つの実施形態では、本発明は、上記の方法により、複数の健常対照対象から収集された生体試料に関する転写解析の尺度である転写プロファイルを提供する。本発明は、複数の情動障害対象から収集された生体試料に関する転写解析の尺度である転写プロファイルも提供する。例えば、本発明は、複数のうつ病対象、重度うつ病対象、または双極性対象から収集された生体試料に関する転写解析の尺度である転写プロファイルも提供する。本発明は、表4にあるような、複数のうつ病対象から収集された生体試料に関する転写解析の尺度である転写プロファイルを提供する。本発明は、表5にあるような、複数の重度うつ病対象から収集された生体試料に関する転写解析の尺度である転写プロファイルを提供

50

する。本発明は、表6にあるような、複数の双極性対象から収集された生体試料に関する転写解析の尺度である転写プロファイルも提供する。本発明は、複数の境界性人格障害対象から収集された生体試料に関する転写解析の尺度である転写プロファイルをさらに提供する。本発明は、複数のPTSD対象から収集された生体試料に関する転写解析の尺度である転写プロファイルを提供する。

【0217】

本発明の1つの実施形態では、生体試料は全血である。

【0218】

本発明は、第1の複数対照対象の集団的尺度を含む転写プロファイルが、例えばデータベースに格納されることも提供する。第2の複数対象、例えば罹患対象の集合的尺度を含む転写プロファイルは、分類アルゴリズムを使用して、第1の複数対照対象の転写プロファイルと比較される。分類アルゴリズムは、対象の各々を分類する出力を提供する。

10

【0219】

本発明のいくつかの態様では、転写プロファイルは、ADA、ARRB1、ARRB2、CD8a、CD8b、CREB1、CREB2、DPP4、ERK1、ERK2、Gi2、Gs、GR、IL1b、IL6、IL8、INDO、MAPK14、MAPK8、MKP1、MR、ODC1、P2X7、PBR、PREP、RGS2、S100A10、SERT、およびVMAT2からなる群から選択される遺伝子の転写解析から決定される。

【0220】

別の実施形態では、転写プロファイルは、ADA、ARRB1、ARRB2、CD8a、CD8b、CREB1、CREB2、DPP4、ERK1、ERK2、Gi2、Gs、GR、IL1b、IL6、IL8、INDO、MAPK14、MAPK8、MKP1、MR、ODC1、P2X7、PBR、PREP、RGS2、S100A10、SERT、およびVMAT2からなる群から選択される少なくとも3つ遺伝子の転写解析から決定される。

20

【0221】

いくつかの実施形態では、転写プロファイルは、ARRB1、ARRB2、CD8a、CREB1、CREB2、ERK2、Gi2、MAPK14、ODC1、P2X7、およびPBRからなる群から選択される遺伝子の転写解析から決定される。

【0222】

別の実施形態では、転写プロファイルは、CD8a、ERK1、MAPK14、P2X7、およびPBRからなる群から選択される遺伝子の転写解析から決定される。

30

【0223】

別の実施形態では、転写プロファイルは、Gi2、GR、およびMAPK14からなる群から選択される遺伝子の転写解析から決定される。

【0224】

別の実施形態では、転写プロファイルは、Gi2、GR、MAPK14、およびMRからなる群から選択される遺伝子の転写解析から決定される。

【0225】

別の実施形態では、転写プロファイルは、ARRB1、ARRB2、CD8b、ERK2、IDO、IL-6、MR、ODC1、PREP、およびRGS2からなる群から選択される遺伝子の転写解析から決定される。

40

【0226】

別の実施形態では、転写プロファイルは、ARRB1、CREB1、ERK2、Gs、IL-6、MKP1、およびRGS2からなる群から選択される遺伝子の転写解析から決定される。

【0227】

別の実施形態では、転写プロファイルは、ERK1およびMAPK14からなる群から選択される遺伝子の転写解析から決定される。別の実施形態では、転写プロファイルは、Gi2およびIL1bからなる群から選択される遺伝子の転写解析から決定される。別の

50

実施形態では、転写プロファイルは、ARRB1およびMAPK14からなる群から選択される遺伝子の転写解析から決定される。別の実施形態では、転写プロファイルは、ERK1およびIL1bからなる群から選択される遺伝子の転写解析から決定される。

【0228】

別の実施形態では、転写プロファイルは、ERK1、MAPK14、およびP2X7からなる群から選択される遺伝子の転写解析から決定される。別の実施形態では、転写プロファイルは、Gi2、IL1b、およびPBRからなる群から選択される遺伝子の転写解析から決定される。別の実施形態では、転写プロファイルは、ARRB1、ODC1、およびP2X7からなる群から選択される遺伝子の転写解析から決定される。別の実施形態では、転写プロファイルは、ARRB1、CD8a、およびIL6からなる群から選択される遺伝子の転写解析から決定される。別の実施形態では、転写プロファイルは、CD8b、ERK1、およびMAPK14からなる群から選択される遺伝子の転写解析から決定される。別の実施形態では、転写プロファイルは、ARRB1、ERK1、およびMAPK14からなる群から選択される遺伝子の転写解析から決定される。別の実施形態では、転写プロファイルは、ERK1、MAPK14、およびPBRからなる群から選択される遺伝子の転写解析から決定される。

10

【0229】

本発明のある態様は、対象の転写プロファイルを識別すること、およびそのような転写プロファイルを、対照対象または健常対照対象のグループのプロファイルと比較することを含む対象の情動障害を診断する方法を提供し、それにより転写プロファイルの変化または相異の存在または非存在に基づいて、対象が情動障害を示すかどうかを診断する。

20

【0230】

本発明のいくつかの実施形態では、情動障害は、うつ病、重度うつ病、双極性障害、境界性人格障害からなる群から選択される。いくつかの実施形態では、情動障害は、心的外傷後ストレス障害、または心的外傷後ストレス障害を伴わない心的外傷から選択される。他の実施形態では、情動障害は、急性心的外傷後ストレス障害、または軽減された心的外傷後ストレス障害から選択される。

【0231】

本発明の1つの態様は、対象が情動障害を示すかどうかを診断するための方法であって、

30

(a) 情動障害を罹患している疑いのある対象から生体試料を取得すること、

(b) 該生体試料中のmRNAレベルを測定し、該mRNAレベルが、ADA、ARRB1、ARRB2、CD8a、CD8b、CREB1、CREB2、DPP4、ERK1、ERK2、Gi2、Gs、GR、IL1b、IL6、IL8、INDO、MAPK14、MAPK8、MKP1、MR、ODC1、P2X7、PBR、PREP、RGS2、S100A10、SERT、およびVMAT2からなる群から選択される遺伝子のmRNAレベルであること、

(c) 該mRNAのレベルを、mRNAデータとしてコンピュータ媒体に収集および格納すること、

(d) そのようなmRNAデータを分類アルゴリズムにより処理し、該処理が、健常対照対象のmRNAデータと該mRNAデータが同じであるかまたは異なっているかを決定すること、および

40

(e) 該対象を分類する出力データを提供することを含み、

それにより、該対象が情動障害を示すかどうかを診断する方法を提供する。

【0232】

本発明は、ADA、ARRB1、ARRB2、CD8a、CD8b、CREB1、CREB2、DPP4、ERK1、ERK2、Gi2、Gs、GR、IL1b、IL6、IL8、INDO、MAPK14、MAPK8、MKP1、MR、ODC1、P2X7、PBR、PREP、RGS2、S100A10、SERT、およびVMAT2からなる群から選択される遺伝子の、対象の転写プロファイルを、複数の健常対照対象の前記遺伝子の転

50

写プロファイルと比較することにより、情動障害に対する該対象の感受性を予測するための方法をさらに提供する。

【0233】

本発明の1つの態様は、対象が情動障害の症状を示す可能性を予測するための方法であって、

(a) 対象から生体試料を取得すること、

(b) mRNAレベルを測定し、該mRNAレベルが、ADA、ARRB1、ARRB2、CD8a、CD8b、CREB1、CREB2、DPP4、ERK1、ERK2、Gi2、Gs、GR、IL1b、IL6、IL8、INDO、MAPK14、MAPK8、MKP1、MR、ODC1、P2X7、PBR、PREP、RGS2、S100A10、SERT、およびVMAT2からなる群から選択される遺伝子のmRNAレベルであること、

(c) 該mRNAのレベルを、mRNAデータとしてコンピュータ媒体に収集および格納すること、

(d) そのようなmRNAデータを分類アルゴリズムにより処理し、該処理が、健常対照対象のmRNAデータと該mRNAデータが同じであるかまたは異なっているかを決定すること、および

(e) 該対象を分類する出力データを提供することを含み、

それにより、該対象が情動障害の症状を示す可能性を予測する方法を提供する。

【0234】

別の実施形態では、本方法は、ADA、ARRB1、ARRB2、CD8a、CD8b、CREB1、CREB2、DPP4、ERK1、ERK2、Gi2、Gs、GR、IL1b、IL6、IL8、INDO、MAPK14、MAPK8、MKP1、MR、ODC1、P2X7、PBR、PREP、RGS2、S100A10、SERT、およびVMAT2からなる群から選択される少なくとも2つ遺伝子のmRNAレベルを測定することを含んでいてもよい。

【0235】

他の実施形態では、本方法は、表1Aに列挙されている任意の3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、または28個の遺伝子のmRNAレベルを測定することを含む。

【0236】

他の実施形態では、本方法は、ARRB1、ARRB2、CD8a、CREB1、CREB2、ERK2、Gi2、MAPK14、ODC1、P2X7、およびPBRからなる群から選択される遺伝子のmRNAレベルを測定することを含む。

【0237】

別の実施形態では、本方法は、CD8a、ERK1、MAPK14、P2X7、およびPBRからなる群から選択される遺伝子のmRNAレベルを測定することを含む。

【0238】

別の実施形態では、本方法は、Gi2、GR、およびMAPK14からなる群から選択される遺伝子のmRNAレベルを測定することを含む。

【0239】

別の実施形態では、本方法は、Gi2、GR、MAPK14、およびMRからなる群から選択される遺伝子のmRNAレベルを測定することを含む。

【0240】

別の実施形態では、本方法は、ARRB1、ARRB2、CD8b、ERK2、IDO、IL-6、MR、ODC1、PREP、およびRGS2からなる群から選択される遺伝子のmRNAレベルを測定することを含む。

【0241】

別の実施形態では、本方法は、ARRB1、CREB1、ERK2、Gs、IL-6、

10

20

30

40

50

M K P 1、および R G S 2 からなる群から選択される遺伝子の m R N A レベルを測定することを含む。

【 0 2 4 2 】

別の実施形態では、本方法は、E R K 1 および M A P K 1 4 からなる群から選択される遺伝子の m R N A レベルを測定することを含む。別の実施形態では、本方法は、G i 2 および I L 1 b からなる群から選択される遺伝子の m R N A レベルを測定することを含む。別の実施形態では、本方法は、A R R B 1 および M A P K 1 4 からなる群から選択される遺伝子の m R N A レベルを測定することを含む。別の実施形態では、本方法は、E R K 1 および I L 1 b からなる群から選択される遺伝子の m R N A レベルを測定することを含む。

10

【 0 2 4 3 】

別の実施形態では、本方法は、E R K 1、M A P K 1 4、および P 2 X 7 からなる群から選択される遺伝子の m R N A レベルを測定することを含む。別の実施形態では、本方法は、G i 2、I L 1 b、および P B R からなる群から選択される遺伝子の m R N A レベルを測定することを含む。別の実施形態では、本方法は、A R R B 1、O D C 1、および P 2 X 7 からなる群から選択される遺伝子の m R N A レベルを測定することを含む。別の実施形態では、本方法は、A R R B 1、C D 8 a、および I L 6 からなる群から選択される遺伝子の m R N A レベルを測定することを含む。別の実施形態では、本方法は、C D 8 b、E R K 1、および M A P K 1 4 からなる群から選択される遺伝子の m R N A レベルを測定することを含む。別の実施形態では、本方法は、A R R B 1、E R K 1、および M A P K 1 4 からなる群から選択される遺伝子の m R N A レベルを測定することを含む。別の実施形態では、本方法は、E R K 1、M A P K 1 4、および P B R からなる群から選択される遺伝子の m R N A レベルを測定することを含む。

20

【 0 2 4 4 】

本発明のいくつかの実施形態では、情動障害は、うつ病、重度うつ病、双極性障害、境界性人格障害からなる群から選択される。いくつかの実施形態では、情動障害は、心的外傷後ストレス障害、または心的外傷後ストレス障害を伴わない心的外傷から選択される。他の実施形態では、情動障害は、急性心的外傷後ストレス障害、または軽減された心的外傷後ストレス障害から選択される。

【 0 2 4 5 】

いくつかの実施形態では、上記の方法は、コンピュータに支援された方法である。

30

【 0 2 4 6 】

5 . 7 情動障害

本明細書中に記述されている精神医学的障害または精神障害、およびそれらの臨床徴候は、開業精神科医に公知である。ほとんどの精神科医は、各障害の特有な症状を認識することができる。

【 0 2 4 7 】

米国精神医学会により出版された The Diagnostic and Statistical Manual of Mental Disorders、第4版、改訂版 (DSM-IV-TR (登録商標) (1994年10月、2000年5月改訂) (非特許文献70) は、米国の医師が使用する精神障害の臨床分類用基準である。精神障害 / 精神医学的障害の症候学および診断判断基準は、D S M - I V - T R (登録商標) ガイドラインに示されている。

40

【 0 2 4 8 】

5 . 7 . 1 抑うつ障害

D S M - I V - T R (登録商標) には、うつ病および大うつ病性障害 (M D D) の特定の診断判断基準が列挙されている。

【 0 2 4 9 】

D S M - I V - T R (登録商標) では、同じ2週間の期間中に、少なくとも以下の症状のうち5つが、以前の適切に機能していた状態からの変化として提示および顕在化する (さらに、(1) または (2) のいずれかが症状に含まれていなければならない) 症候群と

50

して、大うつ病エピソードが定義されている。：

- 1．抑うつ気分
- 2．関心または愉悦の低下
- 3．著しい体重の減少または増加
- 4．不眠症または過眠症
- 5．精神運動性の激越または遅滞
- 6．疲労または活力喪失
- 7．自己否定感情
- 8．思考力または集中力の低下；優柔不断
- 9．死について繰り返し考えること、自殺念慮、自殺企図、または自殺の具体的な計画

10

#### 【0250】

DSM-IV-TR（登録商標）には、種々のサブタイプのうつ病に存在しているはずの症状の説明がさらに含まれている。うつ病は、精神病の症状を示す場合があり、または示さない場合もあり、メランコリー性の特徴を示す場合があり、または緊張性の特徴を示す場合もあり、非定型うつ病として分類される場合があることが指摘されている。

#### 【0251】

患者が示す症状の数および重症度に応じて、うつ病エピソードは、軽度、中程度、重度として指定することができる。臨床医は、患者が、定型（メランコリー型）、非定型、緊張性、または精神病性のうつ病を罹患しているかどうかを決定することができる。

#### 【0252】

臨床的には、うつ病は、異種混合性が非常に高い疾患であると考えられている。うつ病患者の遺伝子発現プロファイルは、この異種混合性を反映する場合がある。本発明に基づくと、患者をより良好に分類または診断するために、遺伝子発現プロファイルに基づき、うつ病のこれらのサブタイプをより良好に定義することが可能である。結果的に、薬物の開発および投与を、うつ病のサブタイプを罹患している患者に個別化することができる。

20

#### 【0253】

病歴および症状に関する情報を対照から取得および分析することにより、遺伝子発現プロファイルを使用して、本明細書に記述されている障害の症状を対象が示す可能性も予測される。

#### 【0254】

抑うつ障害、双極性障害、および気分変調性障害は、気分障害の種類の一部と考えられている。

30

#### 【0255】

本主題発明は、軽度、中程度、または重度うつ病などの抑うつ障害を示す転写プロファイルの客観的な尺度を提供する。本主題発明は、抑うつ障害のサブタイプを分類するための転写プロファイルも提供する。本発明は、軽度、中程度、重度うつ病などの抑うつ障害の対象を診断するための方法をさらに提供する。

#### 【0256】

##### 5.7.2 双極性障害

うつ病について記述されているように、双極性障害（BD）は、異種混合性の疾患であり、双極性I、双極性II、および循環気質を含む下位区分またはサブタイプに区分される。双極性障害は、躁うつ病としても知られており、個人の気分、活力、および機能能力に異常な変調を引き起こす脳障害である。個人が皆経験する正常な「気分の起伏」とは異なり、双極性障害の症状は重度であり、その結果として人間関係への悪影響、仕事または学校の成績不良、および自殺さえももたらす場合がある。

40

#### 【0257】

BDは、典型的には、人の生涯にわたって繰り返えられる躁およびうつ断続的エピソードとして現れる。エピソード間では、双極性障害のほとんどの人々は、無症状であるか、または何らかの残留症状を示す場合がある。うつ病エピソードは、存在することが多く、激しい（major）場合もあり、または重度である場合もある。躁病エピソードは、DS

50

M - I V - T R (登録商標)により記述されているように、仕事における障害または患者もしくは他人に対する危険を引き起こすのに十分であり、物質乱用または医学的状態の結果ではない著しい気分障害、睡眠要求の減少、過度の多弁、談話心迫、および/または思考の奔逸 (racing thought) もしくは観念奔逸などの症状を特徴とする。

【0258】

本発明は、双極性障害の対象を診断するための方法を提供する。BD患者は、双極性障害を示す転写プロファイルの客観的尺度から利益を得るだろう。

【0259】

#### 5.7.3 境界性人格障害

境界性人格障害 (BPD) は、自己像、対人関係、および情動に関する、著しく衝動的な不安定性のパターンを含む。この不安定性は、家族および仕事生活および個人の自己同一性を分裂させることが多い。

10

【0260】

DSM - I V - T R (登録商標)では、BPDは、下記のうちの少なくとも5つにより示されると特徴付けられている：

1. 過度の理想化および否定の両極端が交互に訪れることを特徴とする不安定で極端な対人関係のパターン。
2. 潜在的に自己損傷的である、例えば、浪費、セックス、物質使用、万引き、無謀運転、または過食のうち少なくとも2つの区分における衝動。
3. 著しい気分の反応性による情動不安定性。
4. 不適切で極端な怒り、または怒りの抑制の欠如、例えば、かんしゃく、怒りの持続、物理的な喧嘩の反復を頻繁に示すこと。
5. 自殺威嚇、身振り、もしくは拳動の反復、または自傷行為。
6. 同一性障害；著しく持続的に不安定な自己像。
7. 慢性的な空虚感または退屈感。
8. 現実または想像による放棄を回避するための尋常でない努力。
9. 一時的なストレス関連の妄想様観念または重度の解離症状。

20

【0261】

BPDの患者は、心理療法で見られる最も挑戦的で耐治療性の患者のうちに入る。

【0262】

本発明は、BPDの対象を診断するための方法を提供する。BPD患者は、境界性人格障害を示す転写プロファイルの客観的尺度から利益を得るだろう。

30

【0263】

#### 5.7.4 心的外傷後ストレス障害 (PTSD)

DSM - I V - T R (登録商標)では、心的外傷後ストレス障害は、実際の死もしくは重篤な外傷または死もしくは重篤な外傷への脅威を伴う事象を個人的に直接経験することを含む、極端な心的外傷ストレス要因に曝された後で、特徴的症状を発症することとして記述されている。そうした個人は、死、外傷、または別の人の物理的な完全性に対する脅威を伴う事象を目撃したことがある。そうした事象に対する個人の反応は、極度の恐怖、無力感、または戦慄を含む。そうした個人は、映像、観念、または知覚を含む、執拗な事象の想起を繰り返すか、または事象の悪夢を繰り返す場合がある。

40

【0264】

本発明は、急性PTSD、軽減されたPTSD、またはPTSDを伴わない心的外傷を罹患している対象を診断するための方法を提供する。患者/対象は、急性PTSD、軽減されたPTSD、またはPTSDを伴わない心的外傷を示す転写プロファイルの客観的な尺度から利益を得るだろう。

【0265】

上記に記述されている方法により特定された転写プロファイルに基づき、正常な健常対象と、情動障害を罹患している対象とを決定、識別、および/または区別することが可能である。例として、本発明は、以下の実験の詳細によって、より良好に理解されるだろう

50

。当業者であれば、特定の方法およびそこで考察されている結果は、以下に続く特許請求の範囲において、より完全に記述されている本発明の例示に過ぎないことを直ちに理解するだろう。

#### 【実施例】

#### 【0266】

##### 6 実験の詳細

全RNAの単離。ヒト血液を、PAXgene（商標）血液RNAチューブ（PreAnalytiX社製、ホンブレヒティコン（Hombrechtikon）、スイス）に収集し、何回か反転させて混合し、RNA単離の処理をするまで - 20 または - 80 で保管した。処理は、試料を室温で終夜インキュベートすることにより開始し、その後3000×Gで10分間遠心分離した。上清をデカントし、ペレットを5mlの水に再懸濁し、その後さらなる遠心分離ステップを行った。洗浄および遠心分離ステップを2回繰り返し、ペレットを、チューブに残っている残留水（約100ul）に再懸濁した。この溶液に、941μlのAmbion社製TOTALLY RNA（商標）溶解/変性溶液（Ambion社製、オースティン、テキサス州）および59μlの3M酢酸ナトリウム、pH5.5（Ambion社製）を添加し、その後混合した。室温で15分間インキュベーションした後、770μlの酸性フェノール/クロロホルム（Ambion社製）を添加し、チューブをボルテックスして混合した。溶液を、2mlのプラスチック・スクリーキャップ・チューブに移し、室温で5分間インキュベートした。フェノール抽出物を、最大速度（およそ13,000×G）で1分間、微量遠心機で遠心し、水層（1100μl）を、550μlの100%エタノールを含有する新しいチューブに移した。混合した後、溶液を、Ambion社製RNAqueous（登録商標）-96自動化キットろ過プレートの1つのウェルにアプライし、製造業者のプロトコールに従ってRNAを精製した。RNAを溶出した後、試料を、DNase I（Invitrogen社製、カールズバッド、カリフォルニア州）で2回処理して、残基ゲノムDNAを除去した。3ユニットの酵素を加えた1×DNase消化緩衝液中で、RNAを室温で1時間インキュベートした。終濃度が13mMになるようにEDTAを添加し、その後68 で10分間加熱することにより酵素を不活化した。MultiScreen（登録商標）PCRマイクロ<sub>96</sub>プレート（Millipore社製、ビルリカ、マサチューセッツ州）を通過させて混合物を脱塩し、50μlの水で溶出した。1μl等量のRNAを、Agilent社製2100バイオアナライザー（Agilent社製、バルトブロン、ドイツ）で分析し、残りを - 80 で保管した。RNA試料の品質は、バイオアナライザー・ソフトウェアにより計算されたRIN値を使用して評価した。

10

20

30

#### 【0267】

##### cDNA合成

cDNAの合成は、16.5μlの最終容積中で、およそ1μgの全RNAを1.5μlのランダム六量体（Invitrogen社製、500ng/μl）と混合することにより達成した。75 で10分間および25 で10分間インキュベーションした後、6μlの第1鎖緩衝液（Invitrogen社製）、1.5μlの10mM dNTP（Invitrogen社製、各dNTP 10mM）、1.25μlのSuperscript II（商標）（Invitrogen社製、200単位/μl）、および4μlの水を添加した。最終反応容積は30μlであり、インキュベーションは、25 で10分間、42 で1時間、および95 で10分間実行した。70μlの水を添加するまで、反応物を4 に冷却し、その後MultiScreen（登録商標）PCRマイクロ<sub>96</sub>プレートを用いて精製した。100μlの水を用いてcDNAの溶出を実行し、その結果生じた物質を、定量化するまで - 20 で保管した。ある場合には、cDNA反応の容積を2倍にして、物質の収量を増加させた。

40

#### 【0268】

##### cDNAの定量化

染料インターカレーション・アッセイを使用して、cDNA収量を決定した。5μlの

50

cDNAを、47  $\mu$ lの最終容積中で、7  $\mu$ lの0.5N NaOH、50mM EDTAと混合する。混合物を65 で1時間インキュベートしてRNAを加水分解し、その後10  $\mu$ lの1Mトリス、pH7を添加することにより中和した。25  $\mu$ l等量の加水分解反応液中のcDNA濃度は、製造業者の説明書に従ってQuant-iT (商標) Oligreen (登録商標) ssDNA試薬 (Invitrogen社製) を使用して測定した。未知試料を、既知濃度の一本鎖DNAを使用して生成した検量曲線と比較した。蛍光測定はすべて、Fusion (商標) アルファ装置 (パッカー社製、メリダン (Meridan)、コネティカット州) を使用して行った。重複した加水分解反応液から得られた値を、未知cDNA試料の各々について平均した。重複測定が互いに15%以内でなかった場合、第3の試料を実行し、前の2つの測定値と比較して、2つの最も近い値を平均した。

10

## 【0269】

## 定量ポリメラーゼ連鎖反応 (qPCR)

qPCRの実行はすべて、表1Aおよび1Bに示されているプライマー/プローブ・セットを使用して、Applied Biosystems社製7900HT迅速リアルタイムPCRシステム (Applied Biosystems社製、フォスターシティ、カリフォルニア州) またはMX3000P (登録商標) (Stratagene社製、ラホーヤ、カリフォルニア州) のいずれかで実施した。プローブはすべて、5'末端のFAM (商標) (Applied社製、ノーウォーク、コネティカット州) および3'末端のBHQ-1 (登録商標) 消光体で標識されており、Biosearch社 (ノバート、カリフォルニア州) が合成した。各プライマー/プローブ・セットを検証して、アッセイの発現範囲にわたって、PCR増幅の効率がおよそ100%であることを保証した。各ヒト供与体由来のcDNAを1ウエル当たり1ngまたは10ngのいずれかで含有する重複プレート (96ウエル形式) を構築した。プレートは、2つの陰性対照ウエル (「NTC」、水のみ) および10個体の血液に由来する貯溜された市販のcDNA (参照cDNA) の3つのウエルも含有する。各qPCR反応液は25  $\mu$ l (最終容積) であり、下記の成分を含有していた: 12.5  $\mu$ lのBrilliant QPCR Master Mix (登録商標) (Stratagene社製)、400nMの順方向プライマー、400nMの逆方向プライマー、50nMのプローブ、および60nM/300nMのROX (商標) (Applied社製) (MX3000P (登録商標) 7900HT装置)。サイクル条件は、95、10分間、その後95、15秒間; 60、1分間の40サイクルだった。重複qPCRの実行は、各遺伝子について実施した。例外的に、遺伝子の重複プレートが十分に一致しなかった場合、第3のqPCRプレートを実行した。取得されたCt値に応じて、3つすべてのプレートからの値を平均するか、または異常なプレートをさらなる解析から除外した。

20

30

## 【0270】

qPCRの実行に使用した装置では、予備データ分析ステップが求められた。しかしながら、各々の場合、目的は、増幅曲線の間接点付近の増幅閾値を、所与のプレートの試料すべてに使用された同じ閾値に設定することだった。同じ遺伝子の重複プレート実行の閾値は、必ずしも同一でなかったが類似していた。MX3000P (登録商標) の場合、以下の設定を使用して最初に閾値を決定した: 平滑パラメーター = 5、MX4000アルゴリズムを使用した基線計算、およびシグマ乗算器 (sigma multiplier) によるサイクル6からサイクル14を使用したバックグラウンドに基づく閾値は20。必要な場合は、閾値の細かい調整を手動で行って、閾値を増幅プロットのほぼ中央に設定した。7900HTでのプレート実行の場合、装置の初期設定を使用して、最初に閾値を設定した。必要な場合は、その後手動調整を行った。

40

## 【0271】

表 1A: 選択された遺伝子/バイオマーカーのプライマー/プローブ配列。

遺伝子名	略語	遺伝子受入番号 (配列番号)	代表的なプライマー/プローブ配列(5'から 3')†
アデノシン・デアミナーゼ	ADA	NM_000022 (配列番号 88)	F = GGTTGGTGGAGCTGTGTAAGAAGTAC (配列番号 1) R = CTTCTGGGATGGTCTCATCTC (配列番号 2) P = CAGCAGACCCGTGGTAGCCATTGACCT (配列番号 3)
ベータ-アレスチン 1	ARRB1	L04685 (配列番号 89)	F = AGACACGAACCTTGGCCTCTAGC (配列番号 4) R = TTGTAGGAAACAATGATCCCCAG (配列番号 5) P = TTGAGGGAAGGTGCCAACCCGTGAGAT (配列番号 6)
ベータ-アレスチン 2	ARRB2	BC007427 (配列番号 90)	F = TCTTCCATGCTCCGTCACAC (配列番号 7) R = CGAATCTCAAAGTCTACGCCG (配列番号 8) P = AGCCAGGCCCCAGAGGATACAGGAAA (配列番号 9)
CD8 アルファ	CD8a	M12824 (配列番号 91)	F = TTCCGCCGAGAGAACGAG (配列番号 10) R = AAGACCCGCCACGAAAGTGG (配列番号 11) P = TCGGCCCTGAGCAACTCCCATCATGTA (配列番号 12)
CD8 ベータ	CD8b	M37601 (配列番号 92)	F = TGACAGTCACCACGAGTTCCTG (配列番号 13) R = TCTCCTGTCCACCTTCCACC (配列番号 14) P = CTCTGGGATTCCGCCAAAAGGGACTAT (配列番号 15)

10

20

30

40

遺伝子名	略語	遺伝子受入番号 (配列番号)	代表的なプライマー/プローブ配列(5'から3')†
cAMP 応答配列結合タンパク質 1	CREB1	NM_134442 (配列番号 93)	F = CTGGCTAACAAATGGTACCGATG (配列番号 16) R = GTGGTCTGTGCATACTGTAGAATGG (配列番号 17) P = CATGACCAATGCAGCAGCCACTCA (配列番号 18)
cAMP 応答配列結合タンパク質 2	CREB2	M86842 (配列番号 94)	F = CACGTTGGATGACACTTGTGATC (配列番号 19) R = CTGGGAGATGGCCAAATTGG (配列番号 20) P = ACTAATAAGCAGCCCCCAGACGGT (配列番号 21)
ジブチル・ペプチダーゼ-IV	DPP4	M74777 (配列番号 95)	F = GTGTCAATTCAGTAAAGAGCGGAAG (配列番号 22) R = CTCAGCCCTTATCAITTCACGC (配列番号 23) P = TTCCGGTCCCTGGTCTGCCCTCTATA (配列番号 24)
細胞外シグナル制御キナーゼ 1	ERK1	M84490 (配列番号 96)	F = TGACGGAGTATGTGGTACGC (配列番号 25) R = CCACAGACCAGATGTCGATGG (配列番号 26) P = CTGGTACCGGGCCCCAGAGATCAT (配列番号 27)
細胞外シグナル制御キナーゼ 2	ERK2	M84489 (配列番号 97)	F = TAACGTTCTGCACCGTGACC (配列番号 28) R = CAGGCCAAAAGTACAGATCTTG (配列番号 29) P = ACCTGCTGCTCAACACCACCTGTGAT (配列番号 30)
グアニン・スクレオチド結合タンパク質アルファ i2	Gi2	X04828 (配列番号 98)	F = AGGCGTCTCCCTGATGAC (配列番号 31) R = GCTCCAGGTCGTTACAGGTAGTAG (配列番号 32) P = AGGCTGCTTTGGCCGCTCAA (配列番号 33)

10

20

30

40

遺伝子名	略語	遺伝子受入番号 (配列番号)	代表的なプライマープローブ配列(5'から 3') <sup>†</sup>
グアニン・ヌクレオチド結合タンパク質アルファ s(長鎖)	Gs	AF493897 (配列番号 99)	F = GACTATGTGCCGAGCGATCAG (配列番号 34) R = GTCCACCTGGAACCTTGGTCTCA (配列番号 35) P = CTGCTTCGCTGCCGTGTCCTGA (配列番号 36)
アルファ-グルコリコイド受容体	GR	X03225 (配列番号 100)	F = TCCCTGGTCGAACAGTTTTTTC (配列番号 37) R = TTTGGGAGGTGGTCCCTGTTG (配列番号 38) P = TGTAAAGCTCTCCTCCATCCAGCTCCTCAA (配列番号 39)
インターロイキン 1、ベータ	IL1b	NM_000576 (配列番号 101)	F = GATGGCCCTAAACAGATGAAAGTG (配列番号 40) R = CCTGAAAGCCCTTGGCTGTAGTG (配列番号 41) P = ATGGCGGCATCCAGCTACGGAATCTC (配列番号 42)
インターロイキン 6	IL6	M14584 (配列番号 102)	F = AGCCACTCACCTTTCAGAACG (配列番号 43) R = CATGTCTCCTTTCAGGGCTG (配列番号 44) P = CAAAATTCGGTACATCCTCGACGGCAT (配列番号 45)
インターロイキン 8	IL8	M28130 (配列番号 103)	F = CTGTAGCCAGGATCCACAAG (配列番号 46) R = CTGTGAGGTAAGATGGTGGCTAATAC (配列番号 47) P = CTTGTTCCACTGTGCCTTGGTTTCTCCTT (配列番号 48)
インドールアミン-ピロール 2,3 ジオキシゲナターゼ	INDO	NM_002164 (配列番号 104)	F = GCTTCGAGAAAAGAGTTGAGAAGTTAAAC (配列番号 49) R = GACCTTTGCCCCACACATATG (配列番号 50) P = CTCACAGACCACAAAGTCACAGCGCCTT (配列番号 51)

10

20

30

40

遺伝子名	略語	遺伝子受入番号 (配列番号)	代表的なプライマー/プローブ配列(5'から3') <sup>†</sup>
p38 ク質キナーゼ 14	MAPK14	L35253 (配列番号 105)	F = CGGCAGGAGCTGAACAAGAC (配列番号 52) R = AGCAGCACACACAGAGCCATAG (配列番号 53) P = CCGAGCGTTACCAGAAACCTGTCTCCA (配列番号 54)
マイトジェン活性化タンパク質 キナーゼ 8	MAPK8	AY893269 (配列番号 106)	F = CCAACACCCGTACATCAATGTC (配列番号 55) R = CACTCTTCTATTGTGTGTTCCCTTTC (配列番号 56) P = CACCACCAAAAGATCCCTGACAAAGCAGTT (配列番号 57)
map 1	MKP1	X68277 (配列番号 107)	F = GCCAGGCAGGCATTTC (配列番号 58) R = ATGCTTCGCCTCTGCTTCAC (配列番号 59) P = TCAGCCACCATCTGCCTTGCTTACCCTT (配列番号 60)
ミネラルコルチコイド受容体	MR	M16801 (配列番号 108)	F = AGCCCAGAGGAAGGGACAAC (配列番号 61) R = TGTGAGCGCTCGTGAGATTG (配列番号 62) P = CTCCTGCAAAAAGAAACCCCTCGGTCAACA (配列番号 63)
オルニチン・デカルボキシラーゼ 1	ODC1	NM_002539 (配列番号 109)	F = CCATGTAGGAAGCGGCTGTAC (配列番号 64) R = TCAGCCCCCATGTCAAAAAC (配列番号 65) P = ATCCTGAGACCTTCGTGCAGGCAATCT (配列番号 66)
プリン受容体 P2X7	P2X7	NM_002562 (配列番号 110)	F = GCTGTCGCTCCCATATTTATCC (配列番号 67) R = CACAATGGACTCGCACTTCTTC (配列番号 68) P = CTGTCAGCCCTGTGTGGTCAACGGAATAC (配列番号 69)

遺伝子名	略語	遺伝子受入番号 (配列番号)	代表的なプライマー/プロンプ配列(5'から3') <sup>†</sup>
ベンゾジアゼピン (benzodiazapine)受容体(末梢性タイプ)	PBR	BC001110 (配列番号 111)	F = CTGGTCTGGAAGAGCTGGG (配列番号 70) R = CAGCAGGAGATCCACCAAGG (配列番号 71) P = CCCCATCTTCTTTGGTGCCTGAC (配列番号 72)
プロリル・エンドブチダーゼ	PREP	D21102 (配列番号 112)	F = GGGAAATGACTACGTGACCAATG (配列番号 73) R = GGATCCCTGAAGTCAATGTTGATC (配列番号 74) P = CATTCAAGACGAAATCGCCAGTCTCCC (配列番号 75)
G タンパク質伝達調節因子 2	RGS2	NM_002923 (配列番号 113)	F = GATTGGAAGACCCGTTTGAGC (配列番号 76) R = CAGGAGAAAGGCTTGATGAAAGC (配列番号 77) P = CTGGGAAGCCCAAAACCCGGCAA (配列番号 78)
S100 カルシウム結合タンパク質 A10(p11)	S100A10	NM_002966 (配列番号 114)	F = AGGAGTTCCTGGATTTTGG (配列番号 79) R = GCCCACCTTGCCATCTACAC (配列番号 80) P = CAAAAAGACCCCTCGGCTGGACAAAA (配列番号 81)
セロトニン輸送体	SERT	NM_001045 (配列番号 115)	F = CATGGCTGAGATGAGGAATGAAG (配列番号 82) R = GCTGGCATGTTGGCTATCG (配列番号 83) P = ACGCAGGTCCCAGCCCTCCCTTCAT (配列番号 84)
小胞モノアミン輸送体 2	VMAT2	L23205 (配列番号 116)	F = TGGATTCGTCAATGATGCCTATC (配列番号 85) R = ATGCCACATCCGCAATGG (配列番号 86) P = AGACCTGGGCACGTGTCCTGCTA (配列番号 87)

<sup>†</sup>F=順方向プライマー配列;R=逆方向プライマー配列;P=プロンプ配列

10

20

30

40

50

【 0 2 7 2 】

遺伝子発現の標準化

異なる試料間の遺伝子発現プロファイルを有効に比較するために、あらゆる根本的な生物学的変化を遮蔽する可能性がある変動を制御することが望ましい。例えば、酵素反応の効率、装置の性能、およびピペティングの日差はすべて、所与の日に得られるシグナルに影響を及ぼすだろう。これらの変動の影響を最小限にするための好ましい方法は、多重正規化遺伝子の使用による (Andersen, C. L.ら、Cancer Res、2004年、64巻 : 5245 ~ 5250頁 (非特許文献 7 1) ; Jin, P.ら、BMC Genomics、2004年、5巻 : 55頁 (非特許文献 7 2) ; Huggett, J.ら、Genes and Immunity、2005年、6巻 : 279 ~ 284頁 (非特許文献 7 3) )。理想的な標準化遺伝子は、便利に測定されるレベルで発現され、実験計画の一部で

ある操作によっては変化しない。標準化遺伝子の使用は普通のことであるが、研究者らは、自身が使用する遺伝子が自身の実験系で安定的に発現されるかどうかを検証しないことが多い。この問題を回避するために、市販のソフトウェア・プログラムである Genorm (商標) (Primer Design Ltd社製、サウスハンプトン、英国)を使用した。その方法は、Vandesompele, J.ら、Genome Biol、2002年、3巻(7号): RESEARCH0034.1-0034.11 (電子出版 2002年6月18日) (非特許文献74) による公表研究に基づいており、標準化遺伝子候補が安定的に発現されるかどうかを決定することを可能にする。標準化遺伝子を選択するためには、まず、血液試料を用いて実施された実験に重点をおいて文献を走査し、ヒトでの遺伝子発現を標準化するために研究者により以前に使用された遺伝子を特定する (Vandesompele, J.ら、Genome Biol、電子出版 2002年6月18日、3巻(7号): RESEARCH0034.1-0034.11 (非特許文献74)、特に0034.5頁、表3; Applied Biosystems Application Note 2006年、publication 127AP08-01 (非特許文献75)、特に3頁、図1)。この探索から、表1Bに示されている遺伝子を特定した。これら遺伝子が、本実験における標準化に有効だったことを確認するために、7つの遺伝子の発現プロファイルを、正常対象、薬物治療を受けていないうつ病患者、および薬物治療を受けているうつ病患者を含む異なる実験セットに由来する血液試料を使用して、Genorm (商標) で分析した。すべてのセットにおいて、7つの遺伝子の組合せは、0.15以下の対比変動値(V)により決定したところ、良好な標準化を達成した (Vandesompele, J.ら、Genome Biol、電子出版 2002年6月18日、3巻(7号): RESEARCH0034.1-0034.11 (非特許文献74))。

10

20

#### 【0273】

Genorm (商標) によると、標準化のためには2つまたは3つの最良遺伝子を使用することが必要なだけであると記されているが、いくつかの理由で、3つを超える標準化遺伝子の組合せを考慮すべきである。第1に、より多くの標準化遺伝子を使用することにより、新しい薬物治療、遺伝的バックグラウンド、または疾患状態が、標準化遺伝子の発現に影響を及ぼす場合があることを考慮して、予測が支援されるだろう。3つを超える標準化遺伝子は、特定の試験で安定的に発現されないあらゆる遺伝子の影響を低下させることにより、プロセスを向上させることが期待される。また、発現データを標準化するために一貫して3つを超える遺伝子を使用することにより、長い期間をかけて実施されたすべての研究から発現結果を比較することができる。臨床試料は、適切な対照と必ずしも一致するとは限らないため、3つを超える標準化遺伝子の使用を考慮することは重要である。異なる実験にわたって遺伝子発現を比較する場合、3つを超える遺伝子を用いた標準化が好ましい方法であるが、比較する試料がすべて同じ様式で処理されていれば、任意の特定の試験において2つまたは3つの遺伝子を使用することは依然として有効である。

30

#### 【0274】

【表 8】

表1B：標準化遺伝子。

遺伝子名	略語	遺伝子受入番号 (配列番号)
ベータ-アクチン	ACTB	NM_001101 (配列番号117)
ベータ-2-ミクログロブリン	B2M	NM_004048 (配列番号118)
グリセルアルデヒド-3-リン酸デヒドロゲナーゼ	GAPD	NM_002046 (配列番号119)
グルクロニダーゼ、ベータ	GUSB	NM_000181 (配列番号120)
ヒドロキシメチル-ピラン・シンターゼ	HMBS	NM_000190 (配列番号121)
ヒポキサンチン・ホスホリボシル-トランスフェラーゼI	HPRT1	NM_000194 (配列番号122)
ホスホグリセリン酸キナーゼ	PGK1	NM_000291 (配列番号123)
ペプチジルプロピル・イソメラーゼ A(サイクロフィリンA)	PPIA	NM_021130 (配列番号124)
リボソームタンパク質、大型、P0	RPLP0	NM_001002 (配列番号125)
リボソームタンパク質、L13a	RPL13A	NM_012423 (配列番号126)
コハク酸デヒドロゲナーゼ複合体、サブユニットA	SDHA	NM_004168 (配列番号127)
TATAボックス結合タンパク質(転写因子IID)	TBP	NM_003194 (M34960) (配列番号128)
転移受容体(p90、CD71)	TFRC	NM_003234 (配列番号129)
ユビキチンC	UBC	NM_021009 (M26880) (配列番号130)
チロシン3-モノオキシゲナーゼ/トリプトファン5-モノオキシゲナーゼ活性化タンパク質、ゼータ・ポリペプチド	YWHAZ	NM_003406 (配列番号131)
真核生物18SリボソームRNA	18S	X03205 (配列番号132)

10

20

30

## 【0275】

本明細書に記述されている遺伝子のいずれのプライマーも、上記の5.4.1.2節に記述されているように設計することができる。表1Aおよび表1Bで特定されている遺伝子の公的に利用可能な配列は、遺伝子受入番号(GenBankデータベース)により示されており、参照によりそれらの全体が本明細書に組み込まれる。表1Aおよび表1Bで特定されている遺伝子の配列は、表に示されている適切な配列番号により列挙されているように添付の配列表に開示されている。

## 【0276】

## 転写データ分析

重複PCRプレートに由来する各未知試料の平均Ct(サイクル閾値)値を、各遺伝子ごとに決定した。リアルタイムPCRアッセイでは、陽性反応を、蛍光シグナルの蓄積により検出する。Ctは、蛍光シグナルが閾値を超える(つまり、バックグラウンド・レベルを超過する)のに必要なサイクル数として定義される。Ctレベルは、試料中の標的核酸の量に反比例する(つまり、Ctレベルが低ければ低いほど、試料中の標的核酸の量は大きくなる)。

40

## 【0277】

各未知cDNA試料ならびに参照cDNAの相対的発現レベルを、7つの標準化遺伝子からの平均Ct値を使用して、 $2^{-\Delta Ct}$ 法(Livak, K.およびSchmittgen, T., Methods, 2001年、25巻:402~408頁(非特許文献76))で計算した。次に、参照cDNA

50

の相対的発現レベルを100%に設定し、その後他のすべての試料を、参照の割合として表した。最後に、これら割合を、参照cDNAに含有されていた各遺伝子のコピー数と乗算することにより、これら割合をcDNAの1ng当たりのコピー数に変換した。

#### 【0278】

##### 単変量統計分析およびグラフ化

遺伝子発現値と、患者/対象の質問票から得た臨床パラメーターとの間の相関性を、R統計パッケージを使用して調査した。質問票データを必要に応じてコード化して、比較を容易にした。遺伝子発現データを解析前に対数変換し、パラメトリック分析およびノンパラメトリック分析を両方とも実施した。有意性の閾値を、 $p < 0.05$ に設定した。例えば、表3を参照されたい。単変量検定を使用して、特定の遺伝子が一貫して上方制御または下方制御されているかどうかを、所与の対象集団ごとに決定した。

10

#### 【0279】

対照対象とうつ病患者との間で発現レベルを比較する散布図およびそれに伴う単変量統計分析を、GraphPad Prism4 (登録商標) (GraphPad Software, Inc社、サンディエゴ、カリフォルニア州)を使用して、各遺伝子ごとに生成した。遺伝子発現値は、必ずしも定期的に分布しないため、ノンパラメトリックなマン・ホイットニー検定を使用してグループを比較した。有意性の閾値を、 $p < 0.05$ に設定した。若干の遺伝子およびそれらの血中の相対的発現レベルは、図2~7に例示されている。

20

#### 【0280】

##### 多変量解析

罹患患者と健常対照対象を区別するために、分類アルゴリズムを使用した。分類アルゴリズム、典型的には機械学習アルゴリズムは、以下の2つのステップに従って実行される：(1)mRNA転写データセットから、その遺伝子発現レベルが集団的に最も情報価値の高いことが見出される遺伝子のサブセットを選択する；(2)ステップ(1)で識別されるような遺伝子のサブセットに対してトレーニングされた事前選択タイプの分類アルゴリズムをトレーニングし、返す。

#### 【0281】

##### (1) 遺伝子の選択

第1のステップでは、健常対照対象およびうつ病対象または他の罹患対象からのmRNA転写データセットを、集団的にランダム・フォレスト・アルゴリズム(Breiman, L., 2001年、Machine Learning 45巻(1号):5~32頁(非特許文献60))の入力として使用した。各データセットは、表1Aに列挙されている遺伝子と本明細書に記述されている方法とに基づいて、各対象の血液試料に由来するmRNA転写データを表している。最も重要でない遺伝子を連続的に除去することにより、ランダム・フォレスト・アルゴリズムは、アウト・オブ・バッグ(OOB、out-of-bag)誤差最小化判断基準を使用して、最も重要な遺伝子を含むリストを返す(Liaw, A、およびWiener, M. 2002年、12月、Classification and regression by randomForest. R News Vol. 2/3:18~22頁(非特許文献77))。

30

#### 【0282】

##### (2) トレーニングおよび分類化

第2のステップでは、サポート・ベクター・マシン分類アルゴリズム(Cortes, C.およびVapnik, V. 1995年、Machine Learning、20巻(3号):273~97頁(非特許文献61))などを、ステップ(1)のようにして特定された最も重要な遺伝子に関連する転写プロファイルを使用して調整し、相互検証法に基づいてトレーニングした。

40

#### 【0283】

別の方法では、段階的ロジスティック回帰法を、最も重要であるかまたは最も説明的な遺伝子を選択するステップ(1)、および分類用アルゴリズムを相互検証法によりトレーニングするステップ(2)の両方のために使用する。

#### 【0284】

50

他の解析では、RVM分類器を遺伝的アルゴリズムと共に使用した。データセットを、RVMアルゴリズムを用いてトレーニングし、遺伝的アルゴリズムは、異なるサブセットの候補変数に対してトレーニングおよび試験された多数のRVMを評価して、考え得る遺伝子相互作用を識別した。各変数サブセットの性能は、相互検証により評価した。

【0285】

トレーニング・ステップ中に、リーブ・ワン・アウト相互検証法(LOOCV)または10分割相互検証法などの相互検証法が、アルゴリズムにより実施された。相互検証法は、解析が単一サブセットについて初期に実施されるように、データの試料を別個のサブセットに分別し、他のサブセット(複数可)は、初期解析の確認および検証にその後使用するために保持されるという統計的行為である。データの初期サブセットはトレーニング・

10

【0286】

例えば、全試料に由来するデータ(N)を、2つの別個のサブセットに分割し、データの一方のサブセット(m)を試料の検証に使用する、つまりサブセットmを未知のセットとして使用する。残りのサブセット(N-m)により、分類アルゴリズムがトレーニングされる。そのような相互検証(CV)法を、データセットがすべて未知として処理されるまで繰り返す。正確度の値および予測値は、未知として処理された試料の各々が正確に分類したかどうかに基づいて計算することができる。

【0287】

1つのそのような相互検証法では、分類アルゴリズムは、試料データセットの90%を用いてトレーニングされ、試料データの残りの10%の分類は、トレーニングされたアルゴリズムにより予測される。そのような10分割CVを、10回繰り返す。相互検証法は「動作曲線」を示すことができる。つまり、トレーニングされた分類アルゴリズムは、何らかの無作為選択プロセスよりも良好に、例えば偶然よりも良好に機能する。上記の(1)および(2)で示されている指示に従って構築された分類アルゴリズムの分類誤差を推定するために、正確度、陽性的中率(PPV)、および陰性的中率(NPV)の計算を行って、トレーニングされた分類アルゴリズムが良好に機能した程度を決定した。

20

【0288】

トレーニングされた分類アルゴリズムの正確度は、試料の総数中の正確分類の総数である。

30

【0289】

上記の方法により「罹患」クラスに正確にスコアされたデータセット(つまり、対象)の数は、陽性的中率(PPV)の尺度をもたらす。正確度率とも呼ばれるPPVまたは疾患の事後確率は、正確に診断された陽性検査結果を有する患者の割合である。

【0290】

また、上記の方法により「健常」または「対照」群に正確にスコアされたデータセット(つまり、対象)の数は、陰性的中率(NPV)の尺度をもたらす。陰性的中率は、正確に診断された陰性検査結果を有する患者の割合である。

【0291】

無作為化(並べ替え)データセットの解析

SLRまたはSVMを使用して取得された分類の正確度が有意であったかどうか、つまり偶然よりも良好であったかどうかを決定するために、各データセットを以下のようにさらに解析した:

40

a) 元のデータセットの正確度を、上記に説明されている方法により取得した。

b) 3つの新しい並べ替えデータセットを生成し、元のデータセットと同じような患者の割合を依然として維持しつつ、個々の試料の各々の帰属を無作為に割当てた。

c) その後、各無作為化されたデータセットの正確度を計算した。

d) 10個の正確度(元のデータセットの10分割CVから)を、マン・ホイットニー検定を使用して30個の並べ替え正確度(10分割CVに基づいて行った3つのランダム

50

・セット)と比較した。

e) 0.01未満のp値をもたらす比較は、元のデータセットからの正確度が無作為の偶然によらないことを意味し、つまり対照群および患者群を分離することができることを意味すると解釈された。0.01を超えるp値をもたらす比較は無作為であるとみなされ、患者群および対照群を確信を持って分離されないことを意味する。

【0292】

転写プロファイルの特定に使用された患者/対象

これらの研究の1つの目的は、情動障害などの神経精神障害のリスクにある表現型を特定するのを支援することができる、正常供与体およびサブグループの血液で特定された転写プロファイルを、定義、相関、および関連させることであった。正常供与体の基線転写プロファイルを確立した後、正常集団と、臨床的に診断されたうつ病、重度うつ病、双極性障害、BPD、またはPTSDを罹患している患者との間で比較を行った。これらの研究の別の目的は、対象を、正常対照またはうつ病、重度うつ病、双極性障害、BPD、もしくはPTSDなどの情動障害を罹患している患者のいずれかに分類することができるプロファイルを特定することであった。

10

【0293】

正常集団内のサブグループの存在、例えばリスク・プロファイルを有する対象を決定し、サブグループを全血の転写プロファイルと相関させることができるように、正常志願者の基線データベースを確立した。

【0294】

対照患者/対象(米国)

500個の血液試料を、ペンシルベニア州南東部およびデラウェア州区域に対応している血液銀行で献血する正常志願者から収集した。全供与者からインフォームドコンセントを得た。個人情報不可逆的に匿名化した。

20

【0295】

集団内の分散を最小限にするために、供与者を白人に限定した。集団内では、供与者は性別が均等に分かれていた。供与者について血液銀行で使用された上記のもの以外の追加的な排除要因はなかった。供与者は皆、全身健康状態、医学的問題、薬物使用および乱用、家族歴、ならびに精神医学的問題の特徴付けを支援するための質問票に記入することが求められた。質問票の項目は、著作権の制約無しで使用可能な標準的精神医学尺度に基づいていた。質問票の答えは自己申告であり、供与者は医学的評価または精神医学的評価を受けなかった。質問票は、表2に分類された要因を含む複数の要因を包含していた。

30

【0296】

【表 9】

表2

デモグラフィック	医学一般	家族歴	精神医学的病歴/生活経験	うつ症状
人種	身長/体重	近親者の自殺	過去/最近のストレスの多い生活上の出来事の存在/重大さ	自律神経機能の変化
性別	現在/過去の薬物治療	家族の精神医学的病歴	精神病の既往診断	認知機能の変化
婚姻歴	現在/過去の医学的問題			不安/パニック発作
雇用状態	外科手術			
職業	喫煙			
食事頻度	アルコール摂取			
	薬物乱用			

10

## 【0297】

詳細な質問票を使用して、供与者の精神医学的障害の将来的リスクを増加させる場合がある供与者の病歴または現在の医学的状态における複数の要因に関するデータを取得し、固有の転写プロファイルを、質問票を使用して特定された特定の表現型に関連させた。このデータを使用して、正常集団を区分し、現在使用可能な方法を用いるよりも信頼性が高く一貫して、うつ病患者内の区画を特定した。評価された要因には、（これらに限定されないが）以下が含まれる：最近のストレスの多い生活上の出来事の重大さ、若年期ストレスの存在および重大さ、精神医学的障害の家族歴、ならびに食欲および睡眠パターンの変化を含む前うつ自律神経症状群。必要な場合、複数の質問群のスコアを統合して、複数の否定的要因、つまり症状スコアの影響を評価した。

20

## 【0298】

潜在的に血液転写プロファイルに影響を及ぼし得る喫煙または肥満度指数（BMI）などの共通因数は、極端と考えられる場合があるが、それらの交絡を回避するために、質問票データを使用して、デモグラフィック的、個人的、または医学的属性の特定可能なパターンにより、供与者をグループ化した。これら要因を独立して評価し、転写プロファイルに対するそれらの影響を評価した。これらが前うつ表現型の特定を交絡させる場合があり、そのような要因には、BMI、喫煙、アルコール乱用、薬物使用（および乱用）が含まれるため、供与者の特定および区分は、転写プロファイルに対するそれらの影響を評価するための非精神医学的な要因によるものであった。他の要因の影響も評価した。

30

## 【0299】

対照患者 / 対象（デンマーク）

200人の対象を、デンマーク民族出身（2世代までさかのぼる）に基づき、デンマークを地理的に包含するおよそ1000人の健常志願者（対照対象）からの初期採血から選択した。したがって、出生地（および両親および祖父母の出生地）に関するデータを取得した。全身健康状態および精神医学的病歴を最初に取得した。精神医学的病歴情報を、うつ病の以前のエピソードに簡単なスクリーニングで補完した。200人の対照対象のコホートでは、およそ40歳の平均年齢（18～65歳の範囲）の男性および女性の同様な分布がもたらされた。各対象は、身長、体重の評価、腹部周りおよび臀部周りの測定、ならびにEKGを含むちょっとした健康診断を受けた。各対象は、人格のある種の特性ならびに医学的疾患および精神的疾患のより完全な家族歴に関して対象が特徴付けられた詳細な質問票に記入した。（表2を参照。）

40

## 【0300】

50

上記で言及したような対照対象により提供されるデータを使用して、正常集団を区分し、特定の表現型を、末梢血で特定された転写プロファイルの変化と関連させた。表3 Aおよび3 Bを参照されたい。

#### 【0301】

対照患者 / 対象 (英国)

血液試料を、英国での管理された臨床研究に参加した健常志願者から収集した。全供与者からインフォームドコンセントを得た。この研究には男性および女性が含まれていた。女性は、容認された避妊法(二重障壁避妊)を使用する場合、外科的に避妊されていた場合、または閉経後(2年間月経がないと定義された)であった場合には参加したが、経口避妊薬は許可されなかった。参加した対象は、年齢が18歳および45歳であるが、65歳未満である。この研究に参加した各対象は、研究者の見解では、研究前の身体検査、病歴、生命徴候、ECG、ならびに血液生化学的、血液学的、および血清学的検査の結果、ならびに尿検査に基づき健康である。

10

#### 【0302】

うつ病患者の転写プロファイルの特定

うつ病患者の転写プロファイルの変化を評価するために、うつ病患者、つまり大うつ病性障害(MDD)を罹患している患者に由来する血液を、管理された臨床研究で取得した。全供与者からインフォームドコンセントを得た。

#### 【0303】

患者選択基準:

研究に適格な患者 / 対象は、基線来院でMADRS合計スコア26およびCGI-Sスコア4を有する中程度のMDDを罹患している男性または女性の外来患者だった。MDDの初期診断は、DSM-IV-TR(登録商標)基準によるものでなければならない。患者は、年齢が18~65歳であり(両端の年齢を含む)、精神科外来診療所および一般開業医で募集した。強迫性障害( OCD)、心的外傷後ストレス障害(PTSD)、またはパニック障害(PD)以外に、二次的共存性不安障害を罹患している患者(DSM-IV-TR(登録商標)基準)は、この研究に参加することができた。さらに、それ以外では、患者は、研究者の見解では、身体検査、病歴、および生命徴候に基づいて健康だった。研究者の見解で、臨床研究プロトコルを遵守しない可能性の高い患者、または任意の理由で不適格であった患者は、研究から除外してもよい。

20

30

#### 【0304】

うつ病患者の転写プロファイルの特定

重度大うつ病性障害(SMDD)を罹患している患者の転写プロファイルの変化を評価するために、これら患者に由来する血液を、管理された臨床研究で取得した。全供与者からインフォームドコンセントを得た。

#### 【0305】

患者選択基準:

この研究に適格な患者 / 対象は、精神科外来診療所で募集されたSMDDを罹患している外来患者であり、年齢が18~65歳(両端の年齢を含む)の男性または女性だった。この研究に参加した患者は皆、30以上のMADRS合計スコアを有するべきだった(つまり、より重度のうつ病患者)。選択された患者は、DSM-IV-TR(登録商標)基準による初期診断として大うつ病エピソード(MDE)を罹患している(精神疾患簡易構造化面接法(MINI、Mini International Neuropsychiatric Interview)で評価された現在のエピソード)。現在のMDEが報告された期間は、少なくとも3カ月であり、基線で12カ月未満である。患者は、中程度のうつ病患者に関して上記に説明されているような基準に基づいて、研究に参加する / 除外される。研究者の見解で、臨床研究プロトコルを遵守しない可能性の高い患者、または任意の理由で不適格であった患者は、研究から除外することができた。

40

#### 【0306】

双極性患者の転写プロファイルの特定

50

双極性患者の転写プロファイルの変化の評価するため、双極性患者から血液を取得した。これら患者は、精神科医による詳細な評価を受けており、医療を受けていた。全供与者からインフォームドコンセントを得た。

【0307】

患者選択基準：

患者/対象が、このプロトコルに基づいて血液を供与することができる前に、以下の基準が満たされていなければならない：

a) 患者は、DSM-IV-TR (登録商標) により、中程度または重度大うつ病または双極性Iと診断されている。患者の87パーセントは、双極性I型障害のDSM-IV-TR (登録商標) 基準を満たした。

b) 血液収集時に、患者は、いかなる精神薬理学的薬物も服用しておらず、いかなる精神薬理学的薬物も少なくとも2週間服用していない。加えて、患者は誰も、少なくとも2カ月の間、フルオキセチン、不可逆的MAOI、または蓄積性神経遮断薬で治療されていない。

c) 患者は、他の急性精神的症状、例えば物質乱用を罹患していない。

d) 可能な場合は常に、女性患者からの血液試料は、月経開始の2週間以内に収集されるべきである。いかなる場合でも、最後の月経の1日目の日付は記録されることになる。

e) 患者は、過去6カ月間、いかなる非合法薬物/乱用薬物を摂取していない。

f) 患者は、過去6カ月間、アルコールを乱用していない。

g) 女性患者は妊娠しておらず、授乳していない。

h) 患者は、現在のところ(前週を含む)、他のいかなる一般的な急性医学的状態(軽度の状態、例えば感冒を含む)も罹患していない。

i) 患者は、現在のところ(前週を含む)、いかなる定期的な薬物治療(経口避妊薬、薬草療法、栄養補助剤、ビタミンを含む)も受けていない。

j) 患者は、血液試料収集前の週内に、いかなる薬物治療(経口避妊薬、薬草療法、栄養補助剤、ビタミンを含む)も受けるべきではない。薬物を摂取した場合、例えば急性の頭痛の場合、血液試料収集を1週間遅延させるべきである。

k) 患者が喫煙を示す場合、1日当たりの平均量に関する情報を提供する必要がある。

l) 患者が、乱用ではないアルコール消費を示す場合、1週当たりの平均量に関する情報を提供する必要がある。

m) 患者は、血液試料収集に伴う質問票を提出した。

n) 患者は、患者情報を読み理解した。

o) 患者は、インフォームドコンセントに署名した。

【0308】

このプロトコルに基づき血液を供与するすべての患者から、以下の情報を取得しなければならない：詳細な精神医学的病歴および一般的病歴、精神医学的家族歴、現在症状の詳細な臨床的記述、少なくとも過去3カ月間の薬物治療歴、ならびに少なくとも過去6カ月間の非合法および合法薬物の乱用に関する情報。

【0309】

境界性人格障害患者の転写プロファイルの特定

境界性人格障害(BPD)の患者の転写プロファイルの変化を評価するために、境界性人格障害患者から血液を取得した。これら患者は、精神科医による詳細な評価を受けており、医療を受けていた。全供与者からインフォームドコンセントを得た。

【0310】

BPD研究用の患者/対象選択基準：

患者が、このプロトコルに基づいて血液を供与することができる前に、以下の基準が満たされていなければならない：

a) 患者は、DSM-IV (登録商標) により、境界性人格障害と診察されている。

b) 未治療患者群の場合、血液収集時に、患者は、いかなる精神薬理学的薬物も服用しておらず、いかなる精神薬理学的薬物も少なくとも2週間服用していない。フルオキセチ

10

20

30

40

50

ン、不可逆的MAOI、または蓄積性神経遮断薬で過去に治療された患者は、血液収集前の少なくとも4週間は、これら薬物治療のいずれも受けていない。

c) 患者の小さなコホート(およそ25人の患者)からの血液試料は、原発性精神障害(境界性人格障害)の急性精神医学的増悪期に収集されるだろう。他のすべての患者は、血液収集時に急性精神医学的増悪を罹患していないだろう。血液が急性増悪期に試料採取される患者でのみ、第2の試料が寛解中に収集されるだろう。医学的に可能な場合は常に、2時点での取扱いは同じだろう。

d) 患者は、他の急性精神科的症状、例えば物質乱用を罹患していない。

e) 可能な場合は常に、女性患者からの血液試料は、月経開始の2週間以内に収集されるべきである。いかなる場合でも、最後の月経の1日目の日付は記録されることになる。

f) 患者は、過去6カ月間、いかなる非合法薬物/乱用薬物を摂取していない。

g) 患者は、過去6カ月間、アルコールを乱用していない。

h) 女性患者は妊娠しておらず、授乳していない。

i) 患者は、現在のところ(前週を含む)、他のいかなる一般的な急性医学的状态(軽度の状態、例えば感冒を含む)も罹患していない。

j) 患者は、現在のところ(前週を含む)、処方されたベンラファクシンまたはデュロキセチン以外のいかなる定期的な薬物治療(経口避妊薬、薬草療法、栄養補助剤、ビタミンを含む)も受けていない。

k) 患者が、ベンラファクシンまたはデュロキセチンで治療されている場合、治療は、少なくとも3カ月間は現在の用量で与えられていなければならない。

l) 患者は、血液試料収集前の週内に、いかなる薬物治療(経口避妊薬、薬草療法、栄養補助剤、ビタミンを含む)も受けるべきではない。薬物を摂取した場合、例えば急性の頭痛の場合、血液試料収集を1週間遅延させるべきである。

m) 患者が喫煙を示す場合、1日当たりの平均量に関する情報を提供する必要がある。

n) 患者が、乱用ではないアルコール消費を示す場合、1週当たりの平均量に関する情報を提供する必要がある。

o) 患者は、血液試料収集に伴う質問票を提出した。

p) 患者は、患者情報を読み理解した。

q) 患者は、インフォームドコンセントに署名した。

#### 【0311】

このプロトコールに基づいて血液を供与するすべての患者から、家族歴、臨床的記述、ならびに薬物治療および薬物記録を含む詳細な精神医学的病歴を取得した。

#### 【0312】

患者は、転写プロファイルを交絡させる場合がある要因、例えば薬物使用、一般的医学的状态に特に取り組むために開発された質問票に記入した。患者は研究者に質問票を提出した。質問票は、転写解析サイトの人員に患者の身元が開示されないことを保証するために、血液試料および他の臨床データと同じコードでコード化された。質問票は、血液試料と一緒に転写解析サイトに移動させた。

#### 【0313】

心的外傷後ストレス障害(PTSD)患者の転写プロファイル

PTSDの患者の転写プロファイルの変化の評価するため、PTSD患者から血液を取得した。これら患者は、精神科医による詳細な評価を受けており、医療を受けていた。全供与者からインフォームドコンセントを得た。

#### 【0314】

PTSD研究用の患者/対象選択基準:

この研究の対象は、以下の基準を満たした男性だった:

a) 対象は、急性PTSDまたは軽減されたPTSDと診断されているか(DSM-IV(登録商標)による)、または心的外傷に曝されておりPTSDを発症していないか、または対照と分類される。心的外傷に曝されておらず、同じ地理的區域の出身だった対照を、この研究に選択した。

10

20

30

40

50

b) 血液収集時に、患者は、いかなる精神薬理学的薬物も服用しておらず、いかなる精神薬理学的薬物も少なくとも2週間服用していない。フルオキセチン、不可逆的MAOI、または蓄積性神経遮断薬で過去に治療された患者は、血液収集前の少なくとも4週間は、これら薬物治療のいずれも受けていない。

c) 患者は、他の急性精神科的症状、例えば物質乱用を罹患していない。

d) 患者は、過去6カ月間、いかなる非合法薬物/乱用薬物を摂取していない。

e) 患者は、過去6カ月間、アルコールを乱用していない。

f) 患者は、現在のところ(前週を含む)、他のいかなる一般的な急性医学的状态(軽度の状態、例えば感冒を含む)も罹患していない。

g) 患者は、血液試料収集前の週内に、いかなる薬物治療(薬草療法、栄養補助剤、ビタミンを含む)も受けるべきではない。薬物を摂取した場合、例えば急性の頭痛の場合、血液試料収集を1週間遅延させるべきである。

h) 患者が喫煙を示す場合、1日当たりの平均量に関する情報を提供する必要がある。

i) 患者が、乱用ではないアルコール消費を示す場合、1週当たりの平均量に関する情報を提供する必要がある。

j) 患者は、現在のところ(前週を含む)、いかなる定期的な薬物治療(薬草療法、栄養補助剤、ビタミンを含む)も受けていない。

#### 【0315】

上記に記述されているような臨床データおよびデモグラフィック・データはすべて、血液収集サイトで収集し、その後転写解析サイト(Lundbeck Research USA社、パラマス、ニュージャージー州)に情報を移動させた。臨床的特徴と転写プロファイルとの任意の関係性の探検解析は、Lundbeck Research USA社において、患者の身元を知ることなく実施された。

#### 【0316】

結果および考察

対照対象の転写プロファイルの同定。

表1Aに列挙されている29個の遺伝子の遺伝子発現レベルを、2つの対照群(米国およびデンマーク)からの対象を含む対照対象からの血液試料で測定した。

#### 【0317】

これら個体は皆、健康であるが、質問票の項目に対する特定の回答と相関する遺伝子発現の傾向が特定された。そのような傾向は、特定された場合、うつ病患者の集団で誇張される可能性がある。

#### 【0318】

質問票回答を統計分析のためにコード化された値に変換する。

米国およびデンマークの対照対象により記入された自己評価質問票には、類似しているが同一でない項目が含まれている。質問票からの情報を使用して、回答と遺伝子発現データとの間で考え得る関連性を探索するためには、統計分析の前に情報をコード化することが必要だった。

#### 【0319】

コード化戦略の例は以下の通りである：

a) 年齢およびBMIなどの連続変数は、対象により報告されたように使用された。あるいは、生スコアを、解析の前に2つまたは3つの区分(高値、中値、低値)に統合した。

b) 性別は、二値応答(0、1)に変換した。

c) 睡眠困難、活力の欠如、または気分の落ち込みなどのうつ病に関連する症状の頻度に関する質問は、言葉による回答(決してない、時々、ほとんどの日、毎日)から数値(0、1、2、3)に変換した。

d) 合併症状スコアは、症状の特定の組合せ値を積算して複合スコアを生成することにより生成した。その後、複合スコアに区分けした。

e) 対象のうつ病/不安症の家族歴に関する質問は、言葉による回答(無、二次親族の

10

20

30

40

50

み、一次親族)から数値(0、1、2)に変換した。

f)対象個人のうつ病/不安症の履歴、またはうつ病/不安症のための薬理的治療の履歴に関する質問は、言葉による回答(無、1つまたは複数)から二値応答(0、1)に変換した。

【0320】

コード化の後、スピアマン相関分析、t検定、およびANOVAを含む種々の統計的検定を使用して、遺伝子発現レベルと特定の臨床変数との間の関連性を探索した。

【0321】

必要に応じて、統計的検定を使用して、対象が自己評価質問票で提供したコード化回答と、各遺伝子の発現を比較して関連性を特定した。合計377個の比較がなされたため(29個の遺伝子×13項目の質問票回答)、有意性の閾値を $p < 0.01$ に設定して、タイプ1の誤差の可能性を最小限しつつ、それでも多数の統計的に有意な結果を保持した。

10

【0322】

表3Aおよび3Bは、分析された質問票回答に基づいて、対照集団内で有意差を示す29個の遺伝子(表1Aから)のうちの15個のみに関する相関性データを示す。残りの遺伝子については、有意差は検出されなかった。表3Aおよび3Bは、13項目の質問票回答のうちの11項目に関するデータを示すが、BMIおよび年齢に関する相関性データは、有意差がなかったため示されていない。著しい遺伝子発現プロファイルと相関する臨床パラメーターのいくつかは、生涯経験、生涯治療、および症状スコアである。

【0323】

20

表3Aおよび3B。2つの対照群における臨床変数と遺伝子発現との間の相関性。

(\* \* =  $p < 0.01$  基準 ; \* \* \* =  $p < 0.001$  基準)

【0324】

## 【表 1 0】

表3A 1)米国対象 2)デンマーク対象

	CREB2	DPP4	ERK1	ERK2	GR	Gs	MAPK8	MAPK14
1)家族歴(D/A/S)		増加**					増加**	
2)家族歴(D/A/S)								
1)喫煙								
2)喫煙								
1)生涯経験(D/A)	増加***	増加***		増加**	増加***		増加***	
2)生涯経験(D/A)		増加***			増加傾向			
1)生涯治療(D/A)	増加**	増加***			増加**		増加***	
2)生涯治療(D/A)					増加傾向	増加**		
1)食欲の変化		増加**						
2)食欲の変化								
1)睡眠障害		増加**						
2)睡眠障害		増加**						
1)10症状スコア(*)		増加***					増加***	
2)10症状スコア(*)			増加**	増加***				
1)自律神経症状		増加**						
2)自律神経症状								
1)最近のストレス		増加**						
2)最近のストレス								
1)若年期ストレス								
2)若年期ストレス								
1)性への関心					増加**			
2)性への関心								増加**

(D/A/S=うつ病/不安症/自殺;D/A=うつ病/不安症)

10

20

## 【 0 3 2 5】

30

## 【表 1 1】

表3B 1)米国対象 2)デンマーク対象

	MKP1	MR	PBR	RGS2	S100A10	SERT	VMAT2
1)家族歴(D/A/S)							
2)家族歴(D/A/S)							
1)喫煙						減少***	
2)喫煙						減少傾向	
1)生涯経験(D/A)				増加**			増加**
2)生涯経験(D/A)				増加傾向			
1)生涯治療(D/A)	増加**						
2)生涯治療(D/A)	増加傾向						
1)食欲の変化						減少**	
2)食欲の変化						減少傾向	
1)睡眠障害							
2)睡眠障害							
1)10症状スコア(*)	増加傾向		減少**				
2)10症状スコア(*)	増加**		増加**				
1)自律神経症状							
2)自律神経症状							
1)最近のストレス							
2)最近のストレス							
1)若年期ストレス							
2)若年期ストレス					増加***		
1)性への関心			減少傾向				
2)性への関心		減少**	増加**				

(D/A/S=うつ病/不安症/自殺;D/A=うつ病/不安症)

## 【0326】

分析された合計377個の組合せのうち、23個の組合せ(6%)が、分析された2つの対照群間の有意差を示す。しかしながら、これら組合せのうち345個(94%)は、同じプロファイルを示す。これら組合せのうち9個は、表3Aおよび3Bの網掛け部分により示されているように、研究された両対照群で、同じ方向への遺伝子発現の変化(つまり遺伝子の上方制御または下方制御)を示した。全体として、この分析は、分析に使用された2つの対照群が、非常に類似した遺伝子発現の傾向またはプロファイルを示していることを示す。

## 【0327】

臨床パラメーターと関連する遺伝子発現プロファイルは、本明細書に記述されている多変量アルゴリズムにより分析することもできる。したがって、転写データと混合された臨床変数を、段階的ロジスティック回帰法またはPELORAなどの、当業者に公知である任意の好適なアルゴリズムにかけることができる。

## 【0328】

うつ病患者の転写プロファイルの特定

抗うつ治療を受けていない174人の中程度のうつ病患者/対象から得られた血液試料を、最初に単変量法で分析した。表1Aから選択される遺伝子の転写レベルを測定し、196人の健常対照対象のそのような遺伝子の発現レベルと比較した。対照と比較した、うつ病患者の代表的な遺伝子の発現プロファイルが、図2A~2Bおよび3A~3Bに示されている。

## 【0329】

RF(選択)およびSVM(トレーニング)を使用した、中程度のうつ病患者と対照との分類は、図8Aに示されているように、結果として88%の高正確度でもたらされた(

10

20

30

40

50

PPV = 89% ; NPV = 88% )。遺伝子選択およびトレーニングを両方とも実施するSLRアルゴリズムを使用した、中程度のうつ病患者と対照との分類は、図8Aに示されているように、結果として93%の高正確度でもたらされた ( PPV = 93% ; NPV = 94% )。

【0330】

両アルゴリズムは、図8Bに示されているように、データセット全体に基づいて選択された遺伝子の良好な一致を示した。各方法の統計パラメーターに基づく分類に最も重要な遺伝子として、ランダム・フォレスト法は14個の遺伝子を選択し、SLRは17個の遺伝子を選択した。ARRB1、ARRB2、CD8a、CREB1、CREB2、ERK2、Gi2、MAPK14、ODC1、P2X7、およびPBRを含む11個の遺伝子が、両方の方法により選択された。

10

【0331】

データセットを無作為化、つまり患者または対照としての試料の帰属を無作為化し、上記のような同じ多変量解析にかけた。無作為化した後、両分類アルゴリズム ( RF / SVM および SLR ) は、実際のデータで取得されたものとは統計的に異なる正確度の値をもたらし、上記 ( 図8A ) に列挙されている値が偶然よりも良好であり、群は統計的に分離可能であることを示す。

【0332】

対象をプロファイルすることができ、表1Aの遺伝子に基づく対象の転写データを、上記に記述されているようなパラメーターを用いてトレーニングされた分類アルゴリズムにかけて、中程度のうつ病の診断を取得した。

20

【0333】

表1Aから選択された遺伝子に関する、うつ病対象の転写プロファイルは、各バイオマーカー ( 例えば、遺伝子転写物 ) の含有量に基づいて表4に示されている。対照対象の転写物値は、比較のために示されている。

【0334】

## 【表 1 2】

表4

バイオマーカー (遺伝子の略語)	うつ病対象群の特徴: 含有量=バイオマーカー の平均転写物値(± SD)	対照対象群の特徴:含 有量=バイオマーカー の平均転写物値(± SD)
ADA	4691±2453	4511±1710
ARRB1	189062±62727	297143±91094
ARRB2	84195±31728	114780±39962
CD8a	8304±5825	14693±8416
CD8b	8145±4394	8687±3880
CREB1	71743±20237	63725±16022
CREB2	63732±14463	77059±15755
DPP4	6649±2331	7169±2890
ERK1	25326±10178	39016±12900
ERK2	58338±18813	54137±18660
Gi2	115117±53383	226358±87609
Gs	262885±112989	303930±139837
GR	73224±23517	80610±26544
IL1b	29631±13692	21006±9313
IL6	348±523	182±221
IL8	45487±106224	28024±19993
INDO	6031±10133	5596±4418
MAPK14	73156±33915	51632±20341
MAPK8	12906±3836	12162±3500
MKP1	525383±268053	499308±220665
MR	2565±1110	2830±887
ODC1	71892±32249	58670±40801
P2X7	1095±432	1542±563
PBR	70854±30278	64439±29328
PREP	6715±2072	7072±2102
RGS2	632976±262593	477280±165907
S100A10	32173±9530	35819±10568
SERT	1400±1164	1711±1317
VMAT2	3469±1602	2792±1344

(SD=標準偏差)

## 【 0 3 3 5】

2 遺伝子の組合せは、対照対象に対するうつ病対象の転写物値比率を比較することによっても評価した。表 4 A にあるように、うつ病対象と対照対象との間に、特定のバイオマーカーの含有量比率の著しい差異が見い出される。

## 【 0 3 3 6】

10

20

30

## 【表 1 3】

表4A

バイオマーカー	うつ病対象群の転写物含有量比率	対照群の転写物含有量比率
ERK1	0.35	0.76
MAPK14		
IL1b	0.26	0.09
Gi2		
MAPK14	0.39	0.17
ARRB1		
ERK1	0.85	1.86
IL1b		

10

## 【0337】

より重度のうつ病患者集団の転写プロファイルの変化を評価するために、120人の重度うつ病患者から血液を取得し、表1Aから選択された遺伝子の遺伝子発現を測定した。遺伝子発現データを、単変量法により統計的に分析した。患者転写データを、196人の対照の転写データと比較し、個々の遺伝子データの代表的な散布図を図4A～4Cに示す。

20

## 【0338】

RF/SVMを使用した分類は、結果として92%の高正確度をもたらした(PPV = 89%; NPV = 94%)。遺伝子選択およびトレーニングを両方とも実施するSLRアルゴリズムの分類は、結果として93%の高正確度をもたらした(PPV = 91%; NPV = 95%)。

## 【0339】

両アルゴリズムは、データセット全体に基づいて選択された遺伝子の良好な一致を示した。各方法の統計パラメーターに基づく分類に最も重要な遺伝子として、ランダム・フォレスト分類法は合計7個の遺伝子を選択し、SLRは合計12個の遺伝子を選択した。CD8a、ERK1、MAPK14、P2X7、およびPBRを含む5個の遺伝子が、両方の方法により選択された。

30

## 【0340】

患者/対照の帰属を無作為化した後、両分類アルゴリズム(RF/SVMおよびSLR)は、実際のデータで取得されたものとは統計的に異なる正確度値をもたらし、上記に列挙されている値が偶然よりも良好であり、群は統計的に分離可能であることを示す。

## 【0341】

対象をプロファイルすることができ、表1Aに含まれている遺伝子に基づく対象の転写データを、上記に記述されているようにトレーニングされた分類アルゴリズムにかけて、重度うつ病の診断を取得した。

## 【0342】

表1Aから選択された遺伝子に関する、重度うつ病対象の転写プロファイルは、各バイオマーカー(例えば、遺伝子転写物)の含有量に基づいて表5に示されている。対照対象の転写物値は、比較のために示されている。

40

## 【0343】

【表 1 4】

表5

バイオマーカー (遺伝子の略語)	重度うつ病対象群の特 徴:含有量=バイオマー カーの平均転写物値(± SD)	対照対象群の特徴:含有 量=バイオマーカーの平 均転写物値(±SD)
ADA	3812±1365	4511±1710
ARRB1	161284±47341	297143±91094
ARRB2	79487±22860	114780±39962
CD8a	7666±4603	14693±8416
CD8b	6897±3320	8687±3880
CREB1	64463±18736	63725±16022
CREB2	71534±12311	77059±15755
DPP4	5873±2194	7169±2890
ERK1	19389±7612	39016±12900
ERK2	48236±17894	54137±18660
Gi2	97344±42195	226358±87609
Gs	185642±82731	303930±139837
GR	75411±24542	80610±26544
IL1b	27643±12046	21006±9313
IL6	153±100	182±221
IL8	38817±29253	28024±19993
INDO	5735±5467	5596±4418
MAPK14	67519±29094	51632±20341
MAPK8	11446±3231	12162±3500
MKP1	615915±307961	499308±220665
MR	2023±893	2830±887
ODC1	55085±30043	58670±40801
P2X7	769±331	1542±563
PBR	67863±24974	64439±29328
PREP	5186±1620	7072±2102
RGS2	571284±270572	477280±165907
S100A10	21812±7985	35819±10568
SERT	795±553	1711±1317
VMAT2	3073±1715	2792±1344

(SD=標準偏差)

## 【 0 3 4 4 】

平均発現レベル(転写物値)が、重度うつ病患者と対照との間で有意に異なっていた( $p < 0.05$ ) 遺伝子は以下の通りである: ADA、ARRB1、ARRB2、CD8a、CD8b、CREB2、DPP4、ERK1、Gi2、Gs、IL1b、IL8、MAPK14、MKP1、MR、P2X7、PREP、RGS2、S100A10、およびSERT(表5A)。

## 【 0 3 4 5 】

10

20

30

40

## 【表 15】

表5A:p値(p<0.05)に基づき、対照対象と比較して重度うつ病対象で有意に異なっていた遺伝子。

バイオマーカー (遺伝子の略語)	p 値
ADA	$3.2673 \times 10^{-6}$
ARRB1	$4.40419 \times 10^{-60}$
ARRB2	$1.61434 \times 10^{-27}$
CD8a	$1.92916 \times 10^{38}$
CD8b	$3.13307 \times 10^8$
CREB2	0.0000507671
DPP4	$1.25015 \times 10^7$
ERK1	$1.12946 \times 10^{-72}$
Gi2	$3.27538 \times 10^{-64}$
Gs	$1.98625 \times 10^{35}$
IL1b	$2.13924 \times 10^{-11}$
IL8	$2.00073 \times 10^{-6}$
MAPK14	$5.2042 \times 10^{-15}$
MKP1	$1.25421 \times 10^{-6}$
MR	$1.73784 \times 10^{-23}$
P2X7	$3.7121 \times 10^{-67}$
PREP	$2.72022 \times 10^{-26}$
RGS2	0.0000152985
S100A10	$2.3756 \times 10^{-53}$
SERT	$4.36216 \times 10^{-26}$

10

20

## 【0346】

これら遺伝子を、算出した -Log(p) 値 (図9) の大きさによりランク付けし、それにより、ERK1、P2X7、Gi2、ARRB1、およびS100A10などのいくつかの遺伝子についての、患者転写物値と対照値との間の著しい差異を示す。

## 【0347】

転写物値間の線形および非線形相互作用を探索するために、関連ベクター・マシン (RVM) 分類アルゴリズムを実施し、その後遺伝的アルゴリズムを、考え得る遺伝子間相互作用の空間を探索し、最もロバストで意味のある相互作用を選択するために使用した。単一遺伝子解もこのセットのアルゴリズムで検討し、患者を対照から分離するための単一遺伝子解の妥当性を確認した。ARRB1 (正確度 = 0.86) およびERK1 (正確度 = 0.85)、次いでP2X7 (正確度 = 0.82) およびGi2 (正確度 = 0.81) が、単一遺伝子解析において高度に情報量が高いことが判明する。例えば、対照に対する中程度のうつ病患者、重度うつ病患者、および双極性患者について、情報量が高い遺伝子発現データが示されている図2~5も参照されたい。

30

## 【0348】

いくつかの2遺伝子解は、うつ病患者および対照を90%以上の正確度で分類することが特定された。ERK1およびMAPK14の転写物値は、うつ病患者と対照とを92%の正確度で分類することが示される。図10には、ERK1およびMAPK14の転写物値のみに基づく、重度うつ病対象および対照の分布が示されている。うつ病対象 (表4にあるようなプロファイルを有する) の分類は、重度うつ病対象の結果と一致している。図11、12、および13には、他の2遺伝子転写プロファイル、それぞれIL1b/Gi2、MAPK14/ARRB1、およびERK1/IL1bの転写物値に基づく、重度うつ病対象および対照の分布が示されている。2遺伝子の組合せは、対照対象に対する重度うつ病対象の転写物値比率を比較することによっても評価した。重度うつ病対象と対照対象との間に含有量比率の著しい差異があることが、表5Bに見られる。

40

## 【0349】

## 【表 1 6】

表5B

バイオマーカー	重度うつ病対象群の転写物含有量比率	対照群の転写物含有量比率
ERK1	0.29	0.76
MAPK14		
IL1b	0.28	0.09
Gi2		
MAPK14	0.42	0.17
ARRB1		
ERK1	0.70	1.86
IL1b		

10

## 【0350】

双極性障害患者の転写プロファイルの同定

双極性障害患者の転写プロファイルの変化を評価するために、23人のうつ病患者（DSM-IV基準により双極性障害と確定的に診察されている20人の患者）から血液を取得し、表1Aから選択された遺伝子の遺伝子発現を測定した。遺伝子発現データを、単変量法により統計的に分析した。患者転写データを、196人の対照の転写データと比較し、個々の遺伝子データの代表的な散布図を図5A～5Cに示す。

20

## 【0351】

RF/SVMを使用した分類は、結果として94%の高正確度をもたらした（PPV = 86%；NPV = 95%）。遺伝子選択およびトレーニングを両方とも実施するSLRアルゴリズムの分類は、結果として97%の高正確度をもたらした（PPV = 90%；NPV = 99%）。

## 【0352】

両アルゴリズムは、データセット全体に基づいて選択された遺伝子の良好な一致を示し、各方法の統計パラメーターに基づく分類に最も重要な遺伝子として、ランダム・フォレスト分類法は合計3個の遺伝子を選択し、SLRは合計5個の遺伝子を選択した。Gi2、GR、およびMAPK14を含む3個の遺伝子が、両方の方法により選択された。

30

## 【0353】

患者/対照の帰属を無作為化した後、両分類アルゴリズム（RF/SVMおよびSLR）は、実際のデータで取得されたものとは統計的に異なる正確度値をもたらし、上記に列挙されている値が偶然よりも良好であり、群は統計的に分離可能であることを示す。

## 【0354】

対象をプロファイルすることができ、表1Aに含まれている遺伝子に基づく対象の転写データを、上記に記述されているようにトレーニングされた分類アルゴリズムにかけて、双極性障害の診断を取得した。

## 【0355】

各遺伝子についての双極性対象の転写のプロファイルは、各バイオマーカー（例えば、遺伝子転写物）の含有量に基づいて表6に示されている。対照対象の転写物値は、比較のために示されている。

40

## 【0356】

【表 17】

表6

バイオマーカー (遺伝子の略語)	双極性対象群の特徴: 含有量=バイオマーカー の平均転写物値(± SD)	対照対象群の特徴:含 有量=バイオマーカー の平均転写物値(± SD)
ADA	4775±1508	4511±1710
ARRB1	292298±89272	297143±91094
ARRB2	111023±39397	114780±39962
CD8a	11668±5573	14693±8416
CD8b	7998±3841	8687±3880
CREB1	62347±18282	63725±16022
CREB2	79456±16778	77059±15755
DPP4	7618±3077	7169±2890
ERK1	34901±15116	39016±12900
ERK2	57832±21427	54137±18660
Gi2	192417±98987	226358±87609
Gs	304202±171505	303930±139837
GR	124054±42231	80610±26544
IL1b	21577±13468	21006±9313
IL6	173±78	182±221
IL8	24568±19226	28024±19993
INDO	5428±3847	5596±4418
MAPK14	66946±25751	51632±20341
MAPK8	12584±3060	12162±3500
MKP1	501068±251853	499308±220665
MR	3409±1094	2830±887
ODC1	67672±50925	58670±40801
P2X7	1322±418	1542±563
PBR	64761±29660	64439±29328
PREP	6806±1677	7072±2102
RGS2	499864±264854	477280±165907
S100A10	42063±12765	35819±10568
SERT	1435±710	1711±1317
VMAT2	2736±1050	2792±1344

(SD=標準偏差)

## 【0357】

境界性人格障害患者の転写プロファイルの同定

境界性人格障害患者の転写プロファイルの変化を評価するために、21人の境界性人格障害患者から血液を取得し、表1Aから選択された遺伝子の遺伝子発現を測定した遺伝子発現データを、単変量法により統計的に分析した。患者転写データを、196人の対照の転写データと比較し、個々の遺伝子データの代表的な散布図を図6A～6Cに示す。

## 【0358】

RF(選択)およびSVM(トレーニング)を使用した分類は、結果として97%の高正確度をもたらした(PPV=87%; NPV=98%)。遺伝子選択およびトレーニングを両方とも実施するSLRアルゴリズムの分類は、結果として98%の高正確度をもたらした(PPV=90%; NPV=100%)。

## 【0359】

両アルゴリズムは、データセット全体に基づいて選択された遺伝子の良好な一致を示し、各方法の統計パラメータに基づく分類に最も重要な遺伝子として、ランダム・フォレスト分類法は合計5個の遺伝子を選択し、SLRは合計4個の遺伝子を選択した。Gi2、GR、MAPK14、およびMRを含む4個の遺伝子が、両方の方法により選択された。

。

10

20

30

40

50

## 【0360】

患者 / 対照の帰属を無作為化した後、両分類アルゴリズム (RF / SVM および SLR) は、実際のデータで取得されたものとは統計的に異なる正確度値をもたらし、上記に列挙されている値が偶然よりも良好であり、群は統計的に分離可能であることを示す。

## 【0361】

対象をプロファイルすることができ、表 1 A に含まれている遺伝子に基づく対象の転写データを、上記に記述されているようにトレーニングされた分類アルゴリズムにかけて、境界性人格障害の診断を取得した。

## 【0362】

PTSD 患者の転写プロファイルの同定

急性 PTSD の患者、軽減された PTSD の患者、および心的外傷イベントに曝されたが PTSD を発症していない個体群の転写プロファイルを評価した。これらの群を合わせて評価することにより、急性 PTSD と関連する発現変化を識別する機会、ならびにこの疾患からの回復またはこの疾患に対する抵抗性に関連する可能性のある差異を定義する機会が提供される。遺伝子発現データを、単変量法により統計的に分析した。66 人の急性 PTSD 患者からの患者転写データを、196 人の対照の転写データと比較し、個々の遺伝子データの代表的な散布図を図 7 A ~ 7 C に示す。

## 【0363】

RF (選択) および SVM (トレーニング) を使用し、対照対象と比較した急性 PTSD 患者の分類は、結果として 77 % の正確度をもたらした (PPV = 64 % ; NPV = 82 % )。遺伝子選択およびトレーニングを両方とも実施する SLR アルゴリズムを使用した分類は、結果として 84 % の正確度をもたらした (PPV = 77 % ; NPV = 87 % )。SLR アルゴリズムは、このセットの試験データを使用する SVM アルゴリズムより優れている。各分類アルゴリズムを、無作為化 (並べ替え) 型のデータセットと比較し、SLR は、並べ替えデータセットを使用して 73 % の正確度値をもたらした (PPV = 39 % ; NPV = 75 % )。統計分析は、実際のデータ対無作為化データで取得された SLR 正確度値が異なることを示し、群が分離可能であることを示した。

## 【0364】

SVM は、並べ替えデータセットを使用して 73 % の正確度値をもたらし (PPV = 10 % ; NPV = 75 % )、並べ替え (無作為化) データでは減少傾向を示した。注目すべきは、SVM アルゴリズムで実際のデータを使用する PPV (この疾患の患者を陽性と予測する能力) が、並べ替えデータの 10 % 精度と比較して、60 % より良好であり、実際のデータを使用してトレーニングされたアルゴリズムは、無作為予測より優れていることを示していることである。

## 【0365】

SLR は、急性 PTSD 患者対対照のデータセット全体に基づく分類のために最も重要な遺伝子として、以下の 10 個の遺伝子の合計を選択した: ARR B 1、ARR B 2、CD 8 b、ER K 2、I D O、I L - 6、MR、O D C 1、P R E P、および R G S 2。

## 【0366】

対象をプロファイルすることができ、表 1 A に含まれている遺伝子に基づく対象の転写データを、上記に記述されているようにトレーニングされた分類アルゴリズムにかけて、急性 PTSD の診断を取得した。

## 【0367】

RF (選択) および SVM (トレーニング) を使用し、対照対象と比較した軽減された PTSD 患者の分類は、結果として 81 % の正確度をもたらした (PPV = 59 % ; NPV = 85 % )。遺伝子選択およびトレーニングを両方とも実施する SLR アルゴリズムの分類は、結果として 80 % の正確度をもたらした (PPV = 33 % ; NPV = 86 % )。しかしながら、分類アルゴリズムをこのデータセットの無作為化型で実行した場合、SVM および SLR は、それぞれ 82 % および 81 % の正確度値をもたらした。これらの値は、実際のデータで取得された値と統計的に異なっておらず、アルゴリズムがこれら群を高

10

20

30

40

50

信頼性で分離することができないことを示している。分離ができないため、この比較に関する遺伝子リストは報告されていない。臨床的見地からすると、アルゴリズムが対照と軽減患者とを識別することができないことは、これらの群間の生物学的な相異が欠如しているため予測されている。軽減患者が疾患の症状をほぼ示さないため、軽減患者の遺伝子発現レベルが正常レベルに戻り、そのためアルゴリズムが有効に群を分離するのが妨げられると仮定することは合理的である。

#### 【0368】

RF（選択）およびSVM（トレーニング）を使用し、対照対象と比較した、心的外傷を受けたがPTSDを発症しなかった対象の分類は、結果として74%の正確度をもたらした（PPV = 61%；NPV = 79%）。遺伝子選択およびトレーニングを両方とも実施するSLRアルゴリズムの分類は、結果として73%の正確度をもたらした（PPV = 59%；NPV = 80%）。多変量解析を無作為化データセットで実施した場合、RF/SVMおよびSLR分類アルゴリズムは両方とも、実際のデータで取得されたものとは統計的に異なる正確度の値をもたらし、上記に報告されている値が偶然よりも良好であり、群は分離可能であることを示す。

10

#### 【0369】

各方法の統計パラメーターに基づき、心的外傷患者および対照からのデータセット全体を使用した分類に最も重要な遺伝子として、ランダム・フォレスト分類法は合計14個の遺伝子を選択し、SLRは合計13個の遺伝子を選択した。ARRB2、CREB1、ERK2、Gs、IL-6、MKP1、およびRGS2を含む7個の遺伝子が、両方の方法により選択された。

20

#### 【0370】

これら個体はPTSDと診断されていないが、アルゴリズムは、正確度、PPV、およびNPV値が、本明細書に提示されている他の比較のいくつかより低いものの、依然としてそれらを対照と識別することができる。興味深いことには、急性PTSD患者に由来するSLR遺伝子リストの遺伝子のうちの6つは、PTSDではない心的外傷患者の対応するリストの遺伝子と一致する（ARRB2、CD8b、ERK2、MR、IL-6、およびRGS2）。心的外傷を受けた患者は、まだ疾患を発症していない一方で、いくつかの遺伝子発現プロファイルを罹患患者と共有しており、心的外傷を受けた患者にリスクがある可能性を示している。

30

#### 【0371】

対象をプロファイルすることができ、表1Aに含まれている遺伝子に基づく対象の転写データを、上記に記述されているようにトレーニングされた分類アルゴリズムにかけて、PTSDではない心的外傷の診断を取得した。

#### 【0372】

##### 7 引用参考文献

本明細書に引用された参考文献はすべて、あたかも個々の刊行物または特許または特許出願の各々が、具体的におよび個々に参照によりその全体が、あらゆる目的のために本明細書中に組み込まれると示されているのと同じ程度に、それらの全体があらゆる目的のために参照により本明細書に組み込まれる。

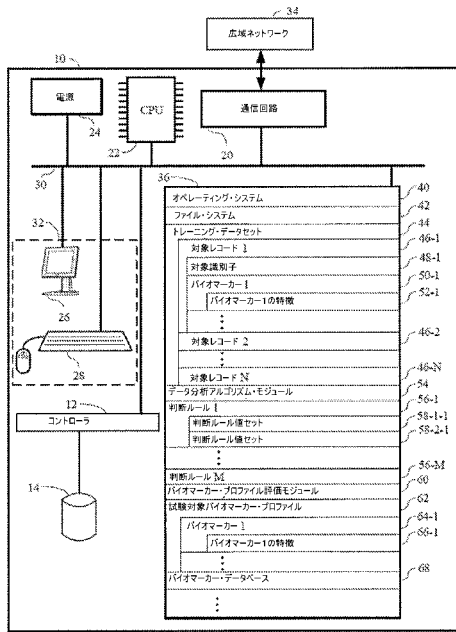
40

#### 【0373】

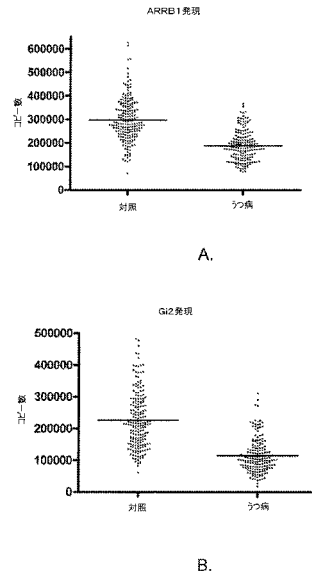
##### 8 改変

当業者であれば明白になるように、本発明の多数の改変および変異を、本発明の趣旨および範囲から逸脱せずになすことができる。本明細書に記述されている特定の実施形態は、例として提供されているに過ぎず、本発明は、そのような特許請求の範囲に与えられる全範囲の等価物に加えて、添付の特許請求の範囲の条件にのみ制限されるべきである。

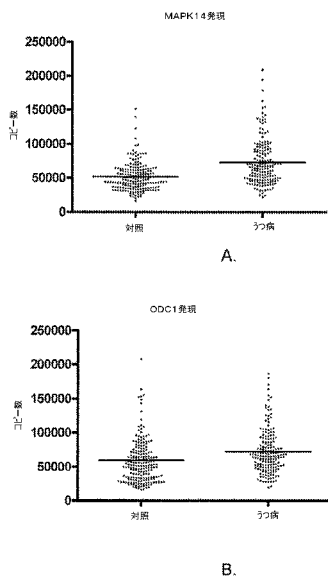
【 図 1 】



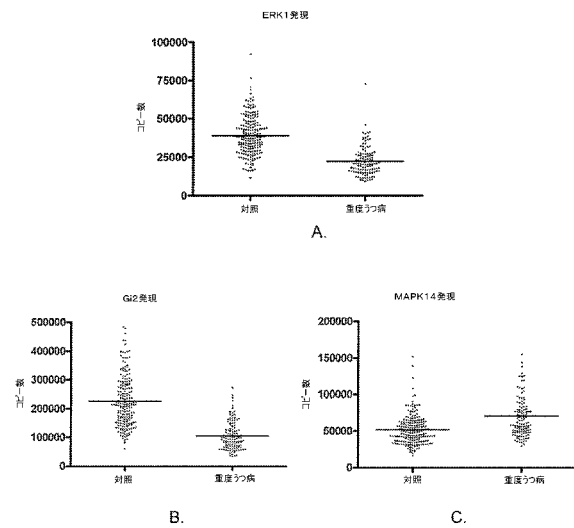
【 図 2 】



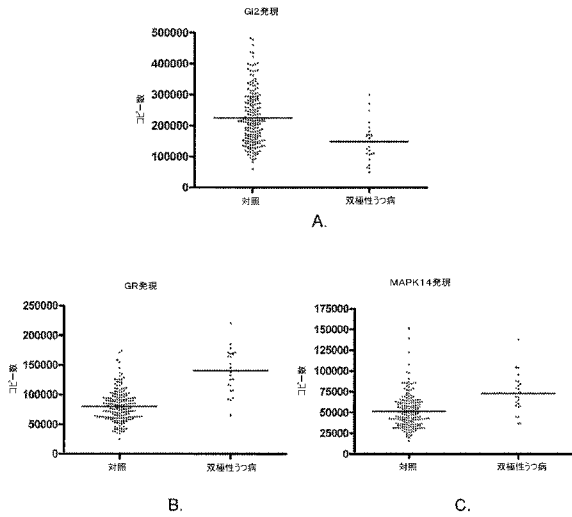
【 図 3 】



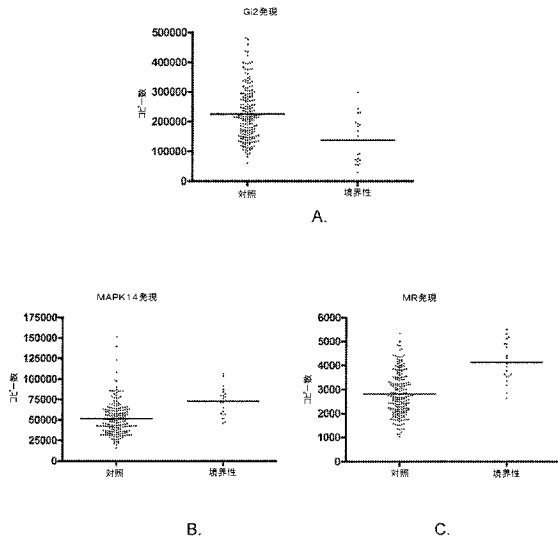
【 図 4 】



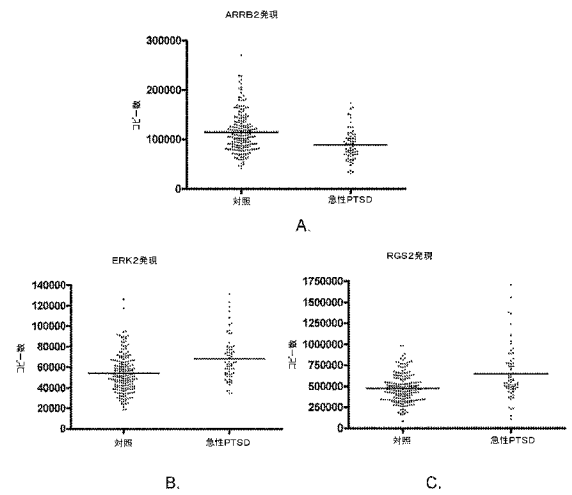
【 図 5 】



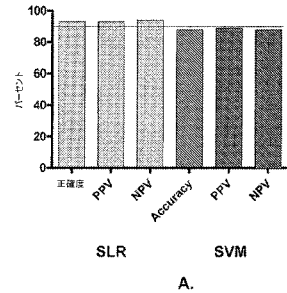
【 図 6 】



【 図 7 】



【 図 8 】

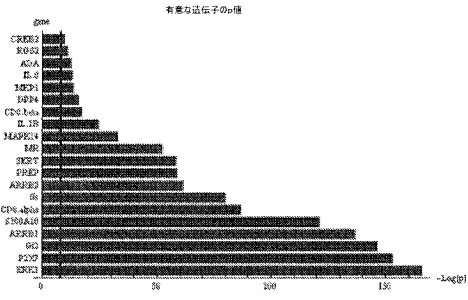


A.

SLR	RF
ARRB1	ARRB1
ARRB2	ARRB2
CD8a	CD8a
CREB1	CREB1
CREB2	CREB2
ERK2	ERK2
Gi2	Gi2
MAPK14	MAPK14
ODC1	ODC1
P2X7	P2X7
PBR	PBR
ADA	ERK1
CD8b	IL-1b
IDO	RGS2
MAPK8	
SERT	
S100A10	

B.

【 図 9 】



【 図 10 】

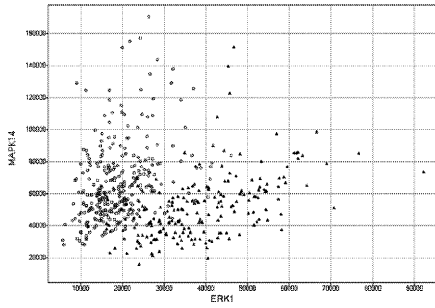


Figure 10.

【 図 11 】

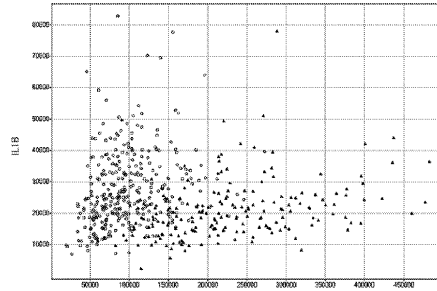


Figure 11.

【 図 12 】

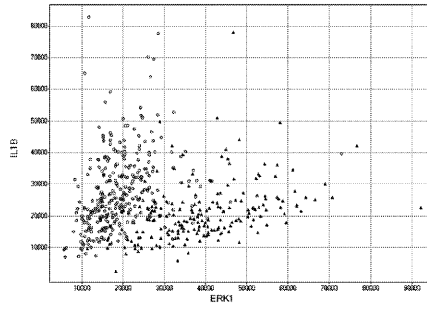
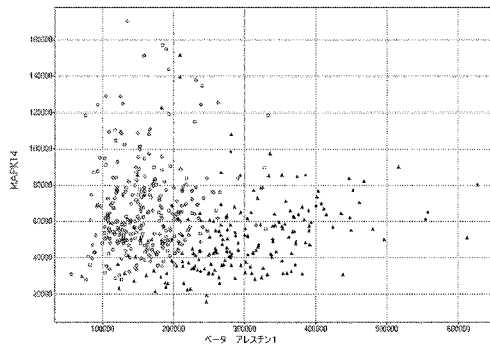


Figure 12.

【 図 13 】



【配列表】

2012501181000001.app

## 【 国際調査報告 】

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2009/055144

A. CLASSIFICATION OF SUBJECT MATTER		
Int. Cl.		
<i>C12Q 1/68</i> (2006.01) <i>G01N 33/68</i> (2006.01) <i>G01N 33/53</i> (2006.01) <i>G06F 19/00</i> (2006.01)		
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)		
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) EPODOC, Medline & WPI; Keywords: biomarker, ERK1, MAPK14, Gi2, diagnostic, screen, micro-array, mental disorder, depression, bipolar, anxiety, panic, dysthenia and similar terms		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P, Y	WO 2008/124428 A1 (INDIANA UNIVERSITY RESEARCH AND TECHNOLOGY CORPORATION) 16 October 2008 See Abstract, [15], [18]-[21], [52]-[58], [87]-[93], [95]-[98], [106], [108], [112], examples 1-2	1-11, 13, 14, 17-35
A	LOFTIS J. M. et al., "Depressive symptoms in patients with chronic hepatitis C are correlated with elevated plasma levels of interleukin-1 $\beta$ and tumor necrosis factor- $\alpha$ ", <i>Neurosci. Lett.</i> , 2008 January 17, vol. 430, no. 3, pages 264-268	
Y	THE WELLCOME TRUST CASE CONTROL CONSORTIUM, "Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls", <i>Nature</i> , 2007 June 7, vol. 447, no. 7145, pages 661-678 See Abstract, "Geographical variation and population structure" on pages 662-663 and "bipolar disorder" on pages 666-667	1-11, 13, 14, 17-35
<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 application or patent 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 22 January 2010		Date of mailing of the international search report - 1 FEB 2010
Name and mailing address of the ISA/AU AUSTRALIAN PATENT OFFICE PO BOX 200, WODEN ACT 2606, AUSTRALIA E-mail address: pct@ipaaustralia.gov.au Facsimile No. +61 2 6283 7999		Authorized officer NATHALIE TOCHON-DANGUY AUSTRALIAN PATENT OFFICE (ISO 9001 Quality Certified Service) Telephone No : 61262833107

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2009/055144

**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 Supplemental Box I

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 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.: 1-14, 17-35 (in part) relating to the biomarkers ERK1, MAPK 14 and Gi2 (invention 9, 11 and 18 respectively)
4.  No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

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

## INTERNATIONAL SEARCH REPORT

International application No. PCT/US2009/055144
--

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	RIBEIRO L. et al., "The brain-derived neurotrophic factor rs6265 (Val66Met) polymorphism and depression in Mexican-Americans", <i>NeuroReport</i> , 2007 August 6, vol. 18, no.12, pages 1291-1293	
A	LIN P.-I. & Mitchell B. D., "Approaches for unravelling the joint genetic determinants of Schizophrenia and Bipolar Disorder", <i>Schizophrenia Bulletin</i> , 2008, vol. 34, no. 4, pages 791-797	
A	MÜLLER D. J. et al., "Brain-derived neurotrophic factor (BDNF) gene and rapid-cycling bipolar disorder", <i>British Journal of Psychiatry</i> , 2006, vol. 189, pages 317-323	
Y	WO 2008/079269 A2 (NIKOLSKY Y. et al.) 3 July 2008 See Abstract, pages 7-8, 10-12, 26; 43, Figures 2, 10, 21 & 22 and claims 1, 14 & 18	1-11, 13, 14, 17-35
Y	KYOSSEVA S.V. et al., "Differential and region-specific activation of mitogen-activated protein kinases following chronic administration of phencyclidine in rat brain", <i>Neuropsychopharmacology</i> , 2001, Vol. 24, pages 267-277 See whole document	1-11, 14, 17-35
Y	SAKAKIBARA H. et al., "Effects of forced swimming stress on rat brain function", <i>The Journal of Medical Investigation</i> , 2005, vol. 52 Supplement, pages 300-301 See whole document	1-11, 14, 17-35
Y	TRONSON N.C. et al., "Regulatory mechanisms of fear extinction and depression-like behavior", <i>Neuropsychopharmacology</i> , 2008 June, vol. 33, no. 7, pages 1570-1583 See whole document	1-11, 14, 17-35
Y	BENEDETTO B.D. et al., "Activation of ERK/MAPK in the lateral amygdala of the mouse is required for acquisition of a fear-potentiated startle response", <i>Neuropsychopharmacology</i> , 2009, vol. 34, pages 356-366, Published online 23 April 2008 See whole document	1-11, 14, 17-35
Y	BEZCHLIBNYK Y. and YOUNG L.T., "The neurobiology of bipolar disorder: focus on signal transduction pathways and the regulation of gene expression", <i>Can J Psychiatry</i> , 2002, vol. 47, no. 2, pages 135-148 Abstract, Table 1	1-10, 12, 17-35
Y	AVISSAR S. et al., "Dynamics of ECT Normalization of low G protein function and immunoreactivity in mononuclear leukocytes of patients with major depression" <i>The American Journal of Psychiatry</i> , 1998, vol. 155, no. 5, pages 666-667 See whole document especially Abstract, Figures	1-10, 12, 17-35

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US2009/055144

C (Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 2008/144371 A1 (THE BOARD OF TRUSTEES OF THE LELAND STANFORD JUNIOR UNIVERSITY) 27 November 2008 See pages 3, 51, 52, examples	1-11, 13, 14, 17-35
A	DONATI R.J. et al., "Postmortem brain tissue of depressed suicides reveals increased G $\alpha$ localization in lipid raft domains where it is less likely to activate adenylyl cyclase", The Journal of Neuroscience, March 2008, vol. 28, no. 12, pages 3042-3050	
Y	VAWTER M.P. et al., "Microarray screening of lymphocyte gene expression differences in multiplex schizophrenia pedigree", Schizophrenia Research, 2004, vol. 67, pages 41-52 Abstract, method	1-11, 13, 14, 17-35
A	STEIN M.B. et al., "G-protein level quantification in platelets and leukocytes from patients with panic disorder", Neuropsychopharmacology, 1996, vol. 15, no. 2, pages 180-186	
A	COLIN S.F. et al., "Chronic lithium regulates the expression of adenylyl cyclase and Gi-protein $\alpha$ subunit in rat cerebral cortex", PNAS, 1991, vol. 88, pages 10634-10637	
Y	WO 2007/044094 A1 (BLANCHETTE ROCKEFELLER NEUROSCIENCES INSTITUTE) 19 April 2007 Whole document	1-11, 13, 14, 17-35
Y	WO 2003/068958 A1 (JAPAN SCIENCE AND TECHNOLOGY CORPORATION) 21 August 2008 Abstract, Pages 9, 16, 19, 20	1-11, 13, 14, 17-35
Y	PRASAD H.C. et al., "Human serotonin transporter variants display altered sensitivity to protein kinase G and p38 mitogen-activated protein kinase", PNAS, 2005, vol. 102, no. 32, pages 11545-11550 Abstract, Introduction	1-11, 13, 17-35

## INTERNATIONAL SEARCH REPORT

International application No. PCT/US2009/055144
--

**Supplemental Box I**

(To be used when the space in any of Boxes I to IV is not sufficient)

**Continuation of Box No: III**

This International Application does not comply with the requirements of unity of invention because it does not relate to one invention or to a group of inventions so linked as to form a single general inventive concept.

In assessing whether there is more than one invention claimed, I have given consideration to those features which can be considered to potentially distinguish the claimed combination of features from the prior art. Where different claims have different distinguishing features they define different inventions.

This International Searching Authority has found that there are different inventions as follows:

- Invention 1: Claims 1-35 are directed to a method of diagnosing an affective disorder in a subject comprising evaluating the presence of adenosine deaminase plus another biomarker chosen from Table 1A. It is considered that the biomarker adenosine deaminase and its use to diagnose an affective disorder comprises a first distinguishing feature.
- Invention 2: Claims 1-35 are directed to a method of diagnosing an affective disorder in a subject comprising evaluating the presence of beta-arrestin 1 plus another biomarker chosen from Table 1A. It is considered that beta-arrestin 1 and its use to diagnose an affective disorder comprises a second distinguishing feature.
- Inventions 3 – 29: *Mutatis mutandis* for the remaining biomarkers listed in Table 1A.

PCT Rule 13.2, first sentence, states that unity of invention is only fulfilled when there is a technical relationship among the claimed inventions involving one or more of the same or corresponding special technical features. PCT Rule 13.2, second sentence, defines a special technical feature as a feature which makes a contribution over the prior art.

It should be noted that given that claim 1 is directed to combinations of at least two biomarkers chosen from the 29 listed in Table 1A, this allows for around half a billion possible combinations. Hence, for the purposes of expediency, each individual biomarker will be considered a distinguishing feature rather than each combination of 2 or more (as claimed).

The only feature common to all of the claims is the use of biomarkers to diagnose affective disorders in a subject. However this concept is not novel and/or not inventive in the light of a vast number of documents of which the following should be considered merely exemplary:

D1: WO 2008/124428 (INDIANA UNIVERSITY RESEARCH AND TECHNOLOGY CORPORATION) 16 October 2008  
 D2: LOFTIS J. M. et al., Neurosci. Lett., 2008 January 17, vol. 430, no. 3, pages 264–268  
 D3: THE WELLCOME TRUST CASE CONTROL CONSORTIUM, Nature, 2007 June 7, vol. 447, no. 7145, pages 661–678.  
 D4: RIBEIRO L. et al., NeuroReport, 2007 August 6, vol. 18, no.12, pages 1291–1293  
 D5: LIN P.-I. & Mitchell B. D., Schizophrenia Bulletin, 2008, vol. 34, no. 4, pages 791–797  
 D6: MÜLLER D. J. et al., British Journal of Psychiatry, 2006, vol. 189, pages 317–323

Further, given the nature of the claimed subject matter, it is appropriate to apply the Markush practice. It is apparent that each of the biomarkers possesses a common activity or property (namely, usefulness in diagnosing an affective disorder); however, there is no significant structural element shared by each of the biomarkers and, further, each of the biomarkers do not belong to a recognised class of compounds. Thus the Markush criteria are not satisfied and the biomarkers cannot be considered to be unified.

This means that the common feature can not constitute a special technical feature within the meaning of PCT Rule 13.2, second sentence, since it makes no contribution over the prior art. Because the common feature does not satisfy the requirement for being a special technical feature it follows that it cannot provide the necessary technical relationship between the identified inventions. Therefore the claims do not satisfy the requirement of unity of invention *a posteriori*.

**INTERNATIONAL SEARCH REPORT**

Information on patent family members

International application No.

**PCT/US2009/055144**

This Annex lists the known "A" publication level patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent Document Cited in Search Report		Patent Family Member					
WO	2008124428	NONE					
WO	2008079269	EP	2094719				
WO	2008144371	NONE					
WO	2007044094	CA	2625300	CN	101322032	EP	1934618
		EP	2031398	KR	20080066785	US	7595167
		US	2007082366	US	2008221042	US	2009029873
		WO	2007043998	WO	2007047029	WO	2009126232
WO	2003068958	JP	2003235557				

Due to data integration issues this family listing may not include 10 digit Australian applications filed since May 2001.

END OF ANNEX

## 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

(72) 発明者 アントーニエーフィク・イリーナ  
アメリカ合衆国、ニュージャージー州 07430、マーワー、リオ・ヴィスタ・ドライブ、1201

(72) 発明者 タム・ジョーゼフ  
アメリカ合衆国、ニュージャージー州 07506、ハウソーン、プエナ・ヴィスタ・アヴェニュー、103

(72) 発明者 アーティーミシシ・ロマン  
アメリカ合衆国、ニュージャージー州 08822、フレミントン、スタントン・ロード、114

(72) 発明者 ジェラル・クリストフ・ペ・ジェ  
アメリカ合衆国、ニュージャージー州 07430、マーワー、リオ・ヴィスタ・ドライブ、1201

(72) 発明者 ヴィスティセン・ヤン・バストホルム  
デンマーク国、9600 アールス、アートフテン、6

Fターム(参考) 2G045 AA25 CA02 CA11 CA12 CA15 CA16 CA20 CA25 DA14 FB01  
FB02 FB12 GC15 JA01 JA06  
4B063 QA01 QA18 QA19 QQ02 QQ42 QQ52 QQ53 QR08 QR32 QR36  
QR42 QR50 QR55 QR62 QR72 QS25 QS28 QS34 QS36 QS39  
QX02 QX10

专利名称(译)	用于测量生物标志物谱的系统和方法		
公开(公告)号	<a href="#">JP2012501181A</a>	公开(公告)日	2012-01-19
申请号	JP2011525187	申请日	2009-08-27
[标]申请(专利权)人(译)	ハーランドベックアクチエゼルスカベット		
申请(专利权)人(译)	她的灵北活化埃泽尔·斯卡赌注		
[标]发明人	アントーニエーフィクイリーナ タムジョーゼフ アーティーミシンロマン ジェラルクリストフベジェ ヴィステイセンヤンバストホルム		
发明人	アントーニエーフィクイリーナ タム・ジョーゼフ アーティーミシン・ロマン ジェラル・クリストフ・ベ・ジェ ヴィステイセン・ヤン・バストホルム		
IPC分类号	C12Q1/68 G01N33/53 G01N33/68 G06F19/24		
CPC分类号	C12Q1/6883 C12Q2600/158 G16B40/00		
FI分类号	C12Q1/68.ZNA.A G01N33/53.M G01N33/68		
F-TERM分类号	2G045/AA25 2G045/CA02 2G045/CA11 2G045/CA12 2G045/CA15 2G045/CA16 2G045/CA20 2G045/CA25 2G045/DA14 2G045/FB01 2G045/FB02 2G045/FB12 2G045/GC15 2G045/JA01 2G045/JA06 4B063/QA01 4B063/QA18 4B063/QA19 4B063/QQ02 4B063/QQ42 4B063/QQ52 4B063/QQ53 4B063/QR08 4B063/QR32 4B063/QR36 4B063/QR42 4B063/QR50 4B063/QR55 4B063/QR62 4B063/QR72 4B063/QS25 4B063/QS28 4B063/QS34 4B063/QS36 4B063/QS39 4B063/QX02 4B063/QX10		
代理人(译)	上西 克礼		
优先权	61/092270 2008-08-27 US		
其他公开文献	JP2012501181A5		
外部链接	<a href="#">Espacenet</a>		

摘要(译)

本发明涉及用于诊断患有情感障碍的患者的方法和系统。该方法还可用于预测受试者中情感障碍的易感性。

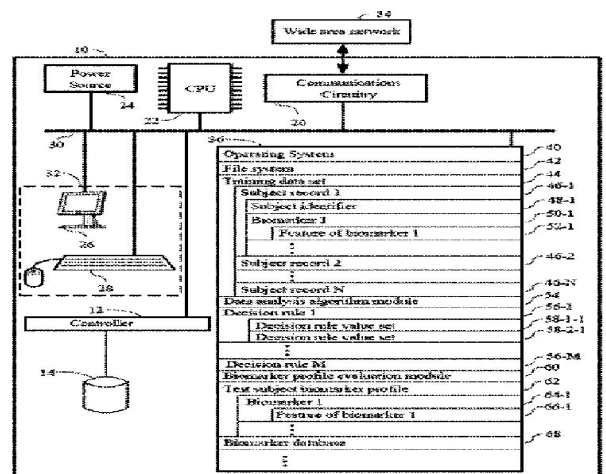


Figure 1.