

(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2015-502740

(P2015-502740A)

(43) 公表日 平成27年1月29日 (2015.1.29)

(51) Int.Cl.	F I	テーマコード (参考)
C 1 2 Q 1/68 (2006.01)	C 1 2 Q 1/68 Z N A A	4 B O 2 4
G O 1 N 33/53 (2006.01)	G O 1 N 33/53 D	4 B O 6 3
G O 1 N 33/543 (2006.01)	G O 1 N 33/53 N	
C 1 2 N 15/09 (2006.01)	G O 1 N 33/53 X	
	G O 1 N 33/53 M	
審査請求 未請求 予備審査請求 未請求 (全 148 頁) 最終頁に続く		

(21) 出願番号	特願2014-537335 (P2014-537335)	(71) 出願人	599132904
(86) (22) 出願日	平成24年10月19日 (2012.10.19)		ネステク ソシエテ アノニム
(85) 翻訳文提出日	平成26年6月18日 (2014.6.18)		スイス国, ブベイ, アブニュー ネスレ
(86) 国際出願番号	PCT/US2012/061202		5 5
(87) 国際公開番号	W02013/059732	(74) 代理人	100088155
(87) 国際公開日	平成25年4月25日 (2013.4.25)		弁理士 長谷川 芳樹
(31) 優先権主張番号	61/550, 293	(74) 代理人	100107456
(32) 優先日	平成23年10月21日 (2011.10.21)		弁理士 池田 成人
(33) 優先権主張国	米国 (US)	(74) 代理人	100162352
(31) 優先権主張番号	61/553, 853		弁理士 酒巻 順一郎
(32) 優先日	平成23年10月31日 (2011.10.31)	(74) 代理人	100140453
(33) 優先権主張国	米国 (US)		弁理士 戸津 洋介
(31) 優先権主張番号	61/567, 096		
(32) 優先日	平成23年12月5日 (2011.12.5)		
(33) 優先権主張国	米国 (US)		
		最終頁に続く	

(54) 【発明の名称】 炎症性腸疾患の診断を改善するための方法

(57) 【要約】

1つ又は複数の血清-遺伝子-炎症マーカーの存在、非存在、レベル及び／又は遺伝子型を検出することによって、炎症性腸疾患（IBD）及び亜型（例えば、潰瘍性大腸炎（UC）、クローン病（CD））を予測及び診断するための方法及びシステムが開示される。これらの方法を使用すれば、IBD対非IBDの診断を提供すること、CD及びDCに関して不確定のIBDを除外すること、及びCDとDCとを高い精度で鑑別することが可能となる。

【選択図】 なし

【特許請求の範囲】

【請求項 1】

個体における炎症性腸疾患（IBD）及び／又はその臨床亜型を診断するための方法であって、

（a）前記個体から得たサンプルを分析して、前記サンプル中の、血清学的マーカー、遺伝子マーカー、炎症マーカー、及びそれらの組み合わせからなる群から選択される 1 つ又は複数のマーカーの存在、レベル又は遺伝子型を決定して、マーカープロファイルを得るステップと、

（b）第 1 のランダムフォレスト統計解析を前記マーカープロファイルに適用して、前記マーカープロファイルが IBD サンプルであるか非 IBD サンプルであるかの決定を得るステップと、

（c）任意選択で、IBD サンプルとして指定された前記サンプルに決定木又はルールセットを適用して、前記 IBD サンプルが不確定サンプルとしてカテゴリー化されるかどうかを決定するステップと、

（d）任意選択で、第 2 のランダムフォレスト統計解析を前記 IBD サンプルに適用して、IBD の臨床亜型を決定するステップと、を含む方法。

【請求項 2】

前記炎症マーカーは、急性期タンパク質、アポリポタンパク質、成長因子、細胞接着分子、及びそれらの組み合わせからなる群から選択される、請求項 1 に記載の方法。

【請求項 3】

前記血清学的マーカーは、抗好中球抗体、抗サッカロマイセス・セレビスエ抗体、抗微生物抗体、及びそれらの組み合わせからなる群から選択される、請求項 1 に記載の方法。

【請求項 4】

前記遺伝子マーカーは、ATG16L1、ECM1、NKX2-3、STAT3、及びそれらの組み合わせからなる群から選択される、請求項 1 に記載の方法。

【請求項 5】

前記抗好中球抗体は、抗好中球細胞質抗体（ANCA）、核周囲型抗好中球細胞質抗体（pANCA）、pANCA2、及びそれらの組み合わせからなる群から選択される、請求項 2 に記載の方法。

【請求項 6】

前記抗サッカロマイセス・セレビスエ抗体は、抗サッカロマイセス・セレビスエ免疫グロブリン A（ASCA-IgA）、抗サッカロマイセス・セレビスエ免疫グロブリン G（ASCA-IgG）、及びそれらの組み合わせからなる群から選択される、請求項 2 に記載の方法。

【請求項 7】

前記抗微生物抗体は、抗外膜タンパク質 C（抗OmpC）抗体、抗フラジェリン抗体、及びそれらの組み合わせからなる群から選択される、請求項 2 に記載の方法。

【請求項 8】

前記急性期タンパク質は C 反応性タンパク質（CRP）である、請求項 2 に記載の方法。

【請求項 9】

前記アポリポタンパク質は血清アミロイド A タンパク質（SAA）である、請求項 2 に記載の方法。

【請求項 10】

前記成長因子は血管内皮成長因子（VEGF）である、請求項 2 に記載の方法。

【請求項 11】

前記細胞接着分子は、細胞間接着分子 1（ICAM-1）、血管細胞接着分子 1（VCAM-1）、及びそれらの組み合わせからなる群から選択される、請求項 2 に記載の方法。

10

20

30

40

50

【請求項 12】

前記血清学的マーカーは、pANCA、pANCA2、ANCA、CRP、SAA、VEGF、ICAM-1、VCAM-1、及びそれらの組み合わせからなる群から選択される、請求項1に記載の方法。

【請求項 13】

前記血清学的マーカーは、ASCA-IgA、ASCA-IgG、抗OmpC抗体、抗CBir-1抗体、抗Fla2抗体、抗FlaX抗体、及びそれらの組み合わせからなる群から選択される、請求項1に記載の方法。

【請求項 14】

前記血清学的マーカー又は炎症マーカーの存在又はレベルが、ハイブリダイゼーションアッセイ、増幅ベースのアッセイ、免疫アッセイ、又は免疫組織化学的アッセイを用いて検出される、請求項1に記載の方法。

10

【請求項 15】

前記遺伝子マーカーは、表A～Eに記載されている遺伝子のうちの少なくとも1つである、請求項1に記載の方法。

【請求項 16】

前記遺伝子マーカーの遺伝子型が、前記遺伝子マーカーにおける一塩基多型(SNP)の有無に関するジェノタイピングによって検出される、請求項1に記載の方法。

【請求項 17】

前記SNPは、表B～Eに記載されているSNPのうちの少なくとも1つである、請求項16に記載の方法。

20

【請求項 18】

前記遺伝子マーカーは、ATG16L1、STAT3、NKX2-3、ECM1、及びそれらの組み合わせからなる群から選択される、請求項16に記載の方法。

【請求項 19】

少なくとも2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17又は18種のマーカーの存在、レベル又は遺伝子型を検出することによって、前記マーカープロファイルが決定される、請求項1に記載の方法。

【請求項 20】

第1のランダムフォレストアルゴリズムのためのマーカーが、ANCA、ASCA-A、ASCA-G、pANCA、pANCA2、抗FlaX抗体、SAA、抗Fla2抗体、ICAM、抗OmpC抗体、抗CBir1抗体、VCAM、CRP、NKX2-3、ATG16L1、STAT3、ECM1、VEGF、及びそれらの組み合わせからなる群から選択される、請求項1に記載の方法。

30

【請求項 21】

サンプルは非IBDサンプルであると決定される、請求項1に記載の方法。

【請求項 22】

第1のランダムフォレストアルゴリズムの後にサンプルがIBDサンプルとして指定され、決定木又はルールセットを使用して、サンプルについて、ANCAレベル $\geq$ Q3を有し、pANCA2陽性であり、抗CBir1抗体又は抗Fla2抗体又は抗FlaX抗体が $\geq$ Q3であるかをさらに決定し、次いでIBDサンプルを不確定サンプルとして指定する、請求項20に記載の方法。

40

【請求項 23】

第1のランダムフォレストアルゴリズムの後にサンプルがIBDサンプルとして指定され、決定木又はルールセットを使用して、サンプルについて、pANCA2陽性であり、抗CBir1抗体、抗Fla2抗体及び抗FlaX抗体から選択される3種のマーカーのうちの2種が $\geq$ Q3であるかをさらに決定し、次いでIBDサンプルを不確定として指定する、請求項20に記載の方法。

【請求項 24】

サンプルは不確定サンプルであると決定される、請求項22又は23に記載の方法。

50

【請求項 25】

サンプルに関して不確定が除外される、請求項 22 又は 23 に記載の方法。

【請求項 26】

第 1 のランダムフォレストアルゴリズムの後にサンプルが IBD サンプルとして指定され、不確定が除外され、第 2 のランダムフォレストアルゴリズムを使用して、IBD サンプルがクローン病 (CD) であるか又は潰瘍性大腸炎 (UC) であるかを決定するステップをさらに含む、請求項 20 に記載の方法。

【請求項 27】

第 2 のランダムフォレストアルゴリズムのためのマーカーが、ANCA、ASCA-A、ASCA-G、pANCA、pANCA2、抗FlaX抗体、SAA、抗Fla2抗体、ICAM、抗OmpC抗体、抗CBir1抗体、VCAM、CRP、NKX2-3、ATG16L1、STAT3、ECM1、VEGF、及びそれらの組み合わせからなる群から選択される、請求項 26 に記載の方法。

【請求項 28】

決定が CD である、請求項 27 に記載の方法。

【請求項 29】

決定が UC である、請求項 27 に記載の方法。

【請求項 30】

前記サンプルは、血清、血漿、全血及び便からなる群から選択される、請求項 1 に記載の方法。

【請求項 31】

前記マーカープロファイルは、前記 1 つ又は複数のマーカーの各々に対する四分位数スコアを合計することによって得られる、前記個体に対する合計四分位数スコア (QSS) を含む、請求項 1 に記載の方法。

【請求項 32】

前記第 1 及び／又は前記第 2 のランダムフォレストは、IBD 及び／又は IBD の臨床亜型の既知の転帰を使用する後ろ向きコホートをを用いて確立される、請求項 31 に記載の方法。

【請求項 33】

前記第 1 及び／又は前記第 2 のランダムフォレストは、血清学的モデル、血清-遺伝子-炎症モデル、及びそれらの組み合わせからなる群から選択されるモデルに基づく、請求項 32 に記載の方法。

【請求項 34】

前記血清学的モデルは、前記後ろ向きコホートで決定された 1 つ又は複数の血清学的マーカーの存在又はレベルにランダムフォレストを適用することによって得られる、請求項 33 に記載の方法。

【請求項 35】

前記血清-遺伝子-炎症モデルは、前記後ろ向きコホートで決定された 1 つ若しくは複数の血清学的マーカーの存在若しくはレベル及び 1 つ若しくは複数の遺伝子マーカーの遺伝子型にランダムフォレストを適用することによって得られる、請求項 33 に記載の方法。

【請求項 36】

前記血清学的マーカーは、次に pANCA2 値を決定するための pANCA である、請求項 1 に記載の方法。

【請求項 37】

pANCA 値が 0 又は +1 である場合、pANCA2 値が陰性である、請求項 36 に記載の方法。

【請求項 38】

pANCA 値が +2、+3 又は +4 である場合、pANCA2 値が陽性である、請求項 36 に記載の方法。

10

20

30

40

50

【請求項 39】

前記 pANCA 2 陰性サンプルにおいて、

a) pANCA 2 陰性サンプルがクローン病の直接的な予測となるかどうかを決定するステップであって、該決定は、血清マーカーパネルを測定し、血清マーカーパネルの血清マーカーの各々をカットオフ値と比較して、サンプルがクローン病に一致するかどうかを決定することによって行われ、血清マーカーパネルは、ASCA-IgA、ASCA-IgG 及び OmpC-IgA からなる群のメンバーである、ステップ、あるいは

b) pANCA 2 陰性サンプルがクローン病に一致するかどうかを決定するステップであって、該決定は、CD 計数パネルを測定して CD 計数値を形成することによって行われ、CD 計数パネルは、ASCA-IgA、ASCA-IgG、OmpC-IgA、CBir1-IgG、A4-Fla2-IgG 及び FlaX-IgG からなる群のメンバーであり、CD 計数値が 2 以上である場合、サンプルはクローン病に一致する、ステップ、あるいは

c) pANCA 値がゼロである pANCA 2 陰性サンプルが IBD に一致するかどうかを決定するステップであって、該決定は、ANCA ELISA を測定し、ANCA ELISA 値を参照値と比較し、ANCA ELISA 値に基づいて、サンプルを IBD に一致する又は一致しないものとして指定することによって行われる、ステップ、あるいは

d) pANCA 値がゼロである pANCA 2 陰性サンプルが IBD に一致するか不確定の IBD であるかを決定するステップであって、該決定は、ANCA ELISA を測定し、ANCA ELISA 値を比較し、CD 計数値を考慮して、サンプルが IBD に一致するかどうか、又はサンプルが IBD に一致し、クローン病若しくは潰瘍性大腸炎に関して不確定であるかを決定することによって行われる、ステップ、あるいは

e) pANCA 値が 1 である pANCA 2 陰性サンプルが IBD に一致するかどうかを決定するステップであって、該決定は、ANCA ELISA を測定し、ANCA ELISA 値をカットオフ値と比較して、サンプルが IBD に一致するかどうかを決定することによって行われる、ステップ

をさらに含む、請求項 37 に記載の方法。

【請求項 40】

上記パート a) において、ASCA-IgA の値が 69 EU/mL 以上であるならば、サンプルをクローン病に一致するものとして指定する、請求項 39 に記載の方法。

【請求項 41】

上記パート a) において、ASCA-IgG の値が 40 EU/mL 以上であるならば、サンプルをクローン病に一致するものとして指定する、請求項 39 に記載の方法。

【請求項 42】

上記パート a) において、OmpC-IgA の値が 60 EU/mL 以上であるならば、サンプルをクローン病に一致するものとして指定する、請求項 39 に記載の方法。

【請求項 43】

上記パート b) において、下記 i ~ v の決定を行うことによって CD 計数値を割り当てることをさらに含む、請求項 39 に記載の方法。

i. ASCA-IgA が 8.5 EU/mL 以上であるならば、CD 計数値に +1 を与える。

ii. ASCA-IgG が 17.8 EU/mL 以上であるならば、CD 計数値に +1 を与える。

iii. OmpC-IgA が 10.9 EU/mL 以上であるならば、CD 計数値に +1 を与える。

iv. 2 種以上のフラジェリンマーカーがそれぞれの参照範囲より高ければ、CD 計数値に +1 を与える。

v. 任意のフラジェリンが 100 EU/mL 以上であるならば、CD 計数値に +1 を与え、総 CD 計数値が 2 以上であるならば、サンプルがクローン病に一致すると指定する。

【請求項 44】

C B i r 1 - I g G の参照値が 7 8 . 4 E U / m L 以上である、請求項 4 3 に記載の方法。

【請求項 4 5】

A 4 - F l a 2 - I g G の参照値が 4 4 . 8 E U / m L 以上である、請求項 4 3 に記載の方法。

【請求項 4 6】

F l a X - I g G の参照値が 3 3 . 4 E U / m L 以上である、請求項 4 3 に記載の方法。

【請求項 4 7】

上記パート c において、p A N C A がゼロ（検出されず）である場合、

i . A N C A E L I S A が 2 0 E U / m L 未満であるならば、サンプルを I B D に一致しないものとして指定するステップ、又は

i i . A N C A E L I S A が 2 7 . 4 E U / m L より大きければ、サンプルを潰瘍性大腸炎に一致するものとして指定するステップ、又は

i i i . A N C A E L I S A が 2 0 E U / m L 以上かつ 2 7 . 4 E U / m L 以下であり、C D 計数値がゼロであるならば、サンプルを I B D に一致しないものとして指定するステップ、又は

i v . A N C A E L I S A が 2 0 E U / m L 以上かつ 2 7 . 4 E U / m L 以下であり、C D 計数値が 1 であるならば、サンプルを、I B D に一致するが、クローン病又は潰瘍性大腸炎に関して不確定であるものとして指定するステップ
をさらに含む、請求項 3 9 に記載の方法。

【請求項 4 8】

p A N C A が 1 である場合、

i . A N C A E L I S A が 1 3 . 7 E U / m L 未満であるならば、サンプルを I B D に一致しないものとして指定するステップ、又は

i i . A N C A E L I S A が 1 3 . 7 E U / m L 以上であるならば、サンプルが潰瘍性大腸炎に一致するものとして指定するステップ
をさらに含む、請求項 3 9 に記載の方法。

【請求項 4 9】

個体における炎症性腸疾患（I B D）及び／又はその臨床亜型を診断するための方法であって、

（a）前記個体から得たサンプルを分析して、前記サンプル中の血清学的マーカー、遺伝子マーカー、炎症マーカー、及びそれらの組み合わせからなる群から選択される 1 つ又は複数のマーカーの存在、レベル又は遺伝子型を決定して、マーカープロファイルを得るステップと、

（b）第 1 のランダムフォレスト統計解析を前記マーカープロファイルに適用して、前記マーカープロファイルが I B D サンプルであるか非 I B D サンプルであるかの決定を得るステップと、

（c）I B D サンプルとして指定された前記サンプルに決定木又はルールセットを適用して、前記 I B D サンプルが不確定サンプルであるかどうかを決定するステップと、

（d）サンプルが不確定でなく、I B D サンプルであるならば、第 2 のランダムフォレスト統計解析を前記 I B D サンプルに適用して、臨床亜型を決定するステップと、
を含む方法。

【請求項 5 0】

第 1 のランダムフォレストアルゴリズムのためのマーカーが、A N C A、A S C A - A、A S C A - G、p A N C A、p A N C A 2、抗 F l a X 抗体、S A A、抗 F l a 2 抗体、I C A M、抗 O m p C 抗体、抗 C B i r 1 抗体、V C A M、C R P、N K X 2 - 3、A T G 1 6 L 1、S T A T 3、E C M 1、V E G F、及びそれらの組み合わせからなる群から選択される、請求項 4 9 に記載の方法。

【請求項 5 1】

第1のランダムフォレストアルゴリズムの後にサンプルがIBDサンプルとして指定され、決定木又はルールセットを使用して、サンプルについて、ANCAレベル $\geq Q3$ を有し、pANCA2陽性であり、抗C B i r 1抗体又は抗F l a 2抗体又は抗F l a X抗体が $\geq Q3$ であるかをさらに決定し、次いでIBDサンプルを不確定サンプルとして指定する、請求項49に記載の方法。

【請求項52】

第1のランダムフォレストアルゴリズムの後にサンプルがIBDサンプルとして指定され、決定木又はルールセットを使用して、サンプルについて、pANCA2陽性であり、抗C B i r 1抗体、抗F l a 2抗体及び抗F l a X抗体から選択される3種のマーカーのうちの2種が $\geq Q3$ であるかをさらに決定し、次いでIBDサンプルを不確定サンプルとして指定する、請求項49に記載の方法。

10

【請求項53】

サンプルは不確定サンプルであると決定される、請求項51又は52に記載の方法。

【請求項54】

サンプルに関して不確定サンプルが除外される、請求項51又は52に記載の方法。

【請求項55】

第1のランダムフォレストアルゴリズムの後にサンプルがIBDサンプルとして指定され、不確定が除外され、第2のランダムフォレストアルゴリズムを使用して、IBDサンプルがクローン病(CD)であるか又は潰瘍性大腸炎(UC)であるかを決定するステップをさらに含む、請求項49に記載の方法。

20

【請求項56】

第2のランダムフォレストアルゴリズムのためのマーカーが、ANCA、ASCA-A、ASCA-G、pANCA、pANCA2、抗F l a X抗体、SAA、抗F l a 2抗体、ICAM、抗OmpC抗体、抗C B i r 1抗体、VCAM、CRP、NKX2-3、ATG16L1、STAT3、ECM1、VEGF、及びそれらの組み合わせからなる群から選択される、請求項49に記載の方法。

【請求項57】

決定がCDである、請求項56に記載の方法。

【請求項58】

決定がUCである、請求項56に記載の方法。

30

【請求項59】

前記サンプルは、血清、血漿、全血及び便からなる群から選択される、請求項49に記載の方法。

【請求項60】

前記マーカープロファイルは、前記1つ又は複数のマーカーの各々に対する四分位数スコアを合計することによって得られる、前記個体に対する合計四分位数スコア(QSS)を含む、請求項49に記載の方法。

【請求項61】

前記第1及び／又は前記第2のランダムフォレストは、IBD及び／又はIBDの臨床亜型の既知の転帰を使用する後ろ向きコホートをを用いて確立される、請求項49に記載の方法。

40

【請求項62】

前記第1及び／又は前記第2のランダムフォレストは、血清学的モデル、血清-遺伝子-炎症モデル、及びそれらの組み合わせからなる群から選択されるモデルに基づく、請求項49に記載の方法。

【請求項63】

前記血清学的モデルは、前記後ろ向きコホートで決定された1つ又は複数の血清学的マーカーの存在又はレベルにランダムフォレストを適用することによって得られる、請求項61に記載の方法。

【請求項64】

50

前記血清－遺伝子－炎症モデルは、前記後ろ向きコホートで決定された1つ若しくは複数の血清学的マーカーの存在若しくはレベル及び1つ若しくは複数の遺伝子マーカーの遺伝子型にランダムフォレストを適用することによって得られる、請求項61に記載の方法。

【請求項65】

前記血清学的マーカーは、次にpANCA2値を決定するためのpANCAである、請求項49に記載の方法。

【請求項66】

pANCA値が0又は+1である場合、pANCA2値が陰性である、請求項65に記載の方法。

【請求項67】

pANCA値が+2、+3又は+4である場合、pANCA2値が陽性である、請求項65に記載の方法。

【請求項68】

前記pANCA2陰性サンプルにおいて、

a) pANCA2陰性サンプルがクローン病の直接的な予測となるかどうかを決定するステップであって、該決定は、血清マーカーパネルを測定し、血清マーカーパネルの血清マーカーの各々をカットオフ値と比較して、サンプルがクローン病に一致するかどうかを決定することによって行われ、血清マーカーパネルは、ASCA-IgA、ASCA-IgG及びOmpC-IgAからなる群のメンバーである、ステップ、あるいは

b) pANCA2陰性サンプルがクローン病に一致するかどうかを決定するステップであって、該決定は、CD計数パネルを測定してCD計数値を形成することによって行われ、CD計数パネルは、ASCA-IgA、ASCA-IgG、OmpC-IgA、CBir1-IgG、A4-Fla2-IgG及びFlax-IgGからなる群のメンバーであり、CD計数値が2以上である場合、サンプルはクローン病に一致する、ステップ、あるいは

c) pANCA値がゼロであるpANCA2陰性サンプルがIBDに一致するかどうかを決定するステップであって、該決定は、ANCA ELISAを測定し、ANCA ELISA値を参照値と比較し、ANCA ELISA値に基づいて、サンプルをIBDに一致する又は一致しないものとして指定することによって行われる、ステップ、あるいは

d) pANCA値がゼロであるpANCA2陰性サンプルがIBDに一致するか不確定のIBDであるかを決定するステップであって、該決定は、ANCA ELISAを測定し、ANCA ELISA値を比較し、CD計数値を考慮して、サンプルがIBDに一致するかどうか、又はサンプルがIBDに一致し、クローン病若しくは潰瘍性大腸炎に関して不確定であるかを決定することによって行われる、ステップ、あるいは

e) pANCA値が1であるpANCA2陰性サンプルがIBDに一致するかどうかを決定するステップであって、該決定は、ANCA ELISAを測定し、ANCA ELISA値をカットオフ値と比較して、サンプルがIBDに一致するかどうかを決定することによって行われる、ステップ

をさらに含む、請求項66に記載の方法。

【請求項69】

上記パートaにおいて、ASCA-IgAの値が69EU/mL以上であるならば、サンプルをクローン病に一致するものとして指定する、請求項68に記載の方法。

【請求項70】

上記パートaにおいて、ASCA-IgGの値が40EU/mL以上であるならば、サンプルをクローン病に一致するものとして指定する、請求項68に記載の方法。

【請求項71】

上記パートaにおいて、OmpC-IgAの値が60EU/mL以上であるならば、サンプルをクローン病に一致するものとして指定する、請求項68に記載の方法。

【請求項72】

10

20

30

40

50

上記パート b において、下記 i ~ v の決定を行うことによって C D 計数値を割り当てることをさらに含む、請求項 6 8 に記載の方法。

i . A S C A - I g A が 8 . 5 E U / m L 以上であるならば、C D 計数値に + 1 を与える。

i i . A S C A - I g G が 1 7 . 8 E U / m L 以上であるならば、C D 計数値に + 1 を与える。

i i i . O m p C - I g A が 1 0 . 9 E U / m L 以上であるならば、C D 計数値に + 1 を与える。

i v . 2 種以上のフラジェリンマーカーがそれぞれの参照範囲より高ければ、C D 計数値に + 1 を与える。

v . 任意のフラジェリンが 1 0 0 E U / m L 以上であるならば、C D 計数値に + 1 を与え、総 C D 計数値が 2 以上であるならば、サンプルがクローン病に一致すると指定する。

【請求項 7 3】

C B i r 1 - I g G の参照値が 7 8 . 4 E U / m L 以上である、請求項 7 2 に記載の方法。

【請求項 7 4】

A 4 - F l a 2 - I g G の参照値が 4 4 . 8 E U / m L 以上である、請求項 7 2 に記載の方法。

【請求項 7 5】

F l a X - I g G の参照値が 3 3 . 4 E U / m L 以上である、請求項 7 2 に記載の方法。

【請求項 7 6】

上記パート c において、p A N C A がゼロ（検出されず）である場合、

i . A N C A E L I S A が 2 0 E U / m L 未満であるならば、サンプルを I B D に一致しないものとして指定するステップ、又は

i i . A N C A E L I S A が 2 7 . 4 E U / m L より大きければ、サンプルを潰瘍性大腸炎に一致するものとして指定するステップ、又は

i i i . A N C A E L I S A が 2 0 E U / m L 以上かつ 2 7 . 4 E U / m L 以下であり、C D 計数値がゼロであるならば、サンプルを I B D に一致しないものとして指定するステップ、又は

i v . A N C A E L I S A が 2 0 E U / m L 以上かつ 2 7 . 4 E U / m L 以下であり、C D 計数値が 1 であるならば、サンプルを、I B D に一致するが、クローン病又は潰瘍性大腸炎に関して不確定であるものとして指定するステップをさらに含む、請求項 6 8 に記載の方法。

【請求項 7 7】

p A N C A が 1 である場合、

i . A N C A E L I S A が 1 3 . 7 E U / m L 未満であるならば、サンプルを I B D に一致しないものとして指定するステップ、又は

i i . A N C A E L I S A が 1 3 . 7 E U / m L 以上であるならば、サンプルが潰瘍性大腸炎に一致するものとして指定するステップをさらに含む、請求項 6 8 に記載の方法。

【請求項 7 8】

個体における炎症性腸疾患（I B D）及び／又はその臨床亜型を診断するための方法であって、

（a）前記個体から得たサンプルを分析して、前記サンプル中の血清学的マーカー、遺伝子マーカー、炎症マーカー、及びそれらの組み合わせからなる群から選択される 1 つ又は複数のマーカーの存在、レベル又は遺伝子型を決定して、マーカープロファイルを得るステップと、

（b）回帰分析を前記マーカープロファイルに適用して、前記マーカープロファイルが I B D サンプルであるか非 I B D サンプルであるかの決定を得るステップと、

10

20

30

40

50

(c) IBDサンプルとして指定された前記サンプルに決定木又はルールセットを適用して、前記 IBDサンプルが不確定サンプルであるかどうかを決定するステップと、

(d) サンプルが不確定でなく、IBDサンプルであるならば、第2の回帰分析を前記 IBDサンプルに適用して、臨床亜型を決定するステップと、を含む方法。

【発明の詳細な説明】

【関連出願の相互参照】

【0001】

[0001]本出願は、米国特許仮出願第 61/550293 号（出願日：2011 年 10 月 21 日）、米国特許仮出願第 61/553853 号（出願日：2011 年 10 月 31 日）、米国特許仮出願第 61/567096 号（出願日：2011 年 12 月 5 日）、及び米国特許仮出願第 61/570271 号（出願日：2011 年 12 月 13 日）の優先権を主張するものである。これらの仮出願の開示内容は、参照によりすべての目的のためにそのまま本明細書に組み入れられる。

【発明の背景】

【0002】

[0002]世界的に発生し、何百万人もの人を患わせている炎症性腸疾患（IBD）は、病因が明らかでない 3 つの胃腸管系障害、すなわちクローン病（CD）、潰瘍性大腸炎（UC）、及び分類不能大腸炎（IC）を表すのに用いられる総称である。IBD は、過敏性腸症候群（IBS）と共に、全米国人の半数がその一生のうちに患うことになり、IBD では 26 億ドルを超え、IBS では 80 億ドルを超える費用となっている。これらの高い医療費の主な決定要因は、消化器疾患及びこれらの疾患がどのように進行するかの診断が困難であることである。これらの障害を患っている人々が全国平均よりも年間少なくとも 8 日間追加で仕事を休んでおり、IBD 及び IBS の費用は生産性の損失によって悪化している。

【0003】

[0003]炎症性腸疾患は、腹痛、慢性下痢、体重減少及び痙攣を含む多くの症状が過敏性腸症候群と共通しており、決定的な診断が非常に困難となっている。米国において IBD を患っていると疑われる 500 万人のうち、100 万人のみが IBD に罹患しているとして診断されている。IBD の鑑別診断及びその結果の判定が困難なことで、これらの疾患の早期かつ有効な処置が妨げられる。したがって、IBD 及びその亜型を決定するための迅速かつ高感度の試験方法の必要性が存在する。

【0004】

[0004]長年にわたって、IBD の臨床亜型の正確な診断はほとんど進展していない。したがって、IBD を有する個体を診断するため及びクローン病（CD）又は潰瘍性大腸炎などの亜型を決定するための改善された方法が緊急に必要である。CD 患者の 70% が GI の外科手術を最終的に必要とするので、将来的に不必要な手術を減らすために、GI の外科手術を最終的に必要とする患者を病態初期に診断する能力が重要である。本発明はこれらの必要性を満たし、関連する利点も提供する。

【発明の概要】

【0005】

[0005]本発明は、炎症性腸疾患（IBD）を診断するため、及び／又は潰瘍性大腸炎（UC）及びクローン病（CD）を含む亜型を決定する又はサンプルを不確定（例えば、CD 及び UC に関して不確定な IBD サンプル）としてカテゴリー化するための方法及びシステムを提供する。有利なことに、本発明を使用すれば、IBD を有する患者を診断し、患者が IBD を有する場合は、UC 又は CD のいずれかを診断することが可能となる。

【0006】

[0006]いくつかの態様において、本発明は、一方は IBD 対非 IBD 用であり、他方は UC 対 CD 用である 2 つの別々のランダムフォレストを含む診断モデルを提供する。他の実施形態では、本発明は、4 種の転帰（非 IBD、CD、UC、不確定）を含むモデルを

10

20

30

40

50

提供する。さらに別の実施形態では、本発明は、3種の転帰（非IBD、CD、UC）を含む単一のモデルを提供する。

【0007】

[0007]特定の実施形態において、本発明は、サンプルを分析して、ATG16L1（例えば、rs2241880、rs3828309）、ECM1（例えば、rs7511649、rs3737240、rs13294）、NKX2-3（例えば、rs1190140、rs10883365、rs6584283）及び／又はSTAT3（例えばrs744166）遺伝子の1つ、2つ、3つ、4つ又はそれより多くのバリエーション対立遺伝子（例えば、一塩基多型又はSNP）の有無を決定することによって、IBD又は非IBDを診断する、及び／又はUC及びCDなどのIBDの臨床亜型を鑑別する、又はサンプルを不確定としてカテゴリー化するための方法及びシステムを提供する。これらの実施形態のいくつかの態様では、本発明は、IBD又は非IBDを診断する、及び／又はIC、UC及びCDなどのIBD臨床亜型を鑑別する、又はサンプルを不確定としてカテゴリー化するために、サンプルを分析して、ASCA-A、ASCA-G、抗OmpC、抗Cbir1、抗Fla2、抗FlaX、VEGF、CRP、SAA、ICAM、VCAM、ANCAなどの1つ又は複数の血清学的マーカー及び炎症マーカー（例えばELISAによる）、及び／又はpANCA（例えば、IBD特異的なpANCAの核周囲パターン及び／又はpANCAのDNase感受性を検出するための1つ又は複数の間接蛍光抗体（IFA）アッセイによる）、pANCA2の存在（若しくは非存在）又は濃度レベルを決定することをさらに含むことができる。

【0008】

[0008]一実施形態において、本発明は、個体における炎症性腸疾患（IBD）及び／又はその臨床亜型を診断するための方法であって、臨床亜型が潰瘍性大腸炎（UC）又はクローン病（CD）又は不確定であり、

（a）個体から得たサンプルを分析して、サンプル中の血清学的マーカー、遺伝子マーカー、炎症マーカー、及びそれらの組み合わせからなる群から選択される1つ又は複数のマーカーの存在、レベル又は遺伝子型を決定して、マーカープロファイルを得るステップと、

（b）第1のランダムフォレスト統計解析をマーカープロファイルに適用して、マーカープロファイルがIBDサンプルであるか非IBDサンプルであるかの決定を得るステップと、

（c）任意選択で、IBDサンプルとして指定されたサンプルに決定木又はルールセットを適用して、IBDサンプルが不確定サンプルであるかどうかを決定するステップと、

（d）任意選択で、第2のランダムフォレスト統計解析をIBDサンプルに適用して、臨床亜型を決定するステップと、を含む方法を提供する。

【0009】

[0009]いくつかの実施形態では、IBDの診断アッセイは、個体の標本を、そのpANCA2値に応じて決定木法又はランダムフォレストアルゴリズムのいずれかに導くための第1ステップとして、サンプルがpANCA2陰性であるかどうかを決定するための初期ノードを含む。いくつかの例では、IBDの臨床的予測を行うために、陰性（0又は「検出されず」）又は弱い（+1）のpANCA決定のいずれかを有するpANCA2サンプルは取り出され、決定行列又はルールセットにかけられる。

【0010】

[0010]他の実施形態において、本発明は、個体における炎症性腸疾患（IBD）及び／又はその臨床亜型を診断するための方法であって、臨床亜型が潰瘍性大腸炎（UC）又はクローン病（CD）又は不確定であり、

（a）個体から得たサンプルを分析して、サンプル中の血清学的マーカー、遺伝子マーカー、炎症マーカー、及びそれらの組み合わせからなる群から選択される1つ又は複数のマーカーの存在、レベル又は遺伝子型を決定して、マーカープロファイルを得るステップ

と、

(b) 第1のランダムフォレスト統計解析をマーカープロファイルに適用して、マーカープロファイルがIBDサンプルであるか非IBDサンプルであるかの決定を得るステップと、

(c) IBDサンプルとして指定されたサンプルに決定木又はルールセットを適用して、IBDサンプルが不確定サンプルであるかどうかを決定するステップと、

(d) サンプルが不確定でなく、IBDサンプルであるならば、第2のランダムフォレスト統計解析をIBDサンプルに適用して、臨床垂型を決定するステップと、を含む方法を提供する。

【0011】

10

[0011]これら及び他の態様、目的及び実施形態は以下の詳細な説明及び図面からさらに明らかとなる。

【図面の簡単な説明】

【0012】

【図1】図1は、二段階のランダムフォレストアルゴリズムを含む本発明のIBD診断アッセイの例示的实施形態に関するフローチャートである。

【図2】図2は、二段階のランダムフォレストアルゴリズム（右側の経路）及び別々のpANCA決定木（左側の経路）を含む本発明のIBD診断アッセイの他の例示的实施形態に関するフローチャートである。

【図3】図3は、本発明の疾患分類システム（DCS）の例示的实施形態を示す。

20

【図4】図4は、固定した好中球に対する免疫蛍光、それに続くDNAse処理によるpANCA染色パターンを示す。

【図5】図5は、訓練コホート及び検証コホートの特徴並びにsgfアルゴリズム及び個々のマーカーに関するROC曲線下面積を示す。

【図6】図6は、IBD診断を予測するためのロジスティック回帰モデルの例示的实施形態を示す。

【図7】図7Aは、IBDモデルはIBDを予測するための優れた分類子システムであることを示す。図7Bは、UCモデルはUCを予測するための優れた分類子システムであることを示す。

【図8】図8Aは、IBDロジットモデル対Rで作出されたIBDランダムフォレストモデルに関する予測P値の散布図である。図8Bは、Matlabで作出されたIBDランダムフォレストモデル対IBDロジットモデルの散布図である。

30

【図9】図9は、IBDロジットモデル及び2種のIBDランダムフォレストモデルに関する予測P値の密度グラフである。グラフは、非IBD、CD及びUCの予測におけるモデルの類似性を比較する。

【図10】図10Aは、UCロジットモデル及びRで作出されたUCランダムフォレストモデルに関する予測P値の散布図である。図10Bは、UCロジットモデル及びMatlabで作出されたUCランダムフォレストモデルに関する予測P値の散布図である。

【図11】図11は、UCロジットモデル及び2種のUCランダムフォレストモデルに関する予測P値の密度グラフである。グラフは、非IBD、CD及びUCの予測におけるモデルの類似性を比較する。

40

【図12】図12は、IBD対非IBDを予測するためのロジットモデリング対ランダムフォレストモデリングの散布図である。

【図13】図13は、UC対CDを予測するためのロジットモデリング対ランダムフォレストモデリングの散布図である。

【図14】図14は、R又はMatlabのいずれかで作出された2種のランダムフォレストIBD対非IBDモデルの散布図である。

【図15】図15は、R又はMatlabのいずれかで作出された2種のランダムフォレストUC対CDモデルの散布図である。

【図16】図16は、IBDsgfコホートの血清学的マーカーを示す。

50

【図 17】図 17 は、IBD s g i コホートの炎症マーカーを示す。

【図 18】図 18 は、非 IBD 患者よりも多くの IBD 患者が 2 つ以上の遺伝子マーカーに関して陽性である一実施形態を示す。

【図 19】図 19 は、UC からの CD の識別に関する、IBD s g i 及び IBD セロロジー 7 の比較の一実施形態を示す。

【図 20】図 20 は、IBD 対非 IBD 並びに CD 対 UC に関する、IBD s g i 及び IBD セロロジー 7 の全体的な診断能力の比較を示す。

【発明の詳細な説明】

【0013】

I. イントロダクション

本発明は、炎症性腸疾患（IBD）及びその亜型の診断の精度が、個体由来の生体サンプルにおける特定のマーカーの存在、レベル又は遺伝子型を検出することによって実質的に改善することができるという驚くべき発見に一部基づく。すなわち、一実施形態において、本発明は、血清学的及び／又は遺伝的なマーカーパネルに基づく診断プラットフォームを提供する。

【0014】

[0033] 本発明は、IBD 及びその亜群（例えば、分類不能大腸炎（IC）、潰瘍性大腸炎（UC）、クローン病（CD）、並びに CD 及び UC に関して不確定（「不確定」））の診断を改善するための方法及びシステムを提供する。一部の例では、本明細書に記載の方法は、転帰を診断及び予測することが非常に難しいと知られている疾患である IBD を正確に予測する。他の例では、本明細書に記載の方法は、多数の血清学的、タンパク質及び／又は遺伝子マーカーを、単独で又は 1 つ若しくは複数のアルゴリズム若しくは他の種類の統計解析と組み合わせて利用して、価値のある診断又は予後診断の見識を医師に提供する。一部の例では、本発明の方法及びシステムは、生物学的治療に対して予想される患者の応答の指標を提供する。他の例では、本発明の方法及びシステムは、多数のマーカー（例えば、血清学的、タンパク質及び／又は遺伝子マーカー）を統計解析（例えば四分位数分析）と併せて利用して、合併疾患を有する患者若しくは疾患合併症（例えば、内部狭窄性又は内部穿通性疾患）を発症するリスクを有する患者及び／又は外科的介入の必要がある患者を特定することによって予後診断値を提供する一方で、疾患進行速度評価の補助も行う。いくつかの例では、本方法は、IC、UC 及び CD を含む IBD 亜型の連続的なつながりに沿った疾患重症度の分類を可能にする。さらに、本方法は、進行性疾患を有する患者の治療決定を導く。他の例では、多数のマーカー（例えば、血清学的、タンパク質及び／又は遺伝子マーカー）の使用は、非応答者から応答者を識別する能力及び初期の治療の選択肢（例えば、積極的処置を処方するかどうか）を導く能力を提供し、疾患挙動を変化させる可能性を有する。

【0015】

[0034] 図 1 は、二段階のランダムフォレストアルゴリズムを含む本発明の IBD 診断アッセイの例示的实施形態のフローチャートを示す。いくつかの実施形態では、IBD 診断アッセイは、17 種の血清-遺伝子-炎症（s g i）生物学的マーカー（例えば、ANCA、ASCA-A、ASCA-G、抗FlaX、抗A4-Fla2、pANCA、抗OmpC、抗CBir1、ATG16L1、CRP、SAA、ICAM、VCAM、ECM1、STAT3、VEGF、NKX2-3、及びそれらの組み合わせ）からの測定値を適用し、第 1 のランダムフォレストモデルに基づいてスコアを算出して、IBD 対非 IBD を予測する（110）。第 1 のランダムフォレストモデルは、患者が IBD を有するかどうか決定する。スコアが IBD 対非 IBD のカットオフ未満（例えば < 0.64）であるならば、サンプルは、IBD を有する患者由来であると予測される（125）。そうでなければ、サンプルは、非 IBD を有する患者由来と予測される（120）。IBD を有すると予測されるサンプルは、アルゴリズムの次のステップへ進み、このステップは、サンプルを不確定としてカテゴリー化するのを除外するように設計されている決定木又はルールセットである（130）。サンプルが「分類不能」ルールの一つに関与するパターンと

10

20

30

40

50

一致する場合は、アルゴリズムは、IBDを有するものとしてサンプルを予測するが、UC及びCDに関しては不確定である（135）。そうでなければ、サンプルは、アルゴリズムの次のステップへ進み（140）、このステップは、UC対CDを予測するための第2のランダムフォレストモデルである（150）。本IBD診断アッセイは、11種のs g i生物学的マーカー（例えば、ANCA、ASCA-A、ASCA-G、抗FlaX、抗A4-Fla2、pANCA、抗OmpC、抗Cbir1、ECM1、STAT3、VEGF、及びそれらの組み合わせ）由来の測定値を適用して、UC対CDを予測するための第2のランダムフォレストモデルに基づいてモデルスコアを算出する。スコアがUC対CDカットオフ（例えば0.35）未満である場合、アルゴリズムは、CDを有するものとしてサンプルを予測する（153）。そうでなければ、アルゴリズムは、UCを有するものとしてサンプルを予測する（155）。

10

【0016】

[0035]図2は、二段階のランダムフォレストアルゴリズム（右側の経路）及び別々のpANCA決定木（左側の経路）を含む本発明のIBD診断アッセイの他の例示的实施形態のフローチャートを示す。いくつかの実施形態では、IBD診断アッセイは、個体の標本をそのpANCA2値に応じて決定木法（205）又はランダムフォレストアルゴリズム（210）のいずれかに導くための第1ステップとして、サンプルがpANCA2陰性であるかどうかを決定するための初期ノード（201）を含む。いくつかの例では、陰性（0又は「検出されず」）又は弱い（+1）pANCA決定のいずれかを有するpANCA2サンプルは、s g iアルゴリズムの待ち行列（右側の経路）から取り出され、IBDの臨床的予測を行うための決定行列又はルールセットにかけられる（207）。pANCA決定木分析に従ってIBDを予測するための決定行列の例は実施例9に記載されている。すなわち、いくつかの態様では、pANCA決定木（205）は、アッセイ結果がpANCA2（-）である臨床標本において予測的IBD診断を行うための代替経路として実施することができる。

20

【0017】

[0036]図2はまた、pANCA2陽性サンプルを分析するための経路の例を示す（右側の経路）。いくつかの実施形態では、pANCA決定が陽性（+2、+3又は+4）であるpANCA2サンプルは、IBD診断アッセイを使用して処理され、図1に記載の二段階のランダムフォレストアルゴリズムに類似した二段階のランダムフォレストアルゴリズムにかけられる。いくつかの例では、IBD診断アッセイは、17種の血清-遺伝子-炎症（s g i）生物学的マーカー（例えば、ANCA、ASCA-A、ASCA-G、抗FlaX、抗A4-Fla2、pANCA、抗OmpC、抗Cbir1、ATG16L1、CRP、SAA、ICAM、VCAM、ECM1、STAT3、VEGF、NKX2-3、及びそれらの組み合わせ）由来の測定値を適用し、第1のランダムフォレストモデルに基づいてスコアを算出して、IBD対非IBDを予測する（210）。第1のランダムフォレストモデルは、患者がIBDを有するかどうかを決定する。スコアがIBD対非IBDのカットオフ未満（例えば<0.64）であるならば、サンプルはIBDを有する患者由来であると予測される（225）。そうでなければ、サンプルは、非IBDを有する患者由来と予測される（220）。IBDを有するものとして予測されるサンプルはアルゴリズムの次のステップへ進み、このステップは、サンプルを不確定としてカテゴリー化するのを除外するように設計されている決定木又はルールセットである（230）。サンプルが「分類不能」ルールのいずれかに対するパターンに一致するならば、アルゴリズムは、IBDを有するものとしてサンプルを予測するが、UC及びCDに関しては不確定である（235）。そうでなければ、サンプルは、アルゴリズムの次のステップ（240）へ進み、このステップは、UC対CDを予測するための第2のランダムフォレストモデルである（250）。IBD診断アッセイは、11種のs g i生物学的マーカー（例えば、ANCA、ASCA-A、ASCA-G、抗FlaX、抗A4-Fla2、pANCA、抗OmpC、抗Cbir1、ECM1、STAT3、VEGF、及びそれらの組み合わせ）由来の測定値を適用して、UC対CDを予測するための第2のランダムフォレストモデル

30

40

50

に基づいてモデルスコアを算出する。スコアがUC対CDカットオフ（例えば0.35）未満であるならば、アルゴリズムは、CDを有するものとしてサンプルを予測する（253）。そうでなければ、アルゴリズムは、UCを有するものとしてサンプルを予測する（255）。

【0018】

[0037]いくつかの例では、本発明の方法及びシステムは、「変換」又はそれと関連する「機器」を含む。例えば、ELISA法を行って、本明細書に記載のマーカーのうちの多くの存在又は濃度レベルを測定してもよい。ELISAには、マーカー（例えば自己抗体）の、マーカー（例えば自己抗体）と結合剤（例えば抗原）との複合体への変換が含まれ、その後、これは標識した二次抗体で測定することができる。多くの場合、標識は、基質を検出可能な生成物へと変換する酵素である。検出可能な生成物の測定は、分光光度計などのプレートリーダーを用いて行うことができる。他の例では、遺伝子マーカーは、PCRなどの種々の増幅法を用いて決定される。PCRなどの増幅が含まれる方法ステップは、検出のために核酸の一本鎖又は二本鎖から多重鎖への変換をもたらす。検出には、蛍光光度計などの機器を用いて行われる蛍光体の使用を含めることができる。

【0019】

[0038]いくつかの実施形態では、本方法は、IBD診断を援助するために、統計解析の結果（すなわち診断プロファイル）を参照（すなわち診断モデル）と比較することをさらに含む。特定の実施形態では、本方法は、医師に価値のある診断の見識を提供するために、多数の血清学的、タンパク質及び／又は遺伝子マーカーを利用する。

【0020】

[0039]有利なことに、本発明のアッセイ方法及びシステムは、多数のマーカー（例えば、血清学的、遺伝子、炎症性マーカー）からなる診断プロファイルを単独で又は統計解析と併せて使用して、IBD及びIC、UC、CD及び不確定などのIBD亜型を有する患者を特定することによって診断値を提供する。いくつかの例では、本発明は、これらに限定されないが、IC、UC、CD及び不確定を含むIBD亜型の連続的なつながりに沿った疾患重症度の分類を可能にする。

【0021】

II. 定義

本明細書において、以下の用語は、特に断らない限り、それらの用語に割り当てられた意味を有する。

【0022】

[0041]用語「分類すること」には、サンプル又は病状若しくは予後診断を有する個体を「関連づけること」又は「カテゴリー化すること」が含まれる。いくつかの例では、「分類すること」は、統計的証拠、経験的証拠、又は両方に基づいている。いくつかの実施形態では、分類する方法及びシステムでは、既知の病状又は予後診断を有する個体からのサンプルの、いわゆる訓練セットが使用される。確立された後、訓練データセットは、個体において未知の病状进行分类する又は病状の予後診断を提供するために個体からの未知のサンプルの特長を比較する場合に、その比較対象の基礎、モデル、又は鋳型として役割を果たす。一部の例では、「分類すること」とは、病状を診断すること及び／又は病状を他の病状から識別することと同種である。他の例では、「分類すること」とは、病状が診断された個体において病状の予後診断を提供することと同種である。

【0023】

[0042]用語「炎症性腸疾患」又は「IBD」には、例えば、クローン病（CD）、潰瘍性大腸炎（UC）、分類不能大腸炎（IC）、CD対UCに関して不確定なIBD（「不確定」）などの胃腸管系障害が含まれる。炎症性腸疾患（例えば、CD、UC、IC及び不確定）は、過敏性腸症候群（IBS）を含む消化管のすべての他の障害、症候群、及び異常から区別される。表題「炎症性腸疾患の診断方法」の米国特許第7873479号は、参照によりすべての目的のためにそのまま本明細書に組み入れられる。

【0024】

[0043]用語「サンプル」には、個体から入手した任意の生体検体が含まれる。本発明における使用に適したサンプルには、これらに限定されないが、全血、血漿、血清、唾液、尿、大便、涙、他の任意の体液、組織サンプル（例えば生検）、及びその細胞抽出物（例えば赤血球抽出物）が含まれる。好ましい実施形態では、サンプルは血清サンプルである。血清、唾液、及び尿などのサンプルの使用は当分野で周知である（例えば、Hashidaら, J. Clin. Lab. Anal., 11:267~86 (1997)を参照）。当業者には、マーカーレベルを分析する前に血清サンプルなどのサンプルを希釈できることが理解されよう。

【0025】

[0044]用語「マーカー」には、IBDの診断、IBDの確からしい経過及び結果の予測、及び／又は疾患からの回復の見込みの予測に使用することができる、任意の生化学的マーカー、血清学的マーカー、遺伝子マーカー、又は他の臨床若しくは超音波検査の特徴が含まれる。そのようなマーカーの非限定的な例には、抗好中球抗体（例えば、ANCA、pANCA）、抗サッカロマイセス・セレビシエ（*Saccharomyces cerevisiae*）抗体（例えば、ASCA-IgA、ASCA-IgG）、抗微生物抗体（例えば、抗OmpC抗体、抗I2抗体、抗Fla2抗体、抗FlaX抗体、抗CBir1抗体）、急性期タンパク質（例えばCRP）、アポリポタンパク質（例えばSAA）、デフェンシン（例えばβデフェンシン）、成長因子（例えばEGF、VEGF）、サイトカイン（例えばTWEAK、IL-1β、IL-6）、カドヘリン（例えばE-カドヘリン）、細胞接着分子（例えばICAM-1、VCAM-1）などの血清学的マーカー、ATG16L1、ECM1、NKX2-3、STAT3及びNOD2/CARD15などの遺伝子マーカー（例えばSNP）、及びそれらの組み合わせが含まれる。遺伝子マーカーの非限定例としては、GLI1（例えばrs2228224及び／又はrs2228226）、MDR1（例えばrs2032582）、ATG16L1（例えばrs2241880、rs3828309）、ECM1（例えばrs7511649、rs373240、rs13294）、NKX2-3（例えばrs1190140、rs10883365、rs6584283）、STAT3（例えばrs744166）及びNOD2/CARD15（例えばrs2066847、rs2066845、rs5743293）遺伝子のバリエーション対立遺伝子が挙げられる。一部の実施形態では、マーカーを1つ又は複数の（例えば複数の）統計解析と組み合わせて利用して、個体におけるIBDの診断又は予後診断を援助又は提供する。いくつかの例では、診断は、IBD又はクローン病（CD）、潰瘍性大腸炎（UC）、分類不能大腸炎（IC）若しくは不確定IBDなどのその臨床亜型であり得る。他のいくつかの例では、予後診断は、手術の必要性（例えば、小腸手術が必要となる可能性若しくはリスク）、CD若しくはUCの臨床亜型の発生（例えば、CD若しくはUCの特定の臨床亜型（例えば、狭窄性、穿通性若しくは炎症性のCD亜型）に罹患する可能性若しくはリスク）、1つ若しくは複数の臨床学的因子の発生（例えば、特定の臨床学的因子の影響を受ける可能性若しくはリスク）、腸癌の発生（例えば、腸癌に罹患する可能性若しくはリスク）、又は疾患からの回復（例えば、寛解の可能性）であり得る。

【0026】

[0045]本発明は、部分的に、個体から得られたサンプル中の少なくとも1つのマーカーの存在（若しくは非存在）又はレベル（例えば濃度）を決定することに依存する。本明細書において、用語「少なくとも1つのマーカーの存在を検出すること」には、当業者に知られている任意の定量的又は定性的アッセイを使用することによって目的の各マーカーの存在を決定することが含まれる。いくつかの例では、特定の形質、変数、遺伝子型、及び／又は生化学的若しくは血清学的物質（例えば、タンパク質、抗体）の存在又は非存在を決定する定性的アッセイが目的の各マーカーを検出するために適している。他のいくつかの例では、DNA、RNA、タンパク質、抗体、又は活性の存在又は非存在を決定する定量的アッセイが目的の各マーカーを検出するために適している。本明細書において、用語「少なくとも1つのマーカーのレベルを検出すること」には、当業者に知られている任意

の直接又は間接的な定量的アッセイを使用することによって目的の各マーカーのレベルを決定することが含まれる。いくつかの例では、例えばDNA、RNA、タンパク質、抗体、又は活性の相対的又は絶対的な量を決定する定量的アッセイが目的の各マーカーのレベルを検出するために適している。当業者には、マーカーのレベルの検出に有用な任意のアッセイがマーカーの存在又は非存在の検出にも有用であることが理解されよう。

【0027】

[0046]用語「マーカープロファイル」には、1、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又はそれより多くの診断及び／又は予後診断のマーカーが含まれ、マーカーは、血清学的マーカー、タンパク質マーカー、遺伝子マーカーなどであり得る。一部の実施形態では、マーカープロファイルは、統計解析と共に、医師及び介護人に価値のある診断及び予後診断の見識を提供することができる。他の実施形態では、マーカープロファイルは、（任意選択で統計解析と共に、）生物学的治療に対して予想される応答を提供する。本明細書に記載のアッセイは、多数のマーカー（例えば、血清学的、タンパク質、遺伝子マーカー）を統計解析と併せて使用して、IBD又はその臨床亜型を有する患者を特定すること、合併疾患の発症のリスクを予測すること、疾患進行速度（例えば、合併疾患又は手術への進行速度）の評価を補助すること、及び治療の選択を補助することによって、診断的、予後診断的及び治療的な値を提供する。

【0028】

[0047]用語「診断プロファイル」には、1、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又はそれより多くの、個体のマーカーが含まれ、マーカーは、血清学的マーカー、タンパク質マーカー、遺伝子マーカーなどであり得る。統計解析によって、マーカープロファイルは診断プロファイルに変換される。好ましい統計解析は四分位数スコアであり、各マーカーに対する四分位数スコアを合計して、四分位数の和のスコアを生成することができる。

【0029】

[0048]用語「診断モデル」には、血清学的モデル、遺伝子モデル、血清－遺伝子－炎症モデル、及びそれらの組み合わせが含まれる。好ましい態様では、遡及的分析は、既知の合併症及び実施された外科手術を伴う既知の疾患の転帰のコホートに関して行われる。一態様では、診断モデルを開発するために、1つ若しくは複数の血清学的マーカーの存在若しくは濃度レベル及び／又は1つ若しくは複数の遺伝子マーカーの遺伝子型に関して回帰分析（例えばロジスティック回帰）を実施することができる。診断モデルは、例えば参照表、グラフ又は他の表示で示す又は表すことができる。次いで、個体の診断プロファイルを診断モデルと比較することができ、診断（例えば、疾患を有するリスク又は可能性）が決定される。

【0030】

[0049]用語「個体」、「対象」又は「患者」には、典型的にはヒトが含まれるが、例えば、他の霊長類、げっ歯類、イヌ、ネコ、ウマ、ヒツジ、ブタなどの他の動物も含まれる。

【0031】

[0050]本明細書において、用語「実質的に同一のアミノ酸配列」には、天然に存在するアミノ酸配列と類似するが同一ではないアミノ酸配列が含まれる。例えば、I2タンパク質と実質的に同一のアミノ酸配列を有するアミノ酸配列、すなわちポリペプチドは、天然に存在するI2タンパク質のアミノ酸配列に対するアミノ酸付加、欠失又は置換などの1つ又は複数の修飾を有し得る。但し、修飾されたポリペプチドが、免疫応答性などの少なくとも1つのI2の生物活性を実質的に保持するものとする。アミノ酸配列間の実質的な類似性の比較は、約6～100残基、好ましくは約10～100残基、より好ましくは約25～35残基の配列を用いて通常実施される。本発明のポリペプチド又はその断片の特に有用な修飾は、例えば安定性の増大を与える修飾である。1つ又は複数のD-アミノ酸

の導入は、ポリペプチド又はポリペプチド断片の安定性を増大させるのに有用な修飾である。同様に、リジン残基の欠失又は置換は、ポリペプチド又はポリペプチド断片を分解から保護することによって安定性を増大させることができる。

【0032】

[0051]用語「臨床学的因子」には、IBDに関連づけられている、個体における症状が含まれる。臨床学的因子の例には、これらに限定されないが、下痢、腹痛、痙攣、発熱、貧血、体重減少、不安、鬱病、及びそれらの組み合わせが含まれる。一部の実施形態では、IBDの診断又は予後診断は、個体から得られたサンプルを分析して1つ又は複数の統計解析を適用することによって1つ又は複数のマーカーの存在、レベル、又は遺伝子型を決定することと、個体が1つ又は複数の臨床学的因子を有するかどうかを決定することとの組み合わせに基づいている。

10

【0033】

[0052]用語「症状」又は「複数の症状」及びその変形には、個体によって経験される、特定の疾患に関連づけられている又は疾患に伴う、及び疾患の指標としてみなされる、身体機能の任意の感覚、変化又は認知された変化が含まれる。本発明に関して使用される症状に関連し得る疾患には、炎症性腸疾患（IBD）、潰瘍性大腸炎（UC）又はクローン病（CD）が含まれる。

【0034】

[0053]好ましい態様では、本発明の方法は、個体がIBDを診断された後に使用される。しかし、他の例では、本方法は、IBDを診断するために使用することができるか、又は、例えばIBDが疑われている又は他の方法を用いて以前に診断されている場合は、「第二診断」として使用することができる。好ましい態様では、本方法は、UCを診断するため、又はUCとCDとを鑑別するために使用することができる。用語「IBDを診断すること」及びその変形には、IBDの存在又は非存在を決定するための、本明細書に記載の方法及びシステムの使用が含まれる。用語「UCを診断すること」には、UCの存在又は非存在を決定するため、及びUCとCDとを鑑別するための、本明細書に記載の方法及びシステムの使用が含まれる。また、この用語には、個体における疾患活動性のレベルを評価することも含めることができる。一部の実施形態では、統計解析を使用して、True loveら、Br. Med. J., 12:1041~1048 (1955)によって開発された基準に基づいて、軽度、中等度、重度、又は劇症型のIBD又はUCを診断する。他の実施形態では、統計解析を使用して、Hanauerら、Am. J. Gastroenterol., 92:559~566 (1997)によって開発された基準に基づいて、軽度から中等度、中等度から重度、又は重度から劇症の型のIBD又はUCを診断する。当業者は、個体におけるIBD又はUCの重症度を評価するための他の方法を知っているであろう。

20

30

【0035】

[0054]いくつかの例では、本発明の方法は、IBDを診断する、UCを診断する、又はICをUC及び／又はCDから識別する、並びにUCとCDとを鑑別する、又はサンプルを不確定としてカテゴリー化するために使用される。本方法は、疾患の進行及び退行のいずれをモニタリングするためにも使用することができる。用語「IBD又はUCの進行又は退行をモニタリングすること」には、個体の病状（例えば、IBDの存在若しくは重症度又はUCの存在）を決定するための方法及びマーカープロファイルの使用が含まれる。いくつかの例では、統計解析の結果を、それより前の時点に同じ個体で得られた結果と比較する。一部の態様では、本発明の方法は、例えば、個体においてIBD又はUCが迅速に又はゆっくりと進行する可能性をサンプル中の少なくとも1つのマーカーの存在又はレベルに基づいて決定することによって、IBD又はUCの進行を予測するためにも使用することができる。他の態様では、本発明の方法は、例えば、個体においてIBD又はUCが迅速に又はゆっくりと退行する可能性をサンプル中の少なくとも1つのマーカーの存在又はレベルに基づいて決定することによって、IBD又はUCの退行を予測するためにも使用することができる。

40

50

【0036】

[0055]用語「遺伝子」は、ポリペプチド鎖の産生に關与しているDNAのセグメントをいい、コード領域の前及び後の領域（例えば、それぞれプロモーター及び3'非翻訳領域）、並びに個々のコードセグメント（エクソン）間の介在配列（イントロン）が含まれる。

【0037】

[0056]用語「遺伝子型」は、例えば二倍体の生物が1つ又は複数の目的のバリエーション対立遺伝子に關してヘテロ接合性であるかホモ接合性であるかを含む、生物の遺伝的組成をいう。

【0038】

[0057]用語「多型性」は、集団中の2つ以上の遺伝子決定された代替配列又は対立遺伝子の発生をいう。「多型部位」とは、分岐が起こる座位をいう。好ましい多型部位は、各々が集団中で特定の頻度で発生している少なくとも2つの対立遺伝子を有する。多型座位は1つの塩基対（すなわち一塩基多型又はSNP）ほど小さくてもよい。多型マーカーには、制限断片長多型、可変数タンデム反復（VNTR）、超可変領域、ミニサテライト、ジヌクレオチド反復、トリヌクレオチド反復、テトラヌクレオチド反復、単純配列反復、及びAluなどの挿入要素が含まれる。最初に同定された対立遺伝子が自由裁量で参照対立遺伝子と指定され、他方の対立遺伝子が代替対立遺伝子、「バリエーション対立遺伝子」、又は「バリエーション」として指定される。選択された集団中で最も頻繁に発生している対立遺伝子は「野生型」対立遺伝子と呼ばれることもある。二倍体の生物は、バリエーション対立遺伝子に關してホモ接合性又はヘテロ接合性であり得る。バリエーション対立遺伝子は、バリエーション対立遺伝子を保有する個体において観察可能な物理的又は生化学的な特徴（「表現型」）を生じて生じなくてもよい。例えば、バリエーション対立遺伝子は、目的の遺伝子によってコードされているタンパク質の酵素活性を変更してもよい。

【0039】

[0058]用語「一塩基多型（SNP）」及びその変形は、対立遺伝子内でのものを含む、単一のヌクレオチドをポリヌクレオチドで変えることをいう。これには、1つのヌクレオチドの他のヌクレオチドによる置換え、及び単一のヌクレオチドの欠失又は挿入が含まれる場合がある。最も典型的には、SNPは二対立遺伝子マーカーであるが、三及び四対立遺伝子マーカーも存在することができる。非限定的な例として、SNP A/Cを含む核酸分子には、多型位置にC又はAが含まれていてもよい。SNPの組み合わせには、用語「ハプロタイプ」、例えば、互いに連結された、単一のDNA鎖中のSNPの遺伝子型が使用される。一部の実施形態では、用語「ハプロタイプ」は、SNP対立遺伝子の組み合わせ、例えば単一のDNA分子上に一緒に見つかるSNPの対立遺伝子を説明するために使用することができる。さらなる実施形態では、ハプロタイプ中のSNPは互いに連鎖不平衡であり得る。

【0040】

[0059]用語「特異的」又は「特異性」及びその変形には、バリエーション対立遺伝子を検出可能なポリヌクレオチド（例えば、異なる対立遺伝子間を区別することができるポリヌクレオチド）に關して使用される場合、一方のバリエーション対立遺伝子を結合又はハイブリダイズ又は検出せずに、他方のバリエーション対立遺伝子を結合又はハイブリダイズ又は検出する能力が含まれる。一部の実施形態では、特異性とは、野生型対立遺伝子を検出し、ミュータント又はバリエーション対立遺伝子を検出しないポリヌクレオチドの能力をいうことができる。他の実施形態では、特異性とは、ミュータント又はバリエーション対立遺伝子を検出し、野生型対立遺伝子を検出しないポリヌクレオチドの能力をいうことができる。

【0041】

[0060]本明細書において、用語「抗体」には、ポリクローナル若しくはモノクローナルかつ任意のアイソタイプのものであり得る免疫グロブリン分子の集団、又は免疫グロブリン分子の免疫学的に活性のある断片が含まれる。そのような免疫学的に活性のある断片は、抗原と特異的に結合する抗体分子の部分を構成する、重鎖及び軽鎖可変領域を含有する

。例えば、当分野で $F a b$ 、 $F a b'$ 又は $F(a b')_2$ として知られている免疫グロブリン分子の免疫学的に活性のある断片は用語「抗体」に含まれる。

【0042】

[0061] 四分位数分析では、データ範囲を4等分する3つの数(値)が存在する。第1四分位数(「下位四分位数」とも呼ばれる)は、それより下に下位データの25パーセントがある数である。第2四分位数(「中央値」)は、データ範囲を中央で分割し、それより下にデータの50パーセントを有する。第3四分位数(「上位四分位数」とも呼ばれる)は、それより下にデータの75パーセントを有し、それより上にデータの上位25パーセントを有する。非限定例としては、四分位数分析は本明細書に記載の抗体又は他のタンパク質マーカーなどのマーカーの濃度レベルに適用することができ、その結果、第1四分位数(<25%)内のマーカーレベルは1の値が割り当てられ、第2四分位数(25~50%)のマーカーレベルは2の値が割り当てられ、第3四分位数(51%~<75%)のマーカーレベルは3の値が割り当てられ、第4四分位数(75%~100%)のマーカーレベルは4の値が割り当てられる。

10

【0043】

[0062] 本明細書において、「四分位数の和のスコア」又は「QSS」には、すべての目的のマーカーに対する四分位数スコアの和が含まれる。非限定例としては、6マーカー(例えば、血清学的、タンパク質及び／又は遺伝子マーカー)のパネルに対する四分位数の和のスコアは6~24に及んでもよく、個々のマーカーの各々は、マーカーの有無、マーカーの濃度レベル、又はマーカーの遺伝子型に基づいて1~4の四分位数スコア(Q)が割り当てられる。

20

【0044】

III. 実施形態の説明

一実施形態において、本発明は、個体における炎症性腸疾患(IBD)及び／又はその臨床亜型を診断するための方法であって、

(a) 個体から得たサンプルを分析して、サンプル中の血清学的マーカー、遺伝子マーカー、炎症マーカー、及びそれらの組み合わせからなる群から選択される1つ又は複数のマーカーの存在、レベル又は遺伝子型を決定して、マーカープロファイルを得るステップと、

(b) 第1のアルゴリズム(例えば、ランダムフォレスト又はロジスティック回帰)をマーカープロファイルに適用して、マーカープロファイルがIBDサンプルであるか非IBDサンプルであるかの決定を得るステップと、

30

(c) 任意選択で、IBDサンプルとして指定されたサンプルに決定木又はルールセットを適用して、IBDサンプルが不確定サンプルであるかどうかを決定するステップと、

(d) 任意選択で、第2のアルゴリズム(例えば、ランダムフォレスト又はロジスティック回帰)をIBDサンプルに適用して、臨床亜型を決定するステップと、を含む方法を提供する。

【0045】

[0064] 第2の実施形態において、本発明は、個体における炎症性腸疾患(IBD)及び／又はその臨床亜型を診断するための方法であって、

40

(a) 個体から得たサンプルを分析して、サンプル中の血清学的マーカー、遺伝子マーカー、炎症マーカー、及びそれらの組み合わせからなる群から選択される1つ又は複数のマーカーの存在、レベル又は遺伝子型を決定して、マーカープロファイルを得るステップと、

(b) ランダムフォレストなどの第1のアルゴリズムをマーカープロファイルに適用して、マーカープロファイルがIBDサンプルであるか非IBDサンプルであるかの決定を得るステップと、

(c) IBDサンプルとして指定されたサンプルに決定木又はルールセットを適用して、IBDサンプルが不確定サンプルであるかどうかを決定するステップと、

(d) サンプルが不確定でなく、IBDサンプルであるならば、ランダムフォレストな

50

どの第2のアルゴリズムをIBDサンプルに適用して、臨床垂型を決定するステップと、を含む方法を提供する。

【0046】

[0065]いくつかの実施形態では、本方法は、IBD診断を援助するために、統計解析の結果（すなわち診断プロファイル）を参照（すなわち診断モデル）と比較することをさらに含む。特定の実施形態では、本方法は、医師に価値のある診断の見識を提供するために、多数の血清学的、タンパク質及び／又は遺伝子マーカーを利用する。

【0047】

[0066]いくつかの実施形態では、本明細書の方法及びシステムは、pANCAの第1の血清学的マーカーを測定してpANCA2値を決定する。いくつかの態様では、pANCAが検出されない又はpANCAがDNaseに細胞質内で感受性である場合、pANCAは0の値が割り当てられる。pANCAがDNase感受性1+P、DNase感受性2+P、DNase感受性3+P又はDNase感受性4+Pの場合は、pANCAは1の値が割り当てられる。ここで、1+P～4+Pは、視覚的評価に基づく蛍光標識の増大量及び強度を示す。例えば、サンプル中のpANCAの値が+2、+3又は+4であるならば、pANCA2値は陽性である。

【0048】

[0067]他のいくつかの態様では、pANCAが検出されない、pANCAがDNaseに細胞質内で感受性である、又はDNase感受性1+Pである場合、pANCA2は0の値（陰性）が割り当てられる。pANCAがDNase感受性2+P、DNase感受性3+P、又はDNase感受性4+Pである場合、pANCA2は1の値（陽性）が与えられる。

【0049】

[0068]本明細書で詳細に説明するように、pANCA2が陰性である場合、方法及びシステムは、

a) pANCA2陰性サンプルがクローン病の直接的な予測となるかどうかを決定するステップであって、該決定は、血清マーカーパネル（血清マーカーパネルは、ASCA-IgA、ASCA-IgG及びOmpC-IgAからなる群のメンバーである。）を測定し、血清マーカーパネルの血清マーカーの各々をカットオフ値と比較して、サンプルがクローン病に一致するかどうかを決定することによって行われる、ステップ、あるいは

b) pANCA2陰性サンプルがクローン病に一致するかどうかを決定するステップであって、該決定は、CD計数パネル（CD計数パネルは、ASCA-IgA、ASCA-IgG、OmpC-IgA、CBir1-IgG、A4-Fla2-IgG及びFlax-IgGからなる群のメンバーである。）を測定してCD計数値を形成することによって行われ、CD計数値が2以上である場合、サンプルはクローン病に一致する、ステップ、あるいは

c) pANCA値がゼロであるpANCA2陰性サンプルがIBDに一致するかどうかを決定するステップであって、該決定は、ANCAのELISAを測定し、ANCAのELISA値を参照値と比較し、ANCAのELISA値に基づいて、サンプルをIBDに一致する又は一致しないものとして指定することによって行われる、ステップ、あるいは

d) pANCA値がゼロであるpANCA2陰性サンプルがIBDに一致するか不確定のIBDであるかを決定するステップであって、該決定は、ANCAのELISAを測定し、ANCAのELISA値を比較し、CD計数値を考慮して、サンプルがIBDに一致するかどうか、又はサンプルがIBDに一致し、クローン病若しくは潰瘍性大腸炎に関して不確定であるかを決定することによって行われる、ステップ、あるいは

e) pANCA値が1であるpANCA2陰性サンプルがIBDに一致するかどうかを決定するステップであって、該決定は、ANCAのELISAを測定し、ANCAのELISA値をカットオフ値と比較して、サンプルがIBDに一致するかどうかを決定することによって行われる、ステップをさらに含む。

【0050】

[0069]いくつかの例では、上記パートaのASCA-IgAの値が69 EU/mL以上であるならば、サンプルはクローン病に一致する。

【0051】

[0070]他の例では、上記パートaのASCA-IgGの値が40 EU/mL以上であるならば、サンプルはクローン病に一致する。

【0052】

[0071]さらに別の例では、上記パートaのOmpC-IgAの値が60 EU/mL以上であるならば、サンプルはクローン病に一致する。

【0053】

[0072]さらに、本システム及び方法は、クローン病(CD)計数値の割り当てを提供する。例えば、上記パートbでは、CD計数値は下記i~vのルールを用いて決定される。

i. ASCA-IgAが8.5 EU/mL以上である(ASCA-IgAの参照範囲より高い)ならば、CD計数値に+1を与える。

ii. ASCA-IgGが17.8 EU/mL以上である(ASCA-IgGの参照範囲より高い)ならば、CD計数値に+1を与える。

iii. OmpC-IgAが10.9 EU/mL以上である(OmpC-IgAの参照範囲より高い)ならば、CD計数値に+1を与える。

iv. 2種以上のフラジェリンマーカーがそれぞれの参照範囲より高ければ、CD計数値に+1を与える。

v. 任意のフラジェリンが100 EU/mL以上であるならば、CD計数値に+1を与え、総CD計数値が2以上であるならば、サンプルがクローン病に一致すると指定する。

【0054】

[0073]いくつかの例では、CBir1-IgGの参照値は78.4 EU/mL以上である。

【0055】

[0074]他のいくつかの例では、A4-Fla2-IgGの参照値は44.8 EU/mL以上である。

【0056】

[0075]さらに別の例では、Flax-IgGの参照値は33.4 EU/mL以上である。

【0057】

[0076]他のいくつかの態様では、本明細書の方法及びシステムは、上記パートcに関して、pANCAがゼロ(検出されず)であるならば、

i. ANCAのELISAが20 EU/mL未満であるならば、サンプルをIBDに一致しないものとして指定するステップ、又は

ii. ANCAのELISAが27.4 EU/mLより高ければ、サンプルを潰瘍性大腸炎に一致するものとして指定するステップ、又は

iii. ANCAのELISAが20 EU/mL以上かつ27.4 EU/mL以下であり、CD計数値がゼロであるならば、サンプルをIBDに一致しないものとして指定するステップ、又は

iv. ANCAのELISAが20 EU/mL以上かつ27.4 EU/mL以下であり、CD計数値が1であるならば、サンプルを、IBDに一致するが、クローン病又は潰瘍性大腸炎に関して不確定であるものとして指定するステップをさらに含む。

【0058】

[0077]さらに別の態様では、本明細書の方法及びシステムは、pANCAが1である場合、

i. ANCAのELISAが13.7 EU/mL未満であるならば、サンプルをIBDに一致しないものとして指定するステップ、又は

10

20

30

40

50

i i. ANCAのELISAが13.7 EU/mL以上であるならば、サンプルが潰瘍性大腸炎に一致するものとして指定するステップをさらに含む。

【0059】

[0078]本発明は、炎症性腸疾患（IBD）の改善された診断のため、及び潰瘍性大腸炎（UC）とクローン病（CD）とを鑑別するための方法及びシステムを提供する。いくつかの例では、本方法及びシステムは、単にCD又はUCとしてではなく、IBD亜型の連続的なつながりに沿った疾患重症度の分類を可能にする。いくつかの例では、本方法は、不確定なものとしてサンプルをカテゴリー化する。合併疾患を有する患者を同定し、具体的な疾患の種類の評価を補助することによって、本明細書に記載の方法及びシステムは、疾患の重症度及び処置の選択肢を評価するための貴重な情報をもたらす。一部の実施形態では、統計解析を血清学的、タンパク質、及び／又は遺伝子マーカーのプロファイルに適用することで、IBD、UC及びCDを予測することの精度が改善され、生物学的療法、通常療法、手術、又はそれらの組み合わせなどの治療を含む適切な処置の選択も可能となる。

10

【0060】

[0079]本発明は、潰瘍性大腸炎（UC）を診断するため、及びUCとクローン病（CD）とを鑑別するための改善された方法及びシステムを提供する。いくつかの例では、本方法及びシステムは、単にCD又はUCとしてではなく、IBD亜型の連続的なつながりに沿った疾患重症度の分類を可能にする。合併疾患を有する患者を特定し、具体的な疾患の種類の評価を補助することによって、本明細書に記載の方法及びシステムは、疾患の重症度及び処置の選択肢を評価するための貴重な情報を提供する。一部の実施形態では、血清学的、タンパク質及び／又は遺伝子マーカーのプロファイルに統計解析を適用することによって、UC予測の精度が改善され、生物学的療法、通常療法、手術、又はそれらの組み合わせなどの治療を含む適切な処置の選択も可能になる。

20

【0061】

[0080]本発明は、クローン病（CD）の改善された診断のため、及びUCとCDとを鑑別するための方法及びシステムを提供する。いくつかの例では、本方法及びシステムは、単にCD又はUCとしてではなく、IBD亜型の連続的なつながりに沿った疾患重症度の分類を可能にする。合併疾患を有する患者を特定し、具体的な疾患の種類の評価を補助することによって、本明細書に記載の方法及びシステムは、疾患の重症度及び処置の選択肢を評価するための貴重な情報を提供する。一部の実施形態では、血清学的、タンパク質及び／又はマーカーのプロファイルに統計解析を適用することによって、CD予測の精度が改善され、生物学的療法、通常療法、手術、又はそれらの組み合わせなどの治療を含む適切な処置の選択も可能になる。

30

【0062】

[0081]本発明は、炎症性腸疾患（IBD）の改善された診断のため、及びIBDを非IBDから識別するための方法及びシステムを提供する。本発明はまた、非炎症性腸疾患（非IBD）を診断するため、及び非IBDをIBDから識別するための方法及びシステムを提供する。いくつかの例では、本方法及びシステムは、単にCD又はUCとしてではなく、IBD亜型の連続的なつながりに沿った疾患重症度の分類を可能にする。合併疾患を有する患者を特定し、具体的な疾患の種類の評価を補助することによって、本明細書に記載の方法及びシステムは、疾患の重症度及び処置の選択肢を評価するための貴重な情報を提供する。一部の実施形態では、血清学的、タンパク質及び／又は遺伝子マーカーのプロファイルに統計解析を適用することによってIBD及び非IBDの予測の精度が改善され、生物学的療法、通常療法、手術、又はそれらの組み合わせなどの治療を含む適切な処置の選択も可能になる。

40

【0063】

[0082]本発明は、不確定の改善された診断のため、及び不確定をUC及びCDから識別するための方法及びシステムを提供する。いくつかの例では、本方法及びシステムは、単

50

にC D又はU Cとしてではなく、I B D垂型の連続的なつながりに沿った疾患重症度の分類を可能にする。合併疾患を有する患者を特定し、具体的な疾患の種類の評価を補助することによって、本明細書に記載の方法及びシステムは、疾患の重症度及び処置の選択肢を評価するための貴重な情報を提供する。一部の実施形態では、血清学的、タンパク質及び／又は遺伝子マーカーのプロファイルに統計解析を適用することによって、垂型予測精度が改善され、生物学的療法、通常療法、手術、又はそれらの組み合わせなどの治療を含む適切な処置の選択も可能になる。

【0064】

[0083]一部の実施形態では、本発明の方法は、U Cの症状を有する個体において実施される。さらなる実施形態では、U Cの症状には、これらに限定されないが、直腸炎症、直腸出血、直腸痛、下痢、腹部痙攣、腹痛、疲労、体重減少、発熱、結腸破裂、及びそれらの組み合わせが含まれる。

10

【0065】

[0084]一部の実施形態では、本発明の方法は、全血、組織、唾液、頬細胞、毛、体液、血漿、血清、脳脊髄液、頬側スワブ、粘液、尿、便、精子、膺分泌物、リンパ、羊水、胸膜液、涙、及びそれらの組み合わせからなる群から選択される生体サンプルの分析を伴う。

【0066】

[0085]一部の態様において、本発明は、I B Dを診断するため、及びU C及びC DなどのI B Dの垂型を鑑別するための方法及びシステムを提供する。特定の実施形態では、本発明の方法及びシステムは、I B D、C D及び／又はU Cを有する患者を特定すること並びに価値のある診断の見識を医師に提供することを援助又は補助するために、1つ又は複数の（例えば多数の）遺伝子マーカーを、単独で又は1つ若しくは複数の血清学的な及び／又はタンパク質マーカーと組み合わせて、並びに単独で又は1つ若しくは複数のアルゴリズム（例えば、ランダムフォレストモデル及び決定木又はルールセット）若しくは他の種類の統計解析（例えば四分位数分析）と組み合わせて利用する。

20

【0067】

[0086]一部の実施形態では、I B Dを診断する方法はNKX2-3 (rs1190140) バリエント対立遺伝子の検出を含む。他の実施形態では、I B Dを診断する方法は、好ましくはNKX2-3 (rs10883365) バリエント対立遺伝子の検出を用いる。さらに別の実施形態では、I B Dを診断する方法はNKX2-3 (rs6584283) バリエント対立遺伝子の検出を用いる。一部の好ましい実施形態では、I B Dを診断する方法はATG16L1 (rs2241880) バリエント対立遺伝子の検出を含む。さらなる実施形態では、I B Dを診断する方法はATG16L1 (rs3828309) バリエント対立遺伝子の検出を含む。一部の実施形態では、I B Dを診断する方法はECM1 (rs7511649) バリエント対立遺伝子の検出を用いる。他の好ましい実施形態では、I B Dを診断する方法はECM1 (rs373240) バリエント対立遺伝子の検出を含む。さらに別の実施形態では、I B Dを診断する方法はECM1 (rs13294) バリエント対立遺伝子の検出を含む。一部の好ましい実施形態では、I B Dを診断する方法はSTAT3 (rs744166) バリエント対立遺伝子の検出を含む。

30

40

【0068】

[0087]一部の実施形態では、U Cを診断する方法はNKX2-3 (rs1190140) バリエント対立遺伝子の検出を含む。他の実施形態では、U Cを診断する方法はNKX2-3 (rs10883365) バリエント対立遺伝子の検出を含む。さらに別の実施形態では、U Cを診断する方法はNKX2-3 (rs6584283) バリエント対立遺伝子の検出を含む。一部の実施形態では、U Cを診断する方法はATG16L1 (rs2241880) バリエント対立遺伝子の検出を含む。さらなる実施形態では、U Cを診断する方法はATG16L1 (rs3828309) バリエント対立遺伝子の検出を含む。一部の実施形態では、U Cを診断する方法はECM1 (rs7511649) バリエント対立遺伝子の検出を含む。他の実施形態では、U Cを診断する方法はECM1 (rs373

50

240) バリエント対立遺伝子の検出を含む。さらに別の実施形態では、UCを診断する方法はECM1 (rs13294) バリエント対立遺伝子の検出を含む。一部の実施形態では、UCを診断する方法はSTAT3 (rs744166) バリエント対立遺伝子の検出を含む。

【0069】

[0088] 一部の実施形態では、CDを診断する方法はNKX2-3 (rs1190140) バリエント対立遺伝子の検出を含む。他の実施形態では、CDを診断する方法はNKX2-3 (rs10883365) バリエント対立遺伝子の検出を含む。さらに別の実施形態では、CDを診断する方法はNKX2-3 (rs6584283) バリエント対立遺伝子の検出を含む。一部の実施形態では、CDを診断する方法はATG16L1 (rs2241880) バリエント対立遺伝子の検出を用いる。さらなる実施形態では、CDを診断する方法はATG16L1 (rs3828309) バリエント対立遺伝子の検出を含む。一部の実施形態では、CDを診断する方法はECM1 (rs7511649) バリエント対立遺伝子の検出を含む。他の実施形態では、CDを診断する方法はECM1 (rs373240) バリエント対立遺伝子の検出を含む。さらに別の実施形態では、CDを診断する方法はECM1 (rs13294) バリエント対立遺伝子の検出を含む。一部の実施形態では、CDを診断する方法はSTAT3 (rs744166) バリエント対立遺伝子の検出を含む。

10

【0070】

[0089] 他の実施形態では、IBDを診断する方法は、ATG16L1 (例えば、rs2241880、rs3828309)、ECM1 (例えば、rs7511649、rs373240、rs13294)、NKX2-3 (例えば、rs1190140、rs10883365、rs6584283) 及びSTAT3 (例えばrs744166) からなる群から選択される1つ又は複数のバリエント対立遺伝子の検出を含む。

20

【0071】

[0090] 他の実施形態では、ICを診断する及び／又は不確定なものとしてサンプルをカテゴリー化する方法は、ATG16L1 (例えば、rs2241880、rs3828309)、ECM1 (例えば、rs7511649、rs373240、rs13294)、NKX2-3 (例えば、rs1190140、rs10883365、rs6584283) 及びSTAT3 (例えばrs744166) からなる群から選択される1つ又は複数のバリエント対立遺伝子の検出を含む。

30

【0072】

[0091] 他の実施形態では、UCを診断する方法は、ATG16L1 (例えば、rs2241880、rs3828309)、ECM1 (例えば、rs7511649、rs373240、rs13294)、NKX2-3 (例えば、rs1190140、rs10883365、rs6584283) 及びSTAT3 (例えばrs744166) からなる群から選択される1つ又は複数のバリエント対立遺伝子の検出を含む。

【0073】

[0092] 他の実施形態では、CDを診断する方法は、ATG16L1 (例えば、rs2241880、rs3828309)、ECM1 (例えば、rs7511649、rs373240、rs13294)、NKX2-3 (例えば、rs1190140、rs10883365、rs6584283) 及びSTAT3 (例えばrs744166) からなる群から選択される1つ又は複数のバリエント対立遺伝子の検出を含む。

40

【0074】

[0093] 遺伝子マーカー中のバリエント対立遺伝子の存在又は非存在は、本明細書に記載のアッセイを使用して決定することができる。バリエント対立遺伝子の状態を決定するために使用することができるアッセイには、これらに限定されないが、電気泳動分析アッセイ、制限長多型分析アッセイ、配列解析アッセイ、ハイブリダイゼーション分析アッセイ、PCR分析アッセイ、対立遺伝子特異的ハイブリダイゼーション、オリゴヌクレオチドライゲーション対立遺伝子特異的伸長／ライゲーション、対立遺伝子特異的増幅、一塩基

50

伸長、分子反転プローブ、侵襲性切断、選択的終結、制限長多型性、配列決定、一本鎖コンホメーション多型性 (SSCP)、一本鎖の鎖多型性、ミスマッチ切断、変性勾配ゲル電気泳動、及びそれらの組み合わせが含まれる。これらのアッセイは十分に記載されており、標準の方法が当分野で知られている。例えば、Ausubelら, Current Protocols in Molecular Biology, John Wiley & Sons, Inc., New York (1984~2008), 第7章及び補遺47; Theophilusら, "PCR Mutation Detection Protocols", Humana Press (2002); Innisら, PCR Protocols, San Diego, Academic Press, Inc. (1990); Maniatisら, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Lab., New York (1982); Ausubelら, Current Protocols in Genetics and Genomics, John Wiley & Sons, Inc., New York (1984~2008); 及びAusubelら, Current Protocols in Human Genetics, John Wiley & Sons, Inc., New York (1984~2008) (いずれも、参照によりすべての目的のためにそのまま本明細書に組み入れられる。)を参照されたい。

10

【0075】

[0094]特定の実施形態では、本明細書に記載の方法は、ANCA及び／又はpANCAに基づいてUCを診断する方法と比較してUCの診断を改善させる。

20

【0076】

[0095]他の実施形態では、IBD及び／又は垂型 (例えば、UC、CD) を診断する方法は、生体サンプルを血清学的マーカーの存在又はレベルについて分析する追加のステップを含み、血清学的マーカー存在又はレベルの検出と1つ又は複数のバリエーション対立遺伝子の存在とが相まってIBD及び／又は垂型 (例えば、UC、CD) の診断をさらに改善させる。

【0077】

[0096]さらに別の実施形態では、IBD及び／又は垂型 (例えば、UC、CD) を診断する方法は、抗好中球抗体、抗サッカロマイセス・セレビスエ抗体、抗微生物抗体、急性期タンパク質、アポリポタンパク質、デフェンシン、成長因子、サイトカイン、カドヘリン、細胞接着分子、又は本明細書に記載のマーカーの任意の組み合わせから選択される血清学的マーカーの検出を含む。

30

【0078】

[0097]さらなる実施形態では、IBD及び／又は垂型 (例えば、UC、CD) を診断する方法は、ANCA及びpANCAのうちの一つ又はANCAとpANCAの組み合わせから選択される抗好中球抗体を利用する。一実施形態では、抗好中球抗体は、免疫アッセイ (例えばELISA) によって検出されるANCAなどの抗好中球細胞質抗体 (ANCA)、免疫組織化学的アッセイ (例えばIFA) 若しくはDNAse感受性免疫組織化学的アッセイによって検出されるpANCAなどの核周囲型抗好中球細胞質抗体 (pANCA)、又はそれらの組み合わせを含む。

40

【0079】

[0098]さらに別の追加的な実施形態では、IBD及び／又は垂型 (例えば、UC、CD) を診断する方法は、抗サッカロマイセス・セレビスエ免疫グロブリンA (ASCA-IgA)、抗サッカロマイセス・セレビスエ免疫グロブリンG (ASCA-IgG)、及びそれらの組み合わせからなる群から選択される抗サッカロマイセス・セレビスエ抗体を利用する。

【0080】

[0099]さらに別の実施形態では、IBD及び／又は垂型 (例えば、UC、CD) を診断する方法は、抗外膜タンパク質C (抗OmpC) 抗体、抗I2抗体、抗フラジェリン抗体 (例えば、抗CBir1抗体、抗Fla2抗体 (すなわち抗A4-Fla2)、抗Fla

50

X抗体、及びそれらの組み合わせからなる群から選択される抗微生物抗体を利用する。

【0081】

[0100]さらに別の実施形態では、IBD及び／又は垂型（例えば、UC、CD）を診断する方法は、急性期タンパク質（例えばC反応性タンパク質（CRP））を利用する。

【0082】

[0101]さらに別の実施形態では、IBD及び／又は垂型（例えば、UC、CD）を診断する方法は、アポリポタンパク質（例えば血清アミロイドA（SAA））を利用する。

【0083】

[0102]さらに別の実施形態では、IBD及び／又は垂型（例えば、UC、CD）を診断する方法は、βデフェンシン（BD1）、βデフェンシン-2（BD2）、又はそれらの組み合わせからなる群から選択されるデフェンシンを利用する。

10

【0084】

[0103]さらに別の実施形態では、IBD及び／又は垂型（例えば、UC、CD）を診断する方法は、上皮成長因子（EGF）、血管内皮成長因子（VEGF）、又はそれらの組み合わせからなる群から選択される成長因子を利用する。

【0085】

[0104]さらに別の実施形態では、IBD及び／又は垂型（例えば、UC、CD）を診断する方法は、TNF関連のアポトーシスの弱い誘導因子（TWEAK）、IL-1β、IL-6、又はそれらの組み合わせからなる群から選択されるサイトカインを利用する。

【0086】

20

[0105]さらに別の実施形態では、IBD及び／又は垂型（例えば、UC、CD）を診断する方法はカドヘリン（例えばE-カドヘリン）を利用する。

【0087】

[0106]さらに別の実施形態では、IBD及び／又は垂型（例えば、UC、CD）を診断する方法は、細胞間接着分子1（ICAM-1又はICAM）、血管細胞接着分子1（VCAM-1又はVCAM）、又はそれらの組み合わせからなる群から選択される細胞接着分子を利用する。

【0088】

[0107]他の実施形態では、血清学的マーカーは、ANCA、pANCA（例えば、pANCA IFA及び／又はDNAse感受性pANCA IFA）、ASCA-IgA、ASCA-IgG、抗OmpC抗体（OmpC）、抗CBir-1抗体（CBir1）、抗Fla2抗体（Fla2又はA4-Fla2）、抗FlaX抗体（FlaX）、又はそれらの組み合わせを含むか、あるいはそれからなる。

30

【0089】

[0108]他の実施形態では、炎症マーカーは、VEGF、CRP、SAA、ICAM、VCAM、又はそれらの組み合わせを含むか、あるいはそれからなる。

【0090】

[0109]血清学的マーカー又は炎症マーカーの存在又は（濃度）レベルは、ハイブリダイゼーションアッセイ、増幅ベースのアッセイ、免疫アッセイ、免疫組織化学的アッセイ、又はそれらの組み合わせを用いて検出（例えば、決定、測定、分析）することができる。サンプル中の1つ又は複数の血清学的マーカーの存在又はレベルを検出又は決定するためのアッセイ、方法及びキットの非限定例は本明細書に記載されている。サンプル中の1つ又は複数の炎症マーカーの存在又はレベルを検出又は決定するためのアッセイ、方法及びキットの非限定例は本明細書に記載されている。

40

【0091】

[0110]特定の実施形態では、本発明の方法及びシステムは、1つ又は複数の（例えば多数の）遺伝子マーカーを、単独で又は1つ若しくは複数の血清学的及び／又は炎症マーカーと組み合わせて利用する。

【0092】

[0111]一態様において、本発明は、炎症性腸疾患（IBD）と診断されている及び／又

50

は潰瘍性大腸炎（UC）の疑いのある個体におけるUCを診断するための方法を提供する。一部の実施形態では、本方法は、（i）個体から得た生体サンプルを分析して、生体サンプルにおける複数の血清－遺伝子－炎症マーカーの存在、非存在又はレベルを決定して、マーカープロファイルを得るステップ（ここで、遺伝子は、ATG16L1、ECM1、STAT3又はNKX2-3のうちの1つ又は複数、血清学的マーカーは、ANCA、pANCA（例えば、pANCA IFA及び／又はDNase感受性pANCA IFA）、ASCA-IgA、ASCA-IgG、抗OmpC抗体、抗CBir-1抗体、抗Fla2抗体、抗FlaX抗体のうちの1つ又は複数、炎症マーカーは、VEGF、CRP、SAA、ICAM-1又はVCAM-1のうちの1つ又は複数である。）と、（ii）UCの診断を援助するために、マーカープロファイルに統計解析を適用して、個体に対する診断プロファイルを得るステップと、を含む。

10

【0093】

[0112]他の態様において、本発明は、炎症性腸疾患（IBD）と診断されている及び／又はクローン病（CD）の疑いのある個体におけるCDを診断するための方法を提供する。一部の実施形態では、本方法は、（i）個体から得た生体サンプルを分析して、生体サンプルにおける複数の血清－遺伝子－炎症マーカーの存在、非存在又はレベルを決定して、マーカープロファイルを得るステップ（ここで、遺伝子は、ATG16L1、ECM1、STAT3又はNKX2-3のうちの1つ又は複数、血清学的マーカーは、ANCA、pANCA（例えば、pANCA IFA及び／又はDNase感受性pANCA IFA）、ASCA-IgA、ASCA-IgG、抗OmpC抗体、抗CBir-1抗体、抗Fla2抗体、抗FlaX抗体のうちの1つ又は複数、炎症マーカーは、VEGF、CRP、SAA、ICAM-1又はVCAM-1のうちの1つ又は複数である。）と、（ii）CDの診断を援助するためにマーカープロファイルに統計解析を適用して、個体に対する診断プロファイルを得るステップと、を含む。

20

【0094】

[0113]他の態様において、本発明は、炎症性腸疾患（IBD）と診断されている及び／又は分類不能大腸炎（IC）の疑いのある個体におけるICを診断するための方法を提供する。一部の実施形態では、本方法は、（i）個体から得た生体サンプルを分析して、生体サンプルにおける複数の血清－遺伝子－炎症マーカーの存在、非存在又はレベルを決定して、マーカープロファイルを得るステップ（ここで、遺伝子は、ATG16L1、ECM1、STAT3又はNKX2-3のうちの1つ又は複数、血清学的マーカーは、ANCA、pANCA（例えば、pANCA IFA及び／又はDNase感受性pANCA IFA）、ASCA-IgA、ASCA-IgG、抗OmpC抗体、抗CBir-1抗体、抗Fla2抗体、抗FlaX抗体のうちの1つ又は複数、炎症マーカーは、VEGF、CRP、SAA、ICAM-1又はVCAM-1のうちの1つ又は複数である。）と、（ii）ICの診断を援助するために、マーカープロファイルに統計解析を適用して、個体に対する診断プロファイルを得るステップと、を含む。

30

【0095】

[0114]一部の実施形態では、本発明の方法は、UC、CD又はICの診断を援助するために、個体に対する診断プロファイルを診断モデルに関係付けることを含むさらなるステップを含む。

40

【0096】

[0115]一部の実施形態では、IBDと非IBDとを鑑別する方法は、生体サンプルにおける複数の血清－遺伝子－炎症マーカーの存在、非存在又はレベルの検出を含む。特定の実施形態では、複数の血清－遺伝子－炎症マーカーの存在、非存在又はレベルの検出はIBDの指標になるが、非IBDの指標にならない。

【0097】

[0116]特定の実施形態では、UCとCDとを鑑別する方法は、IBDを有するものとして決定された患者由来の生体サンプルにおける複数の血清－遺伝子－炎症マーカーの存在、非存在又はレベルの検出を含む。一部の実施形態では、複数の血清－遺伝子－炎症マ

50

カーの存在、非存在又はレベルの検出はUCの指標になるが、CDの指標にならない。他の実施形態では、複数の血清－遺伝子－炎症マーカーの存在、非存在又はレベルの検出はCDの指標になるが、UCの指標にならない。

【0098】

[0117]他の実施形態において、本発明は、個体群において、抗好中球抗体、抗サッカロマイセス・セレビスエ抗体、抗微生物抗体、急性期タンパク質、アポリポタンパク質、デフェンシン、成長因子、サイトカイン、カドヘリンなどの血清学的マーカー；ATG16L1（例えば、rs2241880、rs3828309）、ECM1（例えば、rs7511649、rs373240、rs13294）、NKX2-3（例えば、rs1190140、rs10883365、rs6584283）及びSTAT3（例えばrs744166）の遺伝子バリエーションなどの遺伝子マーカー；並びにこれらのマーカーの任意の組み合わせからなる群から選択される少なくとも1つの血清－遺伝子－炎症マーカーと潰瘍性大腸炎（UC）の存在との関連性を検出するための方法を提供する。一部の具体的な実施形態では、本方法は、（i）IBDと診断されている及び／又はUCの疑いのある個体群から生体サンプルを得るステップと、（ii）生体サンプルをスクリーニングして、抗好中球抗体、抗サッカロマイセス・セレビスエ抗体、抗微生物抗体、急性期タンパク質、アポリポタンパク質、デフェンシン、成長因子、サイトカイン、カドヘリンなどの血清学的マーカー；ATG16L1（例えば、rs2241880、rs3828309）、ECM1（例えば、rs7511649、rs373240、rs13294）、NKX2-3（例えば、rs1190140、rs10883365、rs6584283）及びSTAT3（例えばrs744166）の遺伝子バリエーションなどの遺伝子マーカー、；並びに任意のそれらの組み合わせからなる群から選択される少なくとも1つの血清－遺伝子－炎症マーカーの存在、非存在又はレベルを決定するステップと、（iii）血清学的マーカーのうちの1つ又は複数がIBDと診断されている及び／又はUCの疑いのある個体群に偏っている統計的に有意な分布を示すかどうかを評価するステップであって、比較は、IBDと診断されている及び／又はUCの疑いのある個体群と対照個体群との間で行われる、ステップと、を含む。

【0099】

[0118]さらに別の実施形態において、本発明は、個体群において、抗好中球抗体、抗サッカロマイセス・セレビスエ抗体、抗微生物抗体、急性期タンパク質、アポリポタンパク質、デフェンシン、成長因子、サイトカイン、カドヘリンなどの血清学的マーカー；ATG16L1（例えば、rs2241880、rs3828309）、ECM1（例えば、rs7511649、rs373240、rs13294）、NKX2-3（例えば、rs1190140、rs10883365、rs6584283）及びSTAT3（例えばrs744166）の遺伝子バリエーションなどの遺伝子マーカー；並びにマーカーの任意の組み合わせからなる群から選択される少なくとも1つの血清－遺伝子－炎症マーカーとクローン病（CD）の存在との関連性を検出するための方法を提供する。一部の具体的な実施形態では、本方法は、（i）IBDと診断されている及び／又はCDの疑いのある個体群から生体サンプルを得るステップと、（ii）生体サンプルをスクリーニングして、抗好中球抗体、抗サッカロマイセス・セレビスエ抗体、抗微生物抗体、急性期タンパク質、アポリポタンパク質、デフェンシン、成長因子、サイトカイン、カドヘリンなどの血清学的マーカー；ATG16L1（例えば、rs2241880、rs3828309）、ECM1（例えば、rs7511649、rs373240、rs13294）、NKX2-3（例えば、rs1190140、rs10883365、rs6584283）及びSTAT3（例えばrs744166）の遺伝子バリエーションなどの遺伝子マーカー；並びに任意のそれらの組み合わせからなる群から選択される少なくとも1つの血清－遺伝子－炎症マーカーの存在、非存在又はレベルを決定するステップと、（iii）血清学的マーカーのうちの1つ又は複数がIBDと診断されている及び／又はCDの疑いのある個体群に偏っている統計的に有意な分布を示すかどうかを評価するステップであって、比較は、IBDと診断されている及び／又はCDの疑いのある個体群と対照個体群との間で行わ

れる、ステップと、を含む。

【0100】

[0119]他の実施形態では、遺伝子マーカーは、表A～E（例えば、表A、B、C、D及び／又はE）に記載されている遺伝子の少なくとも1つである。一部の実施形態では、遺伝子マーカーはNKX2-3である。他の実施形態では、遺伝子マーカーはATG16L1である。さらに別の実施形態では、遺伝子マーカーはECM1である。いくつかの実施形態では、遺伝子マーカーはSTAT3である。遺伝子マーカーの遺伝子型は、例えば1つ又は複数の遺伝子マーカーの1つ又は複数の一塩基多型（SNP）などの1つ又は複数のバリエーション対立遺伝子の有無について、個体をジェノタイピングすることによって検出（例えば、決定、分析）することができる。一部の実施形態では、SNPは、表B～E（例えば、表B、C、D及び／又はE）に記載されているSNPの少なくとも1つである。サンプル中の1つ又は複数の遺伝子マーカーの遺伝子型を検出又は決定するための方法の非限定例は本明細書に記載されている。いくつかの実施形態では、遺伝子マーカーはATG16L1、ECM1、NKX2-3及び／又はSTAT3である。いくつかの例では、1つ又は複数のATG16L1、ECM1、NKX2-3及び／又はSTAT3のSNPの有無は、抗好中球抗体、抗サッカロマイセス・セレビスエ抗体、抗微生物抗体、急性期タンパク質、アポリポタンパク質、デフェンシン、成長因子、サイトカイン、カドヘリン、細胞接着分子、及びそれらの組み合わせからなる群から選択される1つ又は複数のマーカーの存在又はレベルと組み合わせて決定される。非限定例としては、ASCA-IgA、ASCA-IgG、抗OmpC抗体、抗CBir-1抗体、抗フラジェリン抗体、pANCA（例えば、pANCA-IFA及び／又はDNAse感受性pANCA-IFA）、CRP、SAA、ANCA、VEGF、ICAM、VCAM、又はそれらの組み合わせが挙げられる。

10

20

【0101】

[0120]さらなる態様では、本明細書に記載のマーカーのうちの1つ又は複数の測定するためのパネルは、IBDの診断及び／又はUCとCDとの鑑別のための本発明のアプローチに関する関連情報をもたらすために構築されていてもよい。そのようなパネルは、遺伝子、生化学的、血清学的、タンパク質、炎症性、又は本明細書に記載の他のマーカーなどの、少なくとも1個、2個、3個、4個、5個、6個、7個、8個、9個、10個、11個、12個、13個、14個、15個、16個、17個、18個、19個、20個、25個、30個、35個、40個、又はそれより多くの個々のマーカーの存在（若しくは非存在）又はレベルを検出又は決定するために構築されていてもよい。本発明のマーカープロファイルは、1つ又は複数の本明細書に記載のマーカーの測定値を含む。また、単一のマーカー又はマーカーのサブセットの分析は、当業者によって種々の臨床設定において実施することもできる。それらには、例えば、通院、応急手当、救命医療、集中治療、モニタリングユニット、入院患者、外来患者、医院、診療所、及び集団検診の設定が含まれる。

30

【0102】

[0121]一部の実施形態では、マーカーの分析は種々の物理的様式で実施することができる。例えば、多数の試験サンプルの処理を容易にするために、マイクロタイタープレート又は自動化を使用することができる。あるいは、適時の様式での処置、診断、及び予後診断を容易にするために、単一サンプルの様式を開発することができる。

40

【0103】

[0122]いくつかの実施形態では、本方法は、IBD診断を援助するために、統計解析の結果（すなわち診断プロファイル）を参照（すなわち診断モデル）と比較することをさらに含む。特定の実施形態では、本方法は、医師に価値のある診断の見識を提供するために、多数の血清学的、タンパク質及び／又は遺伝子マーカーを利用する。

【0104】

[0123]本発明の方法では、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又は20のマーカーの存在、レベル又は遺伝子型を検出することによって決定することができる。特

50

定の実施形態では、サンプルは、血清、血漿、全血及び／又は便である。他の実施形態では、個体はクローン病（C D）、潰瘍性大腸炎（U C）又は不確定と診断される。

【0105】

[0124] マーカープロファイルに適用される統計解析は、本明細書に記載の種々の統計的な方法、モデル及びアルゴリズムのうちのいずれかを含み得る。特定の実施形態では、統計解析は四分位数分析である。一部の例では、四分位数分析は、各マーカーの存在、レベル又は遺伝子型を四分位数スコアに変換する。非限定例としては、I B D 診断は、各マーカーに対する四分位数スコアを合計することによって得られる、個体に対する合計四分位数スコア（Q S S）に基づいて行うことができる。他の実施形態では、統計解析は、本明細書に記載の1つ又は複数の学習統計的分類子システムを含む。特定の実施形態では、統計解析は少なくとも2つの学習統計的分類子システムの組み合わせを含む。そうした組み合わせの非限定例としては、例えばタンデムで適用される、決定／分類木（例えば、分類及び回帰木（C & R T）、ランダムフォレスト）及び神経回路網が挙げられる。いくつかの例では、本方法は、第1の統計解析（例えば、決定／分類木又はランダムフォレスト）を各マーカーについて決定された存在、レベル又は遺伝子型に適用して、予測又は確率値を生成するステップと、次いで、第2の統計解析（例えば、決定／分類木又はランダムフォレスト）を予測若しくは確率値及び各マーカーについて決定された存在、レベル若しくは遺伝子型に適用して、（例えば、I B D サンプル若しくは非 I B D サンプルとしてサンプル进行分类することによって）I B D 診断を援助する及び／又はU C とC D とを鑑別するステップと、を含む。いくつかの実施形態では、第1のランダムフォレストがマーカープロファイルに適用され、続いて決定木又はルールセット、続いてさらに別のランダムフォレストが適用される。

【0106】

[0125] 他の実施形態では、診断プロファイルは個体に対する合計四分位数スコア（Q S S）であり、Q S S は診断モデル（例えば、血清学的モデル、血清－遺伝子－炎症モデル、標準化されたリスク尺度）と比較される。いくつかの実施形態では、Q S S が、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24 などより大きい（例えば、好ましくは10より大きい）場合は、個体は疾患の確率が高いと予測される。

【0107】

[0126] いくつかの実施形態では、診断モデルは、I B D の臨床亜型（例えば、C D、U C）の既知の転帰を使用する後ろ向きコホートをを用いて確立される。好ましい実施形態では、診断モデルは、血清学的モデル、血清－遺伝子－炎症モデル、遺伝子モデル、及びそれらの組み合わせからなる群から選択される。特定の実施形態では、血清学的モデルは、後ろ向きコホートで決定された1つ又は複数の血清学的マーカーの存在又はレベルに対してロジスティック回帰分析を適用することによって得られる。他の特定の実施形態では、血清－遺伝子－炎症モデルは、後ろ向きコホートで決定された1つ又は複数の血清学的マーカー及び1つ又は複数の遺伝子マーカーの遺伝子型の存在又はレベルに対してロジスティック回帰分析を適用することによって得られる。他の実施形態では、血清－遺伝子－炎症モデルは、後ろ向きコホートで決定された1つ又は複数の血清学的マーカー及び1つ又は複数の遺伝子マーカーの遺伝子型の存在又はレベルに対してランダムフォレストモデルを適用することによって得られる。

【0108】

[0127] 特定の実施形態では、本明細書に記載の方法は、（少なくとも）約5%、10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、又は100%（又はその中の任意の範囲）の割合で、個体の診断プロファイル（例えばQ S S）に基づいて、また、1つ又は複数の遺伝子マーカー（例えば、A T G 1 6 L 1、S T A T 3、E C M 1 及び／又はN K X 2 - 3）における1つ

又は複数のバリエーション対立遺伝子の有無、並びに1つ又は複数の血清学的マーカー（例えば、ASCA-IgA、ASCA-IgG、抗OmpC抗体、抗CBir-1抗体、抗フラジェリン抗体、pANCA（例えば、pANCA IFA及び／又はDNase感受性pANCA IFA）、CRP、SAA、ANCA、VEGF、ICAM、VCAM、又はそれらの組み合わせ）のレベルと任意選択で組み合わせて、UC又はCD診断の予測を提供する。

【0109】

[0128]いくつかの実施形態では、参照（濃度）レベルは、CDを有さない個体（例えば、健常個体、非CD個体、非IBD個体、UC個体）に由来するサンプルにおけるマーカーのうちの一つの（濃度）レベルに対応する。他のいくつかの実施形態では、参照遺伝子型は、遺伝子マーカーのうちの一つの野生型遺伝子型（例えば、非バリエーション対立遺伝子又はSNP）に対応する。

10

【0110】

[0129]特定の実施形態では、本明細書に記載の方法は、（少なくとも）約5%、10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、又は100%（又はその中の任意範囲）の割合で、個体の診断プロファイル（例えばQSS）に基づいて、また、1つ又は複数の遺伝子マーカー（例えば、ATG16L1、ECM1、NKX2-3及び／又はSTAT3）における1つ又は複数のバリエーション対立遺伝子の有無、並びに1つ又は複数の血清学的マーカー（例えば、ASCA-IgA、ASCA-IgG、抗OmpC抗体、抗CBir-1抗体、抗フラジェリン抗体、pANCA（例えば、pANCA IFA及び／又はDNase感受性pANCA IFA）、CRP、SAA、ANCA、VEGF、ICAM、VCAM、又はそれらの組み合わせ）のレベルと任意選択で組み合わせて疾患の予測を提供する。

20

【0111】

[0130]いくつかの実施形態では、参照（濃度）レベルは、IBDを有さない個体（例えば、健常個体、非IBD個体、非CD個体、非UC個体）に由来するサンプルにおけるマーカーのうちの一つの（濃度）レベルに対応する。他のいくつかの実施形態では、参照遺伝子型は、遺伝子マーカーのうちの一つの野生型遺伝子型（例えば、非バリエーション対立遺伝子又はSNP）に対応する。

30

【0112】

[0131]特定の実施形態では、本明細書に記載の方法は、（少なくとも）約5%、10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、又は100%（又はその中の任意範囲）のIBD（又はその臨床亜型）の確率を個体のマーカーレベル及び／又は遺伝子型に基づいて提供する。

【0113】

[0132]いくつかの実施形態では、本方法は、IBD診断を援助するために、統計解析の結果（すなわち診断プロファイル）を参照（すなわち診断モデル）と比較することをさらに含む。一部の例では、診断モデルは、参照表又はグラフなどの表示、印字出力及び／又は報告を含む。他の例では、診断プロファイルは、個体に対する合計四分位数スコア（QSS）であり、QSSは診断モデルと比較される。一部の実施形態では、QSSが、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24などより大きい場合は、個体はIBDを有する確率が高いと予測される。

40

【0114】

[0133]いくつかの好ましい実施形態では、本発明のIBD診断アルゴリズムは、17種の血清-遺伝子-炎症の生物学的マーカー（例えば、ANCA、ASCA-A、ASCA

50

—G、pANCA、FlaX、SAA、Fla2、ICAM、OmpC、Cbir1、VCAM、CRP、NKX2-3、ATG16L1、STAT3、ECM1及びVEGF）由来の測定値を使用して、IBD対非IBDを予測するための第1のランダムフォレストモデルに基づいてモデルスコアを算出する（110）。第1のランダムモデルは、患者のサンプルがIBDであるか又は非IBDサンプルであるかどうかを決定する。スコアが、IBD対非IBDのカットオフ未満（例えば <0.64 ）であるならば、サンプルは、IBDを有する患者由来、すなわちIBDサンプルであると予測される（125）。そうでなければ、サンプルは、IBDを有さない患者由来と予測される（120）。IBDを有するものとして予測されるサンプルは、アルゴリズムの次のステップへ進み、このステップは、分類不能大腸炎（IC）を除外する又はサンプルを不確定としてカテゴリー化するのを除外するように設計されている決定木又はルールセットである（130）。サンプルがICルールのいずれかに対するパターンに一致するならば、アルゴリズムは、IBDサンプルをICであると予測するか、又はアルゴリズムは不確定としてサンプルをカテゴリー化する（135）。そうでなければ、サンプルはアルゴリズムの次のステップに進み（140）、このステップは、UC又はCDを予測するための第2のランダムフォレストモデルである（150）。本発明の診断アルゴリズムは、11種の血清—遺伝子—炎症の生物学的マーカー（例えば、ANCA、ASCA-A、ASCA-G、FlaX、Fla2、pANCA、OmpC、Cbir1、ECM1、STAT3、VEGF、及びそれらの組み合わせ）由来の測定値を使用して、UC又はCDを予測するための第2のランダムフォレストモデルに基づいてモデルスコアを算出する。スコアが、UC対CDカットオフ（例えば 0.35 ）未満であるならば、アルゴリズムは、CDを有するものとしてサンプルを予測する（153）。スコアがカットオフより高ければ、アルゴリズムはUCを有するものとしてサンプルを予測する（155）。

10

20

【0115】

[0134]いくつかの実施形態では、ICルール又は不確定ルールは、 $ANCA \geq Q3$ 、pANCA 2陽性、及び抗Cbir1抗体、Fla2抗体若しくは抗FlaX抗体のいずれかが $\geq Q3$ であり得る。他の実施形態では、ICルールは、pANCA 2陽性、及び血清学的な抗フラジェリン抗体（例えば、抗Cbir1抗体、抗Fla2抗体、抗FlaX抗体）マーカーのうちの任意の2種が $\geq Q3$ であり得る。いくつかの態様では、ICルールは不確定ルールでもある。

30

【0116】

[0135]本発明の方法が、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、又は17種の血清—遺伝子—炎症マーカーをIBD診断試験に使用することを含み、診断試験が診断アルゴリズムを含み、アルゴリズムが、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、又は17種の血清—遺伝子—炎症マーカーを含むサンプルセット（例えば訓練セット）から生じることが、当業者であれば分かるであろう。血清遺伝子マーカーの非限定例としては、例えば、ANCA、ASCA-A、ASCA-G、pANCA、FlaX、SAA、Fla2、ICAM、OmpC、Cbir1、VCAM、CRP、NKX2-3、ATG16L1、STAT3、ECM1、VEGF、及びそれらの組み合わせが挙げられる。

40

【0117】

[0136]一部の実施形態では、診断モデルは、参照表又はグラフなどの表示、印字出力及び／又は報告を含む。特定の実施形態では、参照表又はグラフは、IBDを有する又は有さない個体の累積確率を提供する。他のいくつかの実施形態では、参照表又はグラフは、IBDを有する又は有さない個体の累積確率を提供する。他のいくつかの実施形態では、参照表又はグラフは、ICを有する又は有さない個体の累積確率を提供する。他の実施形態では、参照表又はグラフは、CDを有する又は有さない個体の累積確率を提供する。参照表又はグラフは、UCを有する又は有さない個体の累積確率を提供することもできる。

【0118】

IV. IBDの臨床亜型

50

炎症性腸疾患（IBD）は、大腸及び小腸の一群の炎症状態である。IBDの主な形態は、クローン病（CD）及び潰瘍性大腸炎（UC）である。IBDの他の稀な形態には、例えば、分類不能大腸炎（IC）、コラーゲン大腸炎、リンパ球性大腸炎、虚血性大腸炎、空置大腸炎、ベーチェット症候群、感染性大腸炎が含まれる。CD及びUCに関して不確定のIBDもIBDの形態として本明細書に含まれる。

【0119】

A. クローン病

クローン病（CD）とは、胃腸管の任意の部分に關与する場合がある、慢性炎症の疾患である。一般的に、小腸の遠位部分、すなわち回腸、及び盲腸が冒される。他の事例では、疾患は小腸、結腸、又は肛門直腸の領域に局限されている。CDは、時折十二指腸及び胃に、より稀に食道及び口腔に關与する。

10

【0120】

[0139] CDの変動する臨床徴候は、部分的に、疾患の変動する解剖学的局在化の結果である。CDの最も頻度の高い症状は、腹痛、下痢、及び回帰熱である。CDは、一般的に、腸閉塞又は腸瘻、腸の罹患したループ間の異常な通路に關連づけられている。また、CDには、眼、関節、及び皮膚の炎症、肝疾患、腎結石、並びにアミロイド症などの合併症も含まれる。さらに、CDは、腸癌のリスクの増加に關連づけられている。

【0121】

[0140] いくつかの特長は、CD病理に特徴的である。貫壁性炎症として知られるCDに關連づけられている炎症は、腸壁の全層に關与する。例えば、肥厚及び浮腫も、典型的には腸壁全体にわたって出現し、線維症が疾患の長期にわたる形態で存在する。CDの炎症特徴は、「飛び石病変」として知られる炎症組織の区域が見かけ上正常な腸管によって分離されているという点で不連続である。さらに、介在組織の線状潰瘍形成、浮腫、及び炎症は、CDに特有な腸管粘膜の「敷石状」の外見をもたらす。

20

【0122】

[0141] CDの特徴は、一般に粘膜下層で見つかる、肉芽腫として知られる炎症細胞の孤立性の凝集の存在である。一部のCD症例は典型的な孤立性の肉芽腫を示す一方で、他のものはびまん性肉芽腫性反応又は非特異的な貫壁性炎症を示す。その結果、孤立性の肉芽腫の存在はCDの指標であるが、肉芽腫の非存在も疾患において一貫している。したがって、肉芽腫の存在ではなく、貫壁性又は不連続の炎症がCDの好ましい診断的指標である（Rubin及びFarber, Pathology（第2版）, Philadelphia, J. B. Lippincott Company（1994））。

30

【0123】

[0142] クローン病は、疾患が進行する際の挙動によってカテゴリー化してもよい。これはクローン病のウィーン分類に形式化されている。Gascheら, Inflamm. Bowel Dis., 6: 8~15（2000）を参照されたい。クローン病では3つの疾患提示のカテゴリー、すなわち、（1）狭窄性、（2）穿通性、及び（3）炎症性が存在する。狭窄性疾患は、腸閉塞又は糞便の口径の変化をもたらすことがある、腸の狭小化を引き起こす。穿通性疾患は、腸と皮膚などの他の構造との間に異常な通路（瘻孔）を生じる。炎症性疾患（非狭窄性、非穿通性疾患としても知られる。）は、狭窄又は瘻孔を引き起こさずに炎症を引き起こす。

40

【0124】

[0143] したがって、クローン病は胃腸管を冒すいくつかの異種の疾患亜型を表し、同様の症状を生じることがある。CDに關連して本明細書中で使用される用語「臨床亜型」には、1つのCDの分類を他のものから区別する1セットの臨床基準によって定義されたCDの分類が含まれる。非限定的な例として、CDを有する対象は、本明細書に記載されている狭窄性（例えば内部狭窄性）、穿通性（例えば内部穿通性）又は炎症性疾患に罹患しているとして分類することができる。あるいは、これらの対象は、それに加えて若しくはその代わりに、線維狭窄性疾患、小腸疾患、内部穿孔性疾患、肛門周囲瘻形成疾患、UC様疾患、小腸手術の必要性、UC特長の非存在、又はそれらの組み合わせに罹患している

50

として分類することができる。

【0125】

[0144]いくつかの例では、CDを有する対象は、狭窄性又は穿通性の表現型によって特徴づけられている臨床亜型である合併CDに罹患しているとして分類することができる。他のいくつかの例では、CDを有する対象は、次の合併症：線維狭窄性疾患、及び小腸手術の必要性のうちの1つ又は複数によって特徴づけられているCDの形態に罹患しているとして分類することができる。さらなる例では、CDを有する対象は、小腸手術を必要とする侵襲性形態の線維狭窄性疾患に罹患しているとして分類することができる。これらの亜型に関連する基準は、例えば、Gascheら、Inflamm. Bowel Dis., 6:8~15 (2000); Abreuら、Gastroenterology, 123:679~688 (2002); Vasiliasら、Gut, 47:487~496 (2000); Vasiliasら、Gastroenterology, 110:1810~1819 (1996); 及びGreensteinら、Gut, 29:588~592 (1988)に記載されている。

10

【0126】

[0145]CDの「線維狭窄性亜型」は、線維狭窄性疾患の1つ又は複数の認められている特徴によって特徴づけられているCDの分類である。線維狭窄性疾患のそのような特徴には、これらに限定されないが、記述されている持続性の腸閉塞又は腸閉塞の腸切除が含まれる。CDの線維狭窄性亜型には穿孔、膿瘍、又は瘻孔などの他の症状が伴う場合があり、嘔気、嘔吐、腹部膨満、及び固形食品の摂食不能などの持続性の腸管遮断の症状によってさらに特徴づけることができる。CDの線維狭窄性亜型を有する患者の腸管のX線は、例えば、遮断点より前の腸の膨満感を示す場合がある。

20

【0127】

[0146]CDの線維狭窄性亜型を有する対象における小腸手術の必要性は、この亜型のより侵襲性の形態を示している場合がある。CDの追加の亜型も当分野で知られており、定義された臨床基準を使用して同定することができる。例えば、内部穿孔性疾患は、腸管—腸内若しくは腸管—小胞の瘻孔、腹腔内膿瘍、又は小腸穿孔の現在又は以前の証拠によって定義されたCDの臨床亜型である。肛門周囲穿孔性疾患は、肛門周囲の瘻孔若しくは膿瘍又は直腸腔瘻の現在又は以前の証拠によって定義されたCDの臨床亜型である。CDのUC様臨床亜型は、左側結腸障害、出血又は緊急の症状、及び結腸生検の陰窩膿瘍の現在又は以前の証拠によって定義することができる。疾患の位置は、1つ又は複数の内視鏡、放射線又は病理学の研究に基づいて分類することができる。

30

【0128】

[0147]当業者は、CDの臨床亜型間に重複が存在する場合があり、CDに罹患している対象はCDの複数の臨床亜型を有する場合があることを理解している。例えば、CDに罹患している対象はCDの線維狭窄性亜型を有することができ、小腸手術の必要性又は内部穿孔性疾患亜型によって特徴づけられている臨床亜型の臨床基準も満たすことができる。同様に、本明細書に記載のマーカーはCDの複数の臨床亜型に関連づけることができる。

【0129】

B. 潰瘍性大腸炎

潰瘍性大腸炎(UC)とは、痙攣、腹痛、直腸出血、血液の緩い分泌、膿、及び粘液を伴った慢性下痢によって特徴づけられている大腸の疾患である。UCの徴候は広く変動する。増悪及び寛解のパターンがUC患者の約70%の臨床的経過の典型であるが、寛解のない連続的な症状がUCを有する一部の患者で存在する。UCの局所的及び全身性の合併症には、関節炎、ブドウ膜炎などの眼の炎症、皮膚潰瘍、及び肝疾患が含まれる。さらに、UC、特に疾患の長期にわたる大規模な形態は結腸癌のリスクの増加に関連づけられている。

40

【0130】

[0149]UCは、通常は直腸の最も遠位の部分から可変距離で近位に伸びる、びまん性疾患である。用語「左側大腸炎」とは、脾臓湾曲部にまで伸びる結腸の遠位部分に關与する

50

炎症を説明する。直腸が冒されないこと（sparing）又は結腸の右側（近位部分）が単独で関与することはUCでは珍しい。UCの炎症プロセスは結腸に限定されており、例えば小腸、胃又は食道には及ばない。さらに、UCは、一般に腸壁のより深層は冒さない粘膜の表在性炎症によって区別される。また、変性腸管陰窩が好中球で満たされている陰窩膿瘍もUCにおいて典型的である（Rubin及びFarber（上記））。

【0131】

[0150]いくつかの例では、UCに関して、症状の変性は、疾患の程度（すなわち、炎症を起こしている結腸及び直腸の量）及び炎症の強度の相違を反映する。疾患は直腸から始まり、結腸を「上に」移動して器官のより多くの部分に及ぶ。UCは、関与した結腸の量によってカテゴリー化することができる。典型的には、炎症が直腸及び直腸に隣接する結腸の短い区域に限局されている患者は、より広汎性の結腸の炎症を有する患者よりも軽度な症状及び良好な予後診断を有する。

10

【0132】

[0151]直腸が高頻度で冒されていない、斑状の疾患であるCDに比べて、UCは、通常は近位よりも遠位でより重度である結腸の連続的な炎症によって特徴づけられている。UCにおける炎症は、通常は粘膜層に限定されているという点で表在性であり、好中球及び陰窩膿瘍を伴った急性炎症性浸潤によって特徴づけられている。対照的に、CDは腸壁の厚さ全体を冒し、必ずではないが多くの場合に肉芽腫が存在する。回盲弁で、又はそれに遠位の結腸中で終結する疾患はUCの指標である一方で、回腸末端の関与、敷石状様の外見、孤立性の潰瘍、又は瘻孔はCDを示唆する。

20

【0133】

[0152]種々の潰瘍性大腸炎が炎症の位置及び程度に従って分類されている。UCに関連して本明細書中で使用される用語「臨床亜型」には、1つのUCの分類を他のものから区別する1セットの臨床基準によって定義されたUCの分類が含まれる。非限定的な例として、UCを有する対象は、潰瘍性直腸炎、直腸S状結腸炎、左側大腸炎、汎大腸炎、劇症大腸炎、及びそれらの組み合わせに罹患しているとして分類することができる。これらの亜型に関連する基準は、例えばKornbluthら、Am. J. Gastroenterol., 99:1371~85（2004）に記載されている。

【0134】

[0153]潰瘍性直腸炎とは、直腸に限定されている炎症によって定義されたUCの臨床亜型である。直腸S状結腸炎とは、直腸及びS状結腸を冒すUCの臨床亜型である。左側大腸炎とは、結腸の左側全体の、直腸から結腸が脾臓付近で曲がって上腹部を横切り始める場所（脾臓湾曲部）を冒すUCの臨床亜型である。汎大腸炎とは、結腸全体を冒すUCの臨床亜型である。劇症大腸炎は、汎大腸炎の稀であるが重度の形態である。劇症大腸炎を有する患者は病状が非常に重く、脱水、重度の腹痛、出血を伴う遅延性の下痢、さらにはショックを伴っている。

30

【0135】

[0154]一部の実施形態では、UCの臨床亜型の分類は、有効な処置経過を計画する際に重要である。潰瘍性直腸炎、直腸S状結腸炎、及び左側大腸炎は、ステロイド系又は他の浣腸及び泡沫を含む肛門から導入した局所剤で処置することができる一方で、汎大腸炎は、活性成分が結腸の冒された部分の全体に達することができるように、経口医薬品で処置しなければならない。

40

【0136】

[0155]当業者は、UCの臨床亜型間に重複が存在する場合があります、UCに罹患している対象はUCの複数の臨床亜型を有する場合がありますことを理解している。同様に、本明細書に記載の予後診断マーカーはUCの複数の臨床亜型に関連づけることができる。

【0137】

C. 分類不能大腸炎

分類不能大腸炎（IC）とは、CD及びUCの両方の特長が含まれるIBDの臨床亜型である。両疾患の症状のそのような重複は、ICを有する患者において一時的に（例えば

50

、疾患の初期に)又は持続的に(例えば、疾患が進行する間中)起こる場合がある。臨床的には、ICは、直腸出血を伴う又は伴わない腹痛及び下痢によって特徴づけられている。例えば、正常な粘膜によって分離された断続的な多数の潰瘍形成を有する大腸炎が、この疾患を有する患者で見つかる。組織学的には、貫壁性炎症を伴った重度の潰瘍形成のパターンが存在する。直腸には典型的には疾患が存在せず、リンパ系炎症細胞は凝集を示さない。ICを有する患者では、深いスリット様の亀裂が筋細胞溶解の病巣で観察されるが、介在粘膜は典型的には最小限に鬱血しており、杯細胞は保存されている。いくつかの態様では、決定木又はルールセットは不確定又はICとしてサンプルをカテゴリー化する。

【0138】

V. IBDマーカー

生化学マーカー、血清学的マーカー、タンパク質マーカー、遺伝子マーカー及び臨床的又は超音波検査の他の特徴を含む種々のIBDマーカーは、IBDを診断若しくは予測するため又はIBDの診断若しくは予測を援助若しくは改善するための本発明の方法において使用するのに適する。例えば、米国特許出願公開第20110045476号及び国際出願番号PCT/US2011/039174(その開示内容は、参照によりすべての目的のためにそのまま本明細書に組み入れられる。)を参照されたい。いくつかの態様では、本明細書に記載の診断方法は、アルゴリズム(例えば統計解析)を、IBDマーカーのうちの1つ又は複数について決定された存在、濃度レベル、又は遺伝子型に適用することを利用して、IBDの診断又は予測を援助若しくは補助する。

【0139】

[0158] IBDマーカーの非限定的な例には、(i)例えば、サイトカイン、成長因子、抗好中球抗体、抗サッカロマイセス・セレビシエ抗体、抗微生物抗体、急性期タンパク質、アポリポタンパク質、デフェンシン、カドヘリン、細胞接着分子、及びそれらの組み合わせなどの生化学的、血清学的、及びタンパク質マーカー、(ii)例えば表A~Eに記載されている遺伝子のうちのいずれか1つ又はそれらの遺伝子の組み合わせなどの遺伝子マーカー、及び(iii)血清学的及び遺伝子マーカーの組み合わせが含まれる。

【0140】

A. サイトカイン

サンプル中の少なくとも1つのサイトカインの存在又はレベルの決定は本発明において有用であり得る。本明細書において、用語「サイトカイン」には、種々の免疫系の機能を調節する免疫細胞によって分泌された種々のポリペプチド又はタンパク質のうちの任意のものが含まれ、ケモカインなどの小サイトカインが包含される。また、用語「サイトカイン」には、例えば、体重、造血、血管形成、創傷治癒、インスリン耐性、免疫応答、及び炎症反応の調節において機能する脂肪細胞によって分泌されたサイトカインの群を含むアディポサイトカインも含まれる。

【0141】

[0160]いくつかの態様では、これらに限定されないが、TNF- α 、TNF関連のアポトーシスの弱い誘導因子(TWEAK)、オステオプロテゲリン(OPG)、IFN- α 、IFN- β 、IFN- γ 、IL-1 α 、IL-1 β 、IL-1受容体拮抗剤(IL-1ra)、IL-2、IL-4、IL-5、IL-6、可溶性IL-6受容体(sIL-6R)、IL-7、IL-8、IL-9、IL-10、IL-12、IL-13、IL-15、IL-17、IL-23、及びIL-27を含む少なくとも1つのサイトカインの存在又はレベルをサンプル中で決定する。他のいくつかの態様では、例えば、CXCL1/GRO1/GRO α 、CXCL2/GRO2、CXCL3/GRO3、CXCL4/PF-4、CXCL5/ENA-78、CXCL6/GCP-2、CXCL7/NAP-2、CXCL9/MIG、CXCL10/IP-10、CXCL11/I-TAC、CXCL12/SDF-1、CXCL13/BCA-1、CXCL14/BRAK、CXCL15、CXCL16、CXCL17/DMC、CCL1、CCL2/MCP-1、CCL3/MIP-1 α 、CCL4/MIP-1 β 、CCL5/RANTES、CCL6/C10、CCL7/MCP-3、CCL8/MCP-2、CCL9/CCL10、CCL11/エ

10

20

30

40

50

オタキシ、CCL12/MCP-5、CCL13/MCP-4、CCL14/HCC-1、CCL15/MIP-5、CCL16/LEC、CCL17/TARC、CCL18/MIP-4、CCL19/MIP-3 β 、CCL20/MIP-3 α 、CCL21/SLC、CCL22/MDC、CCL23/MPIF1、CCL24/エオタキシン-2、CCL25/TECK、CCL26/エオタキシン-3、CCL27/CTACK、CCL28/MEC、CL1、CL2、及びCX₃CL1などの少なくとも1つのケモカインの存在又はレベルをサンプル中で決定する。いくつかのさらなる態様では、これらに限定されないが、レプチン、アディポネクチン、レジスチン、活性又は全プラスミノーゲン活性化阻害剤-1 (PAI-1)、ビスファチン、及びレチノール結合タンパク質4 (RBP4)を含む少なくとも1つのアディポサイトカインの存在又はレベルをサンプル中で決定する。

10

【0142】

[0161]いくつかの例では、特定のサイトカインの存在又はレベルは、例えばハイブリダイゼーションアッセイ又は増幅ベースのアッセイなどのアッセイを用いて、mRNA発現のレベルで検出する。他のいくつかの例では、特定のサイトカインの存在又はレベルは、例えば免疫アッセイ (例えばELISA) 又は免疫組織化学的アッセイを使用して、タンパク質発現のレベルで検出する。血清、血漿、唾液、又は尿サンプルなどのサンプル中のサイトカインの存在又はレベルを決定するための適切なELISAキットは、例えば、R&D Systems, Inc. (Minneapolis, MN)、Neogen Corp. (Lexington, KY)、Alpco Diagnostics (Sal-

20

【0143】

B. 成長因子

サンプル中の1つ又は複数の成長因子の存在又はレベルの決定は本発明において有用であり得る。本明細書において、用語「成長因子」には、細胞増殖及び/又は細胞分化を刺激することができる種々のペプチド、ポリペプチド、又はタンパク質のうちの任意のものが含まれる。

30

【0144】

[0163]いくつかの態様では、これらに限定されないが、表皮成長因子 (EGF)、ヘパリン結合表皮成長因子 (HB-EGF)、血管内皮成長因子 (VEGF)、色素上皮由来因子 (PEDF、セルピンF1としても知られる)、アンフィレギュリン (AREG、シュワン腫由来成長因子 (SDGF) としても知られる)、塩基性線維芽細胞成長因子 (bFGF)、肝細胞成長因子 (HGF)、トランスフォーミング成長因子- α (TGF- α)、トランスフォーミング成長因子- β (TGF- β)、骨形成タンパク質 (例えば、BMP1~BMP15)、血小板由来成長因子 (PDGF)、神経成長因子 (NGF)、 β -神経成長因子 (β -NGF)、神経栄養因子 (例えば、脳由来神経栄養因子 (BDNF)、ニューロトロフィン3 (NT3)、ニューロトロフィン4 (NT4))、成長分化因子-9 (GDF-9)、顆粒球-コロニー刺激因子 (G-CSF)、顆粒球-マクロファージコロニー刺激因子 (GM-CSF)、ミオスタチン (GDF-8)、エリスロポエチン (EPO)、及びトロンプオエチン (TPO)を含む少なくとも1つの成長因子の存在又はレベルをサンプル中で決定する。好ましくは、VEGFの存在又はレベルを決定する。

40

50

【0145】

[0164] VEGFは血管内皮成長因子遺伝子 (VEGF; Entrez Gene ID: 7422) によってコードされており、転写物 (例えば、Genbank 受託番号 NM_001025366. 2 (アイソフォーム a)、NM_003376. 5 (アイソフォーム b)、NM_001025367. 2 (アイソフォーム c)、NM_001025368. 2 (アイソフォーム d)、NM_001025369. 2 (アイソフォーム e)、NM_001025370. 2 (アイソフォーム f)、NM_001033756. 2 (アイソフォーム g)、NM_001171622. 1 (アイソフォーム h)、NM_001171623. 1 (アイソフォーム i)、NM_001171624. 1 (アイソフォーム j)、NM_001171625. 1 (アイソフォーム k)、NM_001171626. 1 (アイソフォーム l)、NM_001171627. 1 (アイソフォーム m)、NM_001171628. 1 (アイソフォーム n)、NM_001171629. 1 (アイソフォーム o)、NM_001171630. 1 (アイソフォーム p)、NM_001204384. 1 (アイソフォーム q)、及び NM_001204385. 1 (アイソフォーム r) を参照) の選択的スプライシング、及び前駆体ポリペプチドスプライスアイソフォーム (Genbank 受託番号 NP_001020537. 2 (アイソフォーム a)、NP_003367. 4 (アイソフォーム b)、NP_001020538. 2 (アイソフォーム c)、NP_001020539. 2 (アイソフォーム d)、NP_001020540. 2 (アイソフォーム e)、NP_001020541. 2 (アイソフォーム f)、NP_001028928. 1 (アイソフォーム g)、001165093. 1 (アイソフォーム h)、NP_001165094. 1 (アイソフォーム i)、NP_001165095. 1 (アイソフォーム j)、NP_001165096. 1 (アイソフォーム k)、NP_001165097. 1 (アイソフォーム l)、NP_001165098. 1 (アイソフォーム m)、NP_001165099. 1 (アイソフォーム n)、NP_001165100. 1 (アイソフォーム o)、NP_001165101. 1 (アイソフォーム p)、NP_001191313. 1 (アイソフォーム q)、NP_00191314. 1 (アイソフォーム r)) のプロセッシング後に産生される。

【0146】

[0165] いくつかの例では、特定の成長因子の存在又はレベルは、例えばハイブリダイゼーションアッセイ又は増幅ベースのアッセイなどのアッセイを用いて、mRNA 発現のレベルで検出する。他のいくつかの例では、特定の成長因子の存在又はレベルは、例えば免疫アッセイ (例えば ELISA) 又は免疫組織化学的アッセイを使用して、タンパク質発現のレベルで検出する。血清、血漿、唾液、又は尿サンプルなどのサンプル中の成長因子の存在又はレベルを決定するための適切な ELISA キットは、例えば、Antigenix America Inc. (Huntington Station, NY)、Promega (Madison, WI)、R&D Systems, Inc. (Minneapolis, MN)、Invitrogen (Camarillo, CA)、CHEMICON International, Inc. (Temecula, CA)、Neogen Corp. (Lexington, KY)、PeproTech (Rocky Hill, NJ)、Alpco Diagnostics (Salem, NH)、Pierce Biotechnology, Inc. (Rockford, IL)、及び/又は Abazyme (Needham, MA) から入手可能である。

【0147】

C. 抗好中球抗体

サンプル中の ANCA レベル及び/又は pANCA の存在若しくは非存在の決定は本発明において有用であり得る。本明細書において、用語「抗好中球細胞質抗体」又は「ANCA」には、好中球の細胞質及び/又は核成分に向けられた抗体が含まれる。ANCA 活性は、好中球中の ANCA 染色パターンに基づいていくつかの広義なカテゴリー、すなわち、(1) 核周囲のハイライトのない細胞質好中球染色 (cANCA)、(2) 核の外縁の周りの核周囲染色 (pANCA)、(3) 核の内縁の周りの核周囲染色 (NSNA)、

及び(4)好中球全体にわたって点在する広汎性染色(SAPPA)に分けることができる。いくつかの例では、pANCA染色はDNase処理に感受性がある。用語ANCAには、これらに限定されないが、cANCA、pANCA、NSNA、及びSAPPAを含む、抗好中球反応性のすべての種類が包含される。同様に、用語ANCAには、これらに限定されないが、免疫グロブリンA及びGを含むすべての免疫グロブリンアイソタイプが包含される。好ましくは、ANCA IgGの存在又はレベルが決定される。特定の実施形態では、例えば、別々の間接蛍光抗体(IFA)アッセイを使用して、核周囲のpANCA染色パターンの有無及びDNase処理に感受性であるpANCA染色の有無の両方が検出される。

【0148】

[0167]個体からのサンプル中のANCAレベルは、例えばアルコール固定した好中球を用いた酵素結合免疫吸着アッセイ(ELISA)などの免疫アッセイを用いて決定することができる。pANCAなどのANCAの特定のカテゴリーの存在又は非存在は、例えば間接蛍光抗体(IFA)アッセイなどの免疫組織化学的アッセイを用いて決定することができる。特定の実施形態では、サンプル中のpANCAの存在又は非存在は、DNase処理した固定した好中球を使用する免疫蛍光アッセイを用いて決定する。固定した好中球に加えて、ヒト抗体に向けられた抗体を検出に使用することができる。また、これらに限定されないが、未精製又は部分精製の好中球抽出物、ヒストンH1又はそのANCA反応性断片などの精製タンパク質、タンパク質断片、又は合成ペプチド(例えば、米国特許第6074835号を参照)、ヒストンH1様抗原、ポリリン抗原、バクテロイデス(Bacteroides)抗原、又はそのANCA反応性断片(例えば、米国特許第6033864号を参照)、分泌小胞抗原又はそのANCA反応性断片(例えば、米国特許出願公開第08/804106号を参照)、及び抗ANCAイディオタイプ抗体を含む、ANCAに特異的な抗原も、ANCAレベルの決定に適している。当業者には、ANCAに特異的なさらなる抗原の使用が本発明の範囲内にあることが理解されよう。

【0149】

[0168]いくつかの実施形態では、pANCAバイオマーカー値は陽性又は陰性のいずれかであるので、pANCAバイオマーカーは数値変数ではなく二値である。一部の実施形態では、pANCA陽性状態は、合併症(例えば、内部狭窄性疾患、内部穿通性疾患及び/又は手術)の低い割合及び/又はリスクと関係がある。一部の例では、陽性状態が「1」としてスコア付けされ、陰性状態が「4」としてスコア付けされるように、pANCAに対する四分位数のスコア付けは反転させられる。

【0150】

D. 抗サッカロマイセス・セレビシエ抗体

サンプル中のASCA(例えば、ASCA-IgA、ASCA-IgG、ASCA-IgM)の存在又はレベルの決定も本発明において有用であり得る。用語「抗サッカロマイセス・セレビシエ免疫グロブリンA」又は「ASCA-IgA」には、エス・セレビシエと特異的に反応する免疫グロブリンAアイソタイプの抗体が含まれる。同様に、用語「抗サッカロマイセス・セレビシエ免疫グロブリンG」又は「ASCA-IgG」には、エス・セレビシエと特異的に反応する免疫グロブリンGアイソタイプの抗体が含まれる。

【0151】

[0170]サンプルがASCA-IgA又はASCA-IgGに関して陽性であるかどうかの決定は、ヒト抗体配列に特異的な抗体又はASCAに特異的な抗原を用いて行われる。そのような抗原は、ASCA-IgA及び/又はASCA-IgGによって特異的に結合される任意の抗原又は抗原混合物であり得る。ASCA抗体はエス・セレビシエと結合するその能力によって最初に特徴づけられたが、抗原がASCA抗体と特異的に結合することができる限りは、ASCAによって特異的に結合される抗原をエス・セレビシエ又は他の種々の供給源から得ることができることが当業者には理解されよう。したがって、サンプル中のASCA-IgA及び/又はASCA-IgGのレベルを決定するために使用することができるASCAに特異的な抗原の供給源の例には、これらに限定されないが、サ

ッカロマイセス属 (*Saccharomyces*) 又はカンジダ属 (*Candida*) 細胞などの死滅酵母細胞全体、ホスホペプチドマンナン (PPM) などの酵母細胞壁マンナン、オリゴマンノシドなどのオリゴ糖、ネオ糖脂質、抗 ASCA イディオタイプ抗体などが含まれる。エス・セレビスエ株 Su1、Su2、CBS1315、若しくは BM156、又はカンジダ・アルビカンス (*Candida albicans*) 株 VW32 などの酵母の種々の種及び株が、ASCA-IgA 及び／又は ASCA-IgG に特異的な抗原としての使用に適切である。また、ASCA に特異的な精製抗原及び合成抗原も、サンプル中の ASCA-IgA 及び／又は ASCA-IgG のレベルの決定における使用に適切である。精製抗原の例には、これらに限定されないが、オリゴマンノシドなどの精製オリゴ糖抗原が含まれる。合成抗原の例には、これらに限定されないが、米国特許出願公開第 20030105060 号に記載されているものなどの合成オリゴマンノシド、例えば、 $D\text{-Man}\beta(1-2)D\text{-Man}\beta(1-2)D\text{-Man}\beta(1-2)D\text{-Man-OR}$ 、 $D\text{-Man}\alpha(1-2)D\text{-Man}\alpha(1-2)D\text{-Man}\alpha(1-2)D\text{-Man-OR}$ 、及び $D\text{-Man}\alpha(1-3)D\text{-Man}\alpha(1-2)D\text{-Man}\alpha(1-2)D\text{-Man-OR}$ [式中、R は水素原子、 $C_1 \sim C_{20}$ アルキル、又は任意選択で標識された連結基である。] が含まれる。

【0152】

[0171] 酵母細胞壁マンナン (例えば PPM) の調製物を、サンプル中の ASCA-IgA 及び／又は ASCA-IgG のレベルの決定に使用することができる。そのような水溶性表面抗原は、当分野で知られている任意の適切な抽出法 (例えばオートクレーブ) によって調製することができるか、又は商業的に入手することができる (例えば、Lindberg ら, Gut, 33:909~913 (1992) を参照)。また、PPM の酸安定画分は本発明の統計アルゴリズムにおいても有用である (Sendid ら, Clin. Diag. Lab. Immunol., 3:219~226 (1996))。サンプル中の ASCA レベルの決定に有用な PPM は、例えばエス・ウバルム (*S. uvarum*) 株 ATCC # 38926 に由来する。

【0153】

[0172] オリゴマンノシドなどの精製オリゴ糖抗原も、サンプル中の ASCA-IgA 及び／又は ASCA-IgG のレベルの決定に有用であり得る。精製オリゴマンノシド抗原は、例えば Faille ら, Eur. J. Microbiol. Infect. Dis., 11:438~446 (1992) に記載されているように、好ましくはネオ糖脂質へと変換する。そのようなオリゴマンノシド抗原と ASCA との反応性が、マンノシル鎖の長さ (Frosh ら, Proc Natl. Acad. Sci. USA, 82:1194~1198 (1985))、アノマー立体配置 (Fukazawa ら, "Immunology of Fungal Disease", E. Kurstak (編), Marcel Dekker Inc., New York, 第 37~62 ページ (1989); Nishikawa ら, Microbiol. Immunol., 34:825~840 (1990); Poulain ら, Eur. J. Clin. Microbiol., 23:46~52 (1993); Shibata ら, Arch. Biochem. Biophys., 243:338~348 (1985); Trinel ら, Infect. Immun., 60:3845~3851 (1992))、又は結合の位置 (Kikuchi ら, Planta, 190:525~535 (1993)) を変動させることによって最適化できることは当業者には理解されよう。

【0154】

[0173] 本発明の方法で使用するための適切なオリゴマンノシドには、これらに限定されないが、マンノテトラオース $Man(1-3)Man(1-2)Man(1-2)Man$ を有するオリゴマンノシドが含まれる。そのようなオリゴマンノシドは、例えば Faille ら (上記) のように PPM から精製することができる。ASCA に特異的なネオ糖脂質は、例えば、オリゴマンノシドをその対応する PPM から放出させ、続いて放出されたオリゴマンノシドを 4-ヘキサデシルアニリンなどとカップリングさせることによって構

10

20

30

40

50

築することができる。

【0155】

E. 抗微生物抗体

サンプル中の抗Omp C抗体の存在又はレベルの決定は本発明において有用であり得る。本明細書において、用語「抗外膜タンパク質C抗体」又は「抗Omp C抗体」には、例えば、米国特許第7138237号及び国際公開第01/89361号に記載されている、細菌外膜ポリンに向けられた抗体が含まれる。用語「外膜タンパク質C」又は「Omp C」とは、抗Omp C抗体と免疫反応性の細菌ポリンをいう。用語「抗Omp C抗体」は、免疫グロブリンA及びGなどのすべての免疫グロブリンアイソタイプを包含する。好ましくは、抗Omp C Ig Aの存在又はレベルが決定される。

10

【0156】

[0175] 個体からのサンプル中に存在する抗Omp C抗体のレベルは、Omp Cタンパク質又はその免疫反応性断片などのその断片を用いて決定することができる。サンプル中の抗Omp C抗体レベルを決定するために有用な適切なOmp C抗原には、これらに限定されないが、Omp Cタンパク質、Omp Cタンパク質と実質的に同じアミノ酸配列を有するOmp Cポリペプチド、又はその免疫反応性断片などのその断片が含まれる。本明細書において、Omp Cポリペプチドは、一般に、Omp Cタンパク質と約50%より高い同一性、好ましくは約60%より高い同一性、より好ましくは約70%より高い同一性、さらにより好ましくは約80%、85%、90%、95%、96%、97%、98%、又は99%より高いアミノ酸配列同一性を有するアミノ酸配列を有するポリペプチドを説明し、アミノ酸同一性はCLUSTALWなどの配列アラインメントプログラムを用いて決定する。そのような抗原は、例えば、大腸菌(E. coli)などの腸内細菌からの精製によって、Genbank受託番号K00541などの核酸の組換え発現によって、液相若しくは固相ペプチド合成などの合成手段によって、又はファージディスプレイを使用することによって調製することができる。サンプル中の抗Omp C抗体レベルの決定は、ELISAアッセイ又は組織学的アッセイを用いて行うことができる。

20

【0157】

[0176] サンプル中の抗I 2抗体の存在又はレベルの決定も本発明において有用であり得る。本明細書において、用語「抗I 2抗体」には、例えば、米国特許第6309643号に記載されている、細菌転写調節因子と相同性を共有する微生物抗原に向けられた抗体が含まれる。用語「I 2」とは、抗I 2抗体と免疫反応性の微生物抗原をいう。微生物I 2タンパク質は、シー・パスツールリアヌム(C. pasteurianum)からの予測タンパク質4、マイコバクテリウム・ツベルクロシス(Mycobacterium tuberculosis)からのRv3557c、及びアクリフェックス・エオリカス(Aquifex aeolicus)からの転写調節因子とある程度の類似度の弱い相同性を共有する、100個のアミノ酸のポリペプチドである。I 2タンパク質の核酸及びタンパク質の配列は、例えば米国特許第6309643号に記載されている。

30

【0158】

[0177] 個体からのサンプル中に存在する抗I 2抗体のレベルは、I 2タンパク質又はその免疫反応性断片などのその断片を用いて決定することができる。サンプル中の抗I 2抗体レベルを決定するために有用な適切なI 2抗原には、これらに限定されないが、I 2タンパク質、I 2タンパク質と実質的に同じアミノ酸配列を有するI 2ポリペプチド、又はその免疫反応性断片などのその断片が含まれる。そのようなI 2ポリペプチドは、I 2タンパク質に対してシー・パスツールリアヌムタンパク質4よりも高い配列類似度を示し、そのアイソタイプバリエーション及びホモログが含まれる。本明細書において、I 2ポリペプチドは、一般に、天然に存在するI 2タンパク質と約50%より高い同一性、好ましくは約60%より高い同一性、より好ましくは約70%より高い同一性、さらにより好ましくは約80%、85%、90%、95%、96%、97%、98%、又は99%より高いアミノ酸配列同一性を有するアミノ酸配列を有するポリペプチドを説明し、アミノ酸同一性はCLUSTALWなどの配列アラインメントプログラムを用いて決定する。そのようなI

40

50

2 抗原は、例えば、微生物からの精製によって、I 2 抗原をコードしている核酸の組換え発現によって、液相若しくは固相ペプチド合成などの合成手段によって、又はファージディスプレイを使用することによって調製することができる。サンプル中の抗 I 2 抗体レベルの決定は、例えば、米国特許第 7 8 7 3 4 7 9 号に記載されている組織学的アッセイ又は E L I S A アッセイによって行うことができる。

【0159】

[0178] サンプル中の抗フラジェリン抗体の存在又はレベルの決定も本発明において有用であり得る。本明細書において、用語「抗フラジェリン抗体」には、例えば、米国特許第 7 3 6 1 7 3 3 号及び国際公開第 0 3 / 0 5 3 2 2 0 号に記載されている、細菌鞭毛のタンパク質成分に向けられた抗体が含まれる。用語「フラジェリン」とは、抗フラジェリン抗体と免疫反応性の細菌鞭毛タンパク質をいう。微生物フラジェリンは、例えば、自身を中空の円柱に配置してフィラメントを形成する、細菌鞭毛中に見つかるタンパク質である。

10

【0160】

[0179] 個体からのサンプル中に存在する抗フラジェリン抗体のレベルは、フラジェリンタンパク質又はその免疫反応性断片などのその断片を用いて決定することができる。サンプル中の抗フラジェリン抗体レベルを決定するために有用な適切なフラジェリン抗原には、これらに限定されないが、C B i r - 1 フラジェリン、A 4 - F l a 2 フラジェリン (F l a 2)、フラジェリン X (F l a X)、フラジェリン A、フラジェリン B、その断片、及びそれらの組み合わせなどのフラジェリンタンパク質、フラジェリンタンパク質と実質的に同じアミノ酸配列を有するフラジェリンポリペプチド、又はその免疫反応性断片などのその断片が含まれる。本明細書において、フラジェリンポリペプチドは、一般に、天然に存在するフラジェリンタンパク質と約 5 0 % より高い同一性、好ましくは約 6 0 % より高い同一性、より好ましくは約 7 0 % より高い同一性、さらにより好ましくは約 8 0 %、8 5 %、9 0 %、9 5 %、9 6 %、9 7 %、9 8 %、又は 9 9 % より高いアミノ酸配列同一性を有するアミノ酸配列を有するポリペプチドを説明し、アミノ酸同一性は C L U S T A L W などの配列アラインメントプログラムを用いて決定する。そのようなフラジェリン抗原は、例えば、ヘリコバクター・ビルス (H e l i c o b a c t e r B i l i s)、ヘリコバクター・ムステラエ (H e l i c o b a c t e r m u s t e l a e)、ヘリコバクター・ピロリ (H e l i c o b a c t e r p y l o r i)、ブチリビブリオ・フィブリソルベンス (B u t y r i v i b r i o f i b r i s o l v e n s)、及び盲腸中に見つかる細菌などの細菌からの精製によって、フラジェリン抗原をコードしている核酸の組換え発現によって、液相若しくは固相ペプチド合成などの合成手段によって、又はファージディスプレイを使用することによって調製することができる。サンプル中の抗フラジェリン (例えば、抗 C B i r 1、抗 F l a 2、及び／又は抗 F l a X) 抗体レベルの決定は、E L I S A アッセイ又は組織学的アッセイを用いて行うことができる。好ましくは、抗 C B i r 1 I g G、抗 F l a 2 I g G、及び／又は抗 F l a X I g G の存在又はレベルを決定する。

20

30

【0161】

F. 急性期タンパク質

サンプル中の 1 つ又は複数の急性期タンパク質の存在又はレベルの決定は本発明において有用であり得る。急性期タンパク質とは、その血漿濃度が炎症に応答して増加 (陽性急性期タンパク質) 又は減少 (陰性急性期タンパク質) するタンパク質のクラスである。この応答は急性期反応と呼ばれる (急性期応答とも呼ばれる)。陽性急性期タンパク質の例には、これらに限定されないが、C 反応性タンパク質 (C R P)、D - 二量体タンパク質、マンノース結合タンパク質、アルファ 1 - 抗トリプシン、アルファ 1 - 抗キモトリプシン、アルファ 2 - マクログロブリン、フィブリノーゲン、プロトロンビン、第 V I I I 因子、フォンウィルブランド因子、プラスミノーゲン、補体因子、フェリチン、血清アミロイド P 成分、血清アミロイド A (S A A)、オロソムコイド (アルファ 1 - 酸糖タンパク質、 A G P)、セルロプラスミン、ハプトグロビン、及びそれらの組み合わせが含まれ

40

50

る。陰性急性期タンパク質の非限定的な例には、アルブミン、トランスフェリン、トランスサイレチン、トランスコルチン、レチノール結合タンパク質、及びそれらの組み合わせが含まれる。好ましくは、CRP及び／又はSAAの存在又はレベルを決定する。

【0162】

[0181]いくつかの例では、特定の急性期タンパク質の存在又はレベルは、例えばハイブリダイゼーションアッセイ又は増幅ベースのアッセイなどのアッセイを用いて、mRNA発現のレベルで検出する。他のいくつかの例では、CRP又はSAAなどの急性期タンパク質の存在又はレベルは、例えば免疫アッセイ（例えば、ELISA、免疫電気化学発光アッセイ）又は免疫組織化学的アッセイを使用して、タンパク質発現のレベルで検出する。例えば、Alpco Diagnostics (Salem, NH) から入手可能なサンドイッチ比色ELISAアッセイを使用して、血清、血漿、尿又は糞便サンプル中のCRPのレベルを決定することができる。同様に、Biomedical Corporation (Foster City, CA) から入手可能なELISAキットを使用して、サンプル中のCRPレベルを検出することができる。サンプル中のCRPレベルを決定するための他の方法は、例えば、米国特許第6838250号及び第6406862号、並びに米国特許出願公開第20060024682号及び第20060019410号に記載されている。CRPレベルを決定するためのさらなる方法には、例えば、免疫比濁測定アッセイ、迅速免疫拡散アッセイ、及び視覚的凝集アッセイが含まれる。

【0163】

[0182]C反応性タンパク質（CRP）とは、血液中で炎症に応答して見つかるタンパク質である（急性期タンパク質）。CRPは、典型的には肝臓及び脂肪細胞（fat cell, adipocyte）によって産生される。CRPはペントラキシンのタンパク質ファミリーのメンバーである。ヒトCRPポリペプチド配列は、例えばGenbank受託番号NP_000558に記載されている。ヒトCRP mRNA（コード）配列は、例えばGenbank受託番号NM_000567に記載されている。当業者には、CRPがPTX1、MGC88244、及びMGC149895としても知られていることが理解されよう。

【0164】

G. アポリポタンパク質

サンプル中の1つ又は複数のアポリポタンパク質の存在又はレベルの決定は本発明において有用であり得る。アポリポタンパク質とは、脂肪（脂質）と結合するタンパク質である。これらは、血流を通して食事性脂肪を輸送するリポタンパク質を形成する。食事性脂肪は腸管内で消化され、肝臓へと運ばれる。また、脂肪は肝臓自体中でも合成される。脂肪は脂肪細胞内で貯蔵される。脂肪は格筋、心臓、及び他の器官において必要に応じてエネルギーのために代謝され、母乳中で分泌される。また、アポリポタンパク質は、酵素補因子、受容体リガンド、並びにリポタンパク質の代謝及び組織中へのその取り込みを調節する脂質輸送担体としても役割を果たす。アポリポタンパク質の例には、これらに限定されないが、ApoA（例えば、ApoA-I、ApoA-II、ApoA-IV、ApoA-V）、ApoB（例えば、ApoB48、ApoB100）、ApoC（例えば、ApoC-I、ApoC-II、ApoC-II-I、ApoC-IV）、ApoD、ApoE、ApoH、血清アミロイドA（SAA）、及びそれらの組み合わせが含まれる。好ましくは、SAAの存在又はレベルを決定する。

【0165】

[0184]いくつかの例では、特定のアポリポタンパク質の存在又はレベルは、例えばハイブリダイゼーションアッセイ又は増幅ベースのアッセイなどのアッセイを用いて、mRNA発現のレベルで検出する。他のいくつかの例では、SAAなどの特定のアポリポタンパク質の存在又はレベルは、例えば免疫アッセイ（例えば、ELISA、免疫電気化学発光アッセイ）又は免疫組織化学的アッセイを使用して、タンパク質発現のレベルで検出する。血清、血漿、唾液、尿、大便などのサンプル中のSAAの存在又はレベルを決定するための適切なELISAキットは、例えば、Antigenix America Inc

． (Huntington Station, NY)、Abazyme (Needham, MA)、USCN Life (Missouri City, TX)、及び／又はU. S. Biological (Swampscott, MA) から入手可能である。

【0166】

[0185]血清アミロイドA (SAA) タンパク質とは、血漿中の高密度リポタンパク質 (HDL) に関連づけられているアポリポタンパク質のファミリーである。SAAの種々のアイソフォームが、種々のレベルで構成的に (構成的SAA) 又は炎症性刺激に応答して (急性期SAA) 発現されている。これらのタンパク質は、主に肝臓によって産生される。無脊椎動物及び脊椎動物の全体にわたってこれらのタンパク質が保存されていることは、SAAがすべての動物において非常に重要な役割を果たすことを示唆している。急性期血清アミロイドAタンパク質 (A-SAA) は炎症の急性期中に分泌される。ヒトSAAポリペプチド配列は、例えばGenbank受託番号NP_000322に記載されている。ヒトSAA mRNA (コード) 配列は、例えばGenbank受託番号NM_000331に記載されている。当業者には、SAAがPIG4、TP53I4、MGC111216、及びSAA1としても知られていることが理解されよう。

【0167】

H. デフェンシン

サンプル中の1つ又は複数のデフェンシンの存在又はレベルの決定は本発明において有用であり得る。デフェンシンとは、脊椎動物及び無脊椎動物のいずれにおいても見られるシステインに富んだ小さな陽イオンタンパク質である。これらは、細菌、真菌、並びに多くの外被及び非外被ウイルスに対して活性である。これらは、典型的には、6個 (脊椎動物中) ～8個の保存的なシステイン残基を含む18～45個のアミノ酸からなる。免疫系の細胞は、例えば好中性顆粒球及びほぼすべての上皮細胞において、貪食された細菌の死滅を補助するためにこれらのペプチドを含有する。ほとんどのデフェンシンは、微生物細胞膜と結合することによって機能し、包埋された後、必須のイオン及び栄養素の流出を可能にする孔様の膜欠損を形成する。デフェンシンの非限定的な例には、 α -デフェンシン (例えば、DEFA1、DEFA1A3、DEFA3、DEFA4)、 β -デフェンシン (例えば、 β デフェンシン-1 (DEFB1)、 β デフェンシン-2 (DEFB2)、DEFB103A/DEFB103B～DEFB107A/DEFB107B、DEFB110～DEFB133)、及びそれらの組み合わせが含まれる。

【0168】

[0187]いくつかの例では、特定のデフェンシンの存在又はレベルは、例えばハイブリダイゼーションアッセイ又は増幅ベースのアッセイなどのアッセイを用いて、mRNA発現のレベルで検出する。他のいくつかの例では、特定のデフェンシンの存在又はレベルは、例えば免疫アッセイ (例えば、ELISA、免疫電気化学発光アッセイ) 又は免疫組織化学的アッセイを使用して、タンパク質発現のレベルで検出する。血清、血漿、唾液、尿、又は大便などのサンプル中のデフェンシンの存在又はレベルを決定するための適切なELISAキットは、例えば、Alpco Diagnostics (Salem, NH)、Antigenix America Inc. (Huntington Station, NY)、Peprotech (Rocky Hill, NJ)、及び／又はAlpha Diagnostic Intl. Inc. (San Antonio, TX) から入手可能である。

【0169】

[0188] β -デフェンシンは、微生物コロニー形成に対する上皮表面の耐性に関連づけられている抗微生物ペプチドである。これらはすべてのデフェンシンの中で最も広く分布しており、白血球及び多種の上皮細胞によって分泌される。例えば、これらは舌、皮膚、角膜、唾液腺、腎臓、食道、及び気道上に見つけることができる。ヒトDEFB1ポリペプチド配列は、例えばGenbank受託番号NP_005209に記載されている。ヒトDEFB1 mRNA (コード) 配列は、例えばGenbank受託番号NM_005218に記載されている。当業者には、DEFB1がBD1、HBD1、DEFB-1、D

E F B 1 0 1、及びM G C 5 1 8 2 2としても知られていることが理解されよう。ヒト D E F B 2 ポリペプチド配列は、例えばG e n b a n k 受託番号N P _ 0 0 4 9 3 3に記載されている。ヒトD E F B 2 m R N A (コード)配列は、例えばG e n b a n k 受託番号N M _ 0 0 4 9 4 2に記載されている。当業者には、D E F B 2 がS A P 1、H B D - 2、D E F B - 2、D E F B 1 0 2、及びD E F B 4としても知られていることが理解されよう。

【0170】

I. カドヘリン

サンプル中の1つ又は複数のカドヘリンの存在又はレベルの決定は本発明において有用であり得る。カドヘリンとは、細胞接着において重要な役割を果たしており、組織内の細胞と一緒に結合されていることを確実にする、1型膜貫通タンパク質のクラスである。これらは機能するためにカルシウム (Ca^{2+}) イオンに依存する。カドヘリンスーパーファミリーには、カドヘリン、プロトカドヘリン、デスモグレイン、及びデスモコリンなどが含まれる。構造的には、これらは細胞外 Ca^{2+} -結合ドメインであるカドヘリン反復を共有する。本発明における使用に適したカドヘリンには、これらに限定されないが、C D H 1 - E - カドヘリン (上皮)、C D H 2 - N - カドヘリン (中性)、C D H 1 2 - カドヘリン 1 2、2 型 (N - カドヘリン 2)、C D H 3 - P - カドヘリン (胎盤)、C D H 4 - R - カドヘリン (網膜)、C D H 5 - V E - カドヘリン (血管内皮)、C D H 6 - K - カドヘリン (腎臓)、C D H 7 - カドヘリン 7、2 型、C D H 8 - カドヘリン 8、2 型、C D H 9 - カドヘリン 9、2 型 (T 1 - カドヘリン)、C D H 1 0 - カドヘリン 1 0、2 型 (T 2 - カドヘリン)、C D H 1 1 - O B - カドヘリン (骨芽細胞)、C D H 1 3 - T - カドヘリン - H - カドヘリン (心臓)、C D H 1 5 - M - カドヘリン (筋小管)、C D H 1 6 - K S P - カドヘリン、C D H 1 7 - L I カドヘリン (肝臓-腸管)、C D H 1 8 - カドヘリン 1 8、2 型、C D H 1 9 - カドヘリン 1 9、2 型、C D H 2 0 - カドヘリン 2 0、2 型、及びC D H 2 3 - カドヘリン 2 3 (神経感覚上皮) が含まれる。

【0171】

[0190]いくつかの例では、特定のカドヘリンの存在又はレベルは、例えばハイブリダイゼーションアッセイ又は増幅ベースのアッセイなどのアッセイを用いて、m R N A 発現のレベルで検出する。他のいくつかの例では、特定のカドヘリンの存在又はレベルは、例えば免疫アッセイ (例えば、E L I S A、免疫電気化学発光アッセイ) 又は免疫組織化学的アッセイを使用して、タンパク質発現のレベルで検出する。血清、血漿、唾液、尿、又は大便などのサンプル中のカドヘリンの存在又はレベルを決定するための適切なE L I S A キットは、例えば、R & D S y s t e m s, I n c. (M i n n e a p o l i s, M N) 及び/又はG e n W a y B i o t e c h, I n c. (S a n D i e g o, C A) から入手可能である。

【0172】

[0191]E - カドヘリンとは、カドヘリンスーパーファミリーからの古典的カドヘリンである。これは、5つの細胞外カドヘリン反復、1つの膜貫通領域、及び1つの高度に保存的な細胞質側末端からなるカルシウム依存性の細胞間接着糖タンパク質である。E - カドヘリンの外部ドメインは哺乳動物細胞との細菌接着を媒介し、細胞質ドメインは内部移行に必要である。ヒトE - カドヘリンポリペプチド配列は、例えばG e n b a n k 受託番号N P _ 0 0 4 3 5 1に記載されている。ヒトE - カドヘリン m R N A (コード)配列は、例えばG e n b a n k 受託番号N M _ 0 0 4 3 6 0に記載されている。当業者には、E - カドヘリンがU V O、C D H E、E C A D、L C A M、A r c - 1、C D 3 2 4、及びC D H 1としても知られていることが理解されよう。

【0173】

J. 細胞接着分子 (I g S F C A M)

サンプル中の1つ又は複数の免疫グロブリンスーパーファミリー細胞接着分子の存在又はレベルの決定は本発明において有用であり得る。本明細書において、用語「免疫グロブリンスーパーファミリー細胞接着分子」(I g S F C A M)には、1つ又は複数の免疫

グロブリン様の折り畳みドメインを有しており、細胞間接着及び／又はシグナル伝達に機能する、細胞の表面に位置する種々のポリペプチド又はタンパク質のうちの任意のものが含まれる。多くの場合、IgSF CAMは膜貫通タンパク質である。IgSF CAMの非限定的な例には、神経細胞接着分子（NCAM、例えば、NCAM-120、NCAM-125、NCAM-140、NCAM-145、NCAM-180、NCAM-185）、細胞間接着分子（ICAM、例えば、ICAM-1、ICAM-2、ICAM-3、ICAM-4、ICAM-5）、血管細胞接着分子-1（VCAM-1）、血小板-内皮細胞接着分子-1（PECAM-1）、L1細胞接着分子（L1CAM）、L1CAMに相同性を有する細胞接着分子（L1の近いホモログ）（CHL1）、シアル酸結合Ig様レクチン（SIGLEC、例えば、SIGLEC-1、SIGLEC-2、SIGLEC-3、SIGLEC-4）、ネクチン（例えば、ネクチン-1、ネクチン-2、ネクチン-3）、及びネクチン様分子（例えば、Nec1-1、Nec1-2、Nec1-3、Nec1-4、Nec1-5）が含まれる。好ましくは、ICAM-1及び／又はVCAM-1の存在又はレベルを決定する。

10

【0174】

1. 細胞間接着分子-1（ICAM-1）

ICAM-1とは、白血球及び内皮細胞の膜中に低濃度で連続的に存在する膜貫通細胞接着タンパク質である。サイトカイン刺激の際、濃度は大きく増加する。ICAM-1はIL-1及びTNF α によって誘導することができ、血管内皮、マクロファージ、及びリンパ球によって発現される。IBDでは、炎症誘発性サイトカインは、ICAM-1及びVCAM-1などの接着分子の発現をアップレギュレーションすることによって炎症を引き起こす。増加した接着分子の発現がより多くのリンパ球を感染した組織に動員し、組織の炎症をもたらす（Gokeら、J. Gastroenterol., 32:480（1997）；Rijckenら、Gut, 51:529（2002）を参照）。ICAM-1は細胞間接着分子1遺伝子（ICAM1；Entrez Gene ID:3383；Genbank受託番号NM_000201）によってコードされており、細胞間接着分子1前駆体ポリペプチド（Genbank受託番号NP_000192）のプロセッシング後に産生される。

20

【0175】

2. 血管細胞接着分子-1（VCAM-1）

VCAM-1とは、血管内皮へのリンパ球、単球、好酸球、及び好塩基球の接着を媒介する膜貫通細胞接着タンパク質である。内皮細胞におけるサイトカインによるVCAM-1の上方制御は、増加した遺伝子転写（例えば、腫瘍壊死因子- α （TNF α ）及びインターロイキン-1（IL-1）に応答したもの）の結果として起こる。VCAM-1は血管細胞接着分子1遺伝子（VCAM1；Entrez Gene ID:7412）によってコードされており、転写物（Genbank受託番号NM_001078（バリエーション1）又はNM_080682（バリエーション2））の選択的スプライシング、及び前駆体ポリペプチドスプライスアイソフォーム（Genbank受託番号NP_001069（アイソフォームa）又はNP_542413（アイソフォームb））のプロセッシング後に産生される。

30

40

【0176】

[0195]いくつかの例では、IgSF CAMの存在又はレベルは、例えばハイブリダイゼーションアッセイ又は増幅ベースのアッセイなどのアッセイを用いて、mRNA発現のレベルで検出する。他のいくつかの例では、ICAM-1又はVCAM-1などのIgSF CAMの存在又はレベルは、例えば免疫アッセイ（例えば、ELISA又は免疫電気化学発光アッセイ）又は免疫組織化学的アッセイを使用して、タンパク質発現のレベルで検出する。組織サンプル、生検、血清、血漿、唾液、尿、又は大便などのサンプル中のICAM-1及び／又はVCAM-1の存在又はレベルを決定するための適切な抗体及び／又はELISAキットは、例えば、Invitrogen（Camarillo, CA）、Santa Cruz Biotechnology, Inc.（Santa Cru

50

z, CA)、及び／又はAbcam Inc. (Cambridge, MA) から入手可能である。

【0177】

K. 遺伝子マーカー

サンプル中の1つ又は複数の遺伝子マーカーの対立遺伝子バリエーションの有無の決定も本発明で有用である。遺伝子マーカーの非限定例としては、表A～Eに記載されている遺伝子のいずれか又はそれらの組み合わせが挙げられる。ATG16L1、STAT3、NKX2-3及びECM1遺伝子のうちの1つ又は複数において少なくとも1つの一塩基多型(SNP)の有無が検出されることが好ましい。

【0178】

[0197]表Aは、IBD、UC及びCD遺伝子の例示的リストを提供し、これらの遺伝子における1つ又は複数の対立遺伝子バリエーション(例えばSNP)の有無に関するジェノタイピングは本発明の診断方法において有用である。表Bは、さらなる本発明の診断方法に従ってジェノタイピングされ得る遺伝子マーカーの例及び対応するSNPを提供する。表C～Eは、本明細書に記載の診断方法に従ってジェノタイピングされ得るさらなるIBD、UC及びCDの遺伝子マーカーの例及び対応するSNPを提供する。

【0179】

表A：IBD、CD及びUC遺伝子

【表A】

IBD 遺伝子 (CD 及び UC)	結腸 IBD 遺伝子	UC 遺伝子	CD 遺伝子
IL23R	HLA 領域	ECM1	NOD2
IL12B/p40		IL10	ATG16L1
JAK2		IFNg	IRGM
STAT3		IL22	NLRP3
NKX2.3		IL26	5p13/PTGER4
3p21/MST1		OTUD3	PTPN2
CCNY		PLA2G2E	TNFSF15(TL1A)
IL18RAP		ARPC2	IBD5/5q31
LYRM4			ZNF365
CDKAL4			PTPN22
TNFRSF6B			CCR6
PSMG1			LRRK2
			ICOSLG
			ITLN1
			ORMDL3

【0180】

表B：IBD、CD及びUC遺伝子並びにSNP

【表 B - 1】

遺伝子	SNP
NOD2/CARD15	rs2066847
IL23R	rs11465804
ATG16L1	rs3828309
MST1	rs3197999
PTGER4	rs4613763
IRGM	rs11747270
TNFSF15	rs4263839
ZNF365	rs10995271
NKX2-3	rs11190140
PTPN2	rs2542151
PTPN22	rs2476601
ITLN1	rs2274910
IL12B	rs10045431
CDKAL1	rs6908425
CCR6	rs2301436
JAK2	rs10758669
C11orf30	rs7927894
LRRK2, MUC19	rs11175593
ORMDL3	rs2872507
STAT3	rs744166
ICOSLG	rs762421
GCKR	rs780094
BTNL2, SLC26A3, HLA-DRB1, HLA-DQA1	rs3763313

10

20

30

【表 B - 2】

遺伝子	SNP
PUS10	rs13003464
CCL2, CCL7	rs991804
LYRM4	rs12529198
SLC22A23	rs17309827
IL18RAP	rs917997
IL12RB2	rs7546245
IL12RB1	rs374326
CD3D	rs3212262
CD3G	rs3212262
CD247	rs704853
JUN	rs6661505
CD3E	rs7937334
IL18R1	rs1035127
CCR5	
MAPK14	rs2237093
IL18	rs11214108
IFNG	rs10878698
MAP2K6	rs2905443
STAT4	rs1584945
IL12A	rs6800657
TYK2	rs12720356
ETV5	rs9867846
MAPK8	rs17697885

10

20

30

【 0 1 8 1 】

表 C : C D 遺伝子及び S N P

【表 C - 1】

遺伝子	SNP
NOD2 (R702W)	rs2066844
NOD2 (G908R)	rs2066845
NOD2 (3020insC)	rs5743293
ATG16L1 (T300A)	rs2241880
ATG16L1	rs3828309
IRGM	rs13361189
IRGM	rs4958847
IRGM	rs1000113
IRGM	rs11747270
TL1A/TNFSF15	rs6478109
TL1A/TNFSF15	rs6478108
TL1A/TNFSF15	rs4263839
PTN22	rs2476601
CCR6	rs1456893
CCR6	rs2301436
5p13/PTGER4	rs1373692
5p13/PTGER4	rs4495224
5p13/PTGER4	rs7720838
5p13/PTGER4	rs4613763
ITLN1	rs2274910
ITLN1	rs9286879
ITLN1	rs11584383
IBD5/5q31	rs2188962
IBD5/5q31	rs252057
IBD5/5q31	rs10067603
GCKR	rs780094
TNFRSF6B	rs1736135

10

20

30

【表 C - 2】

遺伝子	SNP
ZNF365	rs224136
ZNF365	rs10995271
C11orf30	rs7927894
LRRK2, MUC19	rs1175593
DLG5	rs2165047
IL-27	rs8049439
TLR2	rs4696480
TLR2	rs3804099
TLR2	rs3804100
TLR2	rs5743704
TLR2	rs2405432
TLR4 (D299G)	rs4986790
TLR4 (T399I)	rs4986791
TLR4 (S360N)	rs4987233
TLR9	rs187084
TLR9	rs352140
NFC4	rs4821544
KIF21B	rs11584383
IKZF1	rs1456893
C11orf30	rs7927894
CCL2, CCL7	rs991804
ICOSLG	rs762421
TNFAIP3	rs7753394
FLJ45139	rs2836754
PTGER4	rs4613763

10

20

30

【 0 1 8 2 】

表 D : U C 遺伝子及び S N P

【表 D】

遺伝子	SNP
ECM1	rs7511649
ECM1 (T130M)	rs3737240
ECM1 (G290S)	rs13294
GLI1 (G933D)	rs2228224
GLI1 (Q1100E)	rs2228226
MDR1 (3435C>T)	rs1045642
MDR1 (A893S/T)	rs2032582
MAGI2	rs6962966
MAGI2	rs2160322
IL26	rs12815372
IFNG, IL26	rs1558744
IFNG, IL26	rs971545
IL26	rs2870946
ARPC2	rs12612347
IL10, IL19	rs3024493
IL10, IL19	rs3024505
IL23R	rs1004819
IL23R	rs2201841
IL23R	rs11209026
IL23R	rs11465804
IL23R	rs10889677
BTLN2	rs9268480
HLA-DRB1	rs660895
MEP1	rs6920863
MEP1	rs2274658
MEP1	rs4714952
MEP1	rs1059276
PUS10	rs13003464
PUS10	rs6706689
RNF186	rs3806308
RNF186	rs1317209
RNF186	rs6426833
FCGR2A, C	rs10800309
CEP72	rs4957048
DLD, LAMB1	rs4598195
CAPN10, KIF1A	rs4676410

10

20

30

40

【 0 1 8 3 】

表 E : I B D 遺伝子及び S N P

【表 E】

遺伝子	SNP
IL23R (R381Q)	rs11209026
IL23R	rs11805303
IL23R	rs7517847
IL12B/p40	rs1368438
IL12B/p40	rs10045431
IL12B/p40	rs6556416
IL12B/p40	rs6887695
IL12B/p40	rs3212227
STAT3	rs744166
JAK2	rs10974914
JAK2	rs10758669
NKX2-3	rs6584283
NKX2-3	rs10883365
NKX2-3	rs11190140
IL18RAP	rs917997
LYRM4	rs12529198
CDKAL1	rs6908425
MAGI2	rs2160322
TNFRSF6B	rs2160322
TNFRSF6B	rs2315008
TNFRSF6B	rs4809330
PSMG1	rs2094871
PSMG1	rs2836878
PTPN2	rs2542151
MST1/3p21	rs9858542
MST1/3p21	rs3197999
SLC22A23	rs17309827
MHC	rs660895
XBP1	rs35873774
ICOSLG1	rs762421
BTLN2	rs3763313
BTLN2	rs2395185
BTLN2	rs9268480
ATG5	rs7746082
CUL2, CREM	rs17582416
CARD9	rs4077515
ORMDL3	rs2872507
ORMDL3	rs2305480

【0184】

[0198] 本発明で有用なさらなる SNP には、例えば、rs2188962、rs9286879、rs11584383、rs7746082、rs1456893、rs1551398、rs17582416、rs3764147、rs1736135、rs4807569、rs7758080 及び rs8098673 が含まれる。例えば、Bar

10

20

30

40

50

rettら, Nat. Genet., 40:955~62 (2008) を参照されたい。

【0185】

1. ATG16L1

ATG16L1 (自己貪食関連16様1としても知られている。)は、細胞質成分をリソソームに送達する細胞内過程である自己貪食と呼ばれる過程に関与するタンパク質である。自己貪食とは、細胞成分を再利用するために細胞によって使用される過程である。また、自己貪食過程は炎症反応にも関与しており、細菌による免疫系の破壊を促進する。ATG16L1タンパク質は、大きなタンパク質複合体のWDリピート含有成分であり、自食胞形成の間、自己貪食性の単離膜と会合する (Mizushimaら, Journal of Cell Science 116 (9):1679~1688 (2003); Hampeら, Nature Genetics 39:207~211 (2006))。ATG16L1はクローン病に関連付けられている (Riouxら, Nature Genetics 39 (5):596~604 (2007))。例えば、Marquezら, Inflamm. Bowel Disease 15 (11):1697~1704 (2009); Mizushimaら, J. Cell Science 116:1679~1688 (2003); Zhengら, DNA Sequence: The Journal of DNA Sequencing and Mapping 15 (4):303~5 (2004) も参照されたい。

10

【0186】

[0200] ATG16L1 遺伝子中の SNP などの対立遺伝子バリエーションの存在又は非存在の決定は本発明において特に有用である。本明細書において、用語「ATG16L1 バリエーション」又はその変形には、野生型の ATG16L1 遺伝子と比較して1つ又は複数の変化を含有する ATG16L1 遺伝子のヌクレオチド配列、又は野生型の ATG16L1 ポリペプチド配列と比較して1つ又は複数の変化を含有する ATG16L1 ポリペプチドのアミノ酸配列が含まれる。ATG16L1 (自己貪食関連16様1としても知られている。)はヒト染色体2に位置決定されている。

20

【0187】

[0201] ATG16L1 の遺伝子位置情報は、例えば GeneID:55054 に記載されている。ヒト ATG16L1 の mRNA (コード) 及びポリペプチド配列は、例えば、NM_017974.3 又は NM_030803.6 及び NP_060444.3 又は NP_110430.5 にそれぞれ記載されている。さらに、ATG16L1 が含まれるヒト染色体2 (2q37.1) の完全配列、GRCh37 一次参照アセンブリーは、例えば GenBank 受託番号 NT_005120.16 に記載されている。さらに、他の種の ATG16L1 の配列は GenBank データベースに見出すことができる。

30

【0188】

[0202] 本発明の方法での使用を見出す rs2241880 SNP は、NP_060444.3 のヌクレオチド位置281位でのスレオニンからアラニンへの変化に対応する A から G への転移として NM_017974.3 の1098位に位置するか、又は NP_110430.5 の300位でのスレオニンからアラニンへの変化に対応する A から G への転移として NM_030803.6 の1155位に位置する。ATG16L1 の rs2241880 SNP 及び他のバリエーションの存在は、例えば対立遺伝子識別アッセイ又は配列解析によって検出することができる。

40

【0189】

2. ECM1

細胞外マトリックスタンパク質としても知られる ECM1 は、小腸及び大腸で発現する糖タンパク質である。ECM1 は基底膜と相互作用し、マトリックスメタロプロテアーゼ9を阻害する。ECM1 はまた、免疫応答の重要な制御因子である NF- κ B のシグナル伝達を強く活性化する。

【0190】

[0204] ECM1 遺伝子の SNP などの対立遺伝子バリエーションの有無の決定は本発明で特

50

に有用である。本明細書において、用語「ECM1バリエーション」又はその変形には、野生型ECM1遺伝子と比較して1つ又は複数の変化を含有するECM1遺伝子のヌクレオチド配列、又は野生型ECM1ポリペプチド配列と比較して1つ又は複数の変化を含有するECM1ポリペプチドのアミノ酸配列が含まれる。ECM1遺伝子は、ヒト染色体1q21上に位置決定されている(Fisherら, Nature Genetics., 40: 710~712 (2008))。

【0191】

[0205] ECM1の遺伝子位置情報は、例えば遺伝子ID: 1893に記載されている。ヒトECM1のmRNA(コード)配列は、例えばNM_001202858. 1、NM_022664. 2及びNM_004425. 3に記載されている。それぞれのポリペプチド配列は、NP_001189787. 1、NP_004416. 2及びNP_073155. 2に記載されている。さらに、ECM1が含まれるヒト染色体1(1q21)の完全配列、GRCh37. p5の一次参照アセンブリーは、例えばGenBank受託番号NC_000001. 10に記載されている。さらに、他の種のECM1の配列はGenBankデータベースに見出すことができる。

10

【0192】

[0206] ECM1遺伝子の近く及びECM1遺伝子内の種々のSNPが文献に記載されている。例えば潰瘍性大腸炎に対する非同義SNPスキャンにおいて同定されたECM1のSNPは、UCに関係付けられている(Fisherら, Nature Genetics., 40: 710~712 (2008))。ECM1のSNPの非限定例としては、rs3737240、rs7511649、rs13294、rs12724450、rs11205386、rs11205387、rs9661040、rs875514及びrs11810419が含まれる。ECM1のrs3737240 SNP及び他のバリエーションの存在は、例えば対立遺伝子識別アッセイ又は配列解析によって検出することができる。

20

【0193】

3. NKX2-3

ホメオドメイン含有転写因子のNKXファミリーのメンバーであるNKX2-3は、消化管、脾臓及び消化管に付随するリンパ組織の発達に関連付けられている。

【0194】

[0208] NKX2-3遺伝子のSNPなどの対立遺伝子バリエーションの有無の決定は本発明で特に有用である。本明細書において、用語「NKX2-3バリエーション」又はその変形には、野生型NKX2-3遺伝子と比較して1つ又は複数の変化を含有するNKX2-3遺伝子のヌクレオチド配列、又は野生型NKX2-3ポリペプチド配列と比較して1つ又は複数の変化を含有するNKX2-3ポリペプチドのアミノ酸配列が含まれる。NKX2-3遺伝子は、ヒト染色体10q24. 2上に位置決定されている。

30

【0195】

[0209] NKX2-3の遺伝子位置情報は、例えばGeneID: 159296に記載されている。ヒトNKX2-3のmRNA(コード)配列及び対応するポリペプチド配列は、例えばNM_145285. 2及びNP_660328. 2にそれぞれ記載されている。さらに、NKX2-3が含まれるヒト染色体10(10q24. 2)の完全配列、GRCh37. p5一次参照アセンブリーは、例えばGenBank受託番号NC_000001. 10に記載されている。さらに、他の種のNKX2-3の配列はGenBankデータベースに見出すことができる。

40

【0196】

[0210] NKX2-3遺伝子の近く及びNKX2-3遺伝子内の種々のSNPが文献に記載されている。例えばNKX2-3のSNP rs10883365は、日本人集団におけるUC及びCD(Araiら, Hum. Immunol., 72: 587~91 (2011))及び西欧人集団におけるCDの罹患率(Meggyesiら, World J. Gastroenterol., 16: 5233~5240 (2010))と関係付けられ

50

ている。NKX2-3のSNPの非限定例としては、rs10883365、rs6584283、rs11190140、rs888208、rs11596008、rs7911680、rs7908704、rs41290504、rs7893840、rs884144、rs10082511、rs60505509及びrs59754112が挙げられる。NKX2-3のSNP rs10883365及び他のバリエーションの存在は、例えば対立遺伝子識別アッセイ又は配列解析によって検出することができる。

【0197】

4. STAT3

STATタンパク質ファミリーのメンバーであるシグナル伝達性転写因子3 (STAT3) は、細胞成長、アポトーシス、細胞運動及びサイトカイン産生などの多くの細胞内過程に参与する種々の遺伝子の発現を制御する転写因子である。サイトカイン及び成長因子に応答して、STAT3はJAKキナーゼによって活性化され、核に移行して転写活性化因子として作用する。研究によって、STAT3が炎症性腸疾患の病変形成を含む種々の免疫異常において重要な役割を果たすことが実証された (例えば、Sugimoto, World J. Gastroenterol., 14:5110~5114, (2008) を参照)。

【0198】

[0212] STAT3の遺伝子位置情報は、例えばGeneID:6774に記載されている。ヒトSTAT3のmRNA (コード) 配列及び対応するポリペプチド配列は、例えばNM_139276.2、NM_003150.3及びNM_213662.1、並びにNP_644805.1、NP_003141.2及びNP_998827.1にそれぞれ記載されている。さらに、STAT3が含まれるヒト染色体17q21.31の完全配列、GRCh37.p5一次アセンブリーは、例えばGenBank受託番号NC_000017.10に記載されている。さらに、他の種のSTAT3の配列はGenBankデータベースに見出すことができる。

【0199】

[0213] STAT3遺伝子の近く及びSTAT3遺伝子内の種々のSNPが文献に記載されている (Franker, Nature Genetics, 40:713~715 (2008); Sato, J. Clin. Immunol., 29:815~25 (2009); Ferguson, Mutat. Res., 690:108~15 (2010); Bolland, Dig. Dis., 28:590~5 (2010) を参照)。STAT3のSNPの非限定例としては、rs744166、rs11547455、rs957970、rs2291281、rs1064111、rs10775、rs9912773、rs2306580、rs1905340、rs12947808、rs12949918、rs34460718、rs1064110、rs1053004、rs3736164、rs7211777、rs1064118、rs12721576、rs35499754、rs4796644、rs3736163、rs3809758、rs3744483、rs17885291、rs6503698及びrs77479856が挙げられる。STAT3のSNP rs744166及び他のバリエーションの存在は、例えば対立遺伝子識別アッセイ又は配列解析によって検出することができる。

【0200】

5. NOD2/CARD15

NOD2/CARD15遺伝子におけるSNPなどの対立遺伝子バリエーションの有無の決定も本発明で有用であり得る。本明細書において、用語「NOD2/CARD15バリエーション」又は「NOD2バリエーション」には、野生型NOD2遺伝子と比較して1つ若しくは複数の変化を含有するNOD2遺伝子のヌクレオチド配列、又は野生型NOD2ポリペプチド配列と比較して1つ若しくは複数の変化を含有するNOD2ポリペプチドのアミノ酸配列が含まれる。CARD15としても知られるNOD2は、染色体16上のIBD1座位に位置決定されており、ポジショナルクローニング (Hugot, Nature, 411:599~603 (2001)) 及び位置的候補遺伝子戦略 (Ogura, Nat

10

20

30

40

50

ure, 411:603~606 (2001); Hampeら, Lancet, 357:1925~1928 (2001))によって同定されている。IBD1座位は、炎症性腸疾患に対して高い多点連鎖スコア(MLS)を有する(16q12のマーカーD16S411においてMLS=5.7)。例えば、Chorら, Inflamm. Bowel Dis., 3:186~190 (1997); Akolkarら, Am. J. Gastroenterol., 96:1127~1132 (2001); Ohmenら, Hum. Mol. Genet., 5:1679~1683 (1996); Parkesら, Lancet, 348:1588 (1996); Cavanaughら, Ann. Hum. Genet., 62:291~8 (1998); Brantら, Gastroenterology, 115:1056~1061 (1998); Curranら, Gastroenterology, 115:1066~1071 (1998); Hampeら, Am. J. Hum. Genet., 64:808~816 (1999); 及びAnnesseら, Eur. J. Hum. Genet., 7:567~573 (1999)を参照されたい。

10

【0201】

[0215]ヒトNOD2のmRNA(コード)及びポリペプチド配列は、例えばGenBank受託番号NM_022162及びNP_071445にそれぞれ記載されている。さらに、NOD2が含まれるヒト染色体16のクローンRP11-327F22の完全配列は、例えばGenBank受託番号AC007728に記載されている。さらに、他の種のNOD2の配列はGenBankデータベースに見出すことができる。

【0202】

[0216]NOD2タンパク質は、NF-κB(NF-κB)を活性化することができるアミノ末端カスパーゼ動員ドメイン(CARD)及びいくつかのカルボキシ末端ロイシンリッチリピートドメインを含有する(Oguraら, J. Biol. Chem., 276:4812~4818 (2001))。NOD2は、アポトーシス制御因子Apaf-1/CED-4及び植物病害抵抗性遺伝子産物種と構造的相同性を有する(Oguraら(上記))。植物病害抵抗性遺伝子産物と同様に、NOD2は、アミノ末端エフェクタードメイン、ヌクレオチド結合ドメイン及びロイシンリッチリピート(LRR)を有する。野生型NOD2は、核内因子NF-κBを活性化し、NF-κBを細菌性リポ多糖(LPS; Oguraら(上記); Inoharaら, J. Biol. Chem., 276:2551~2554 (2001))に対して応答性とする。NOD2はLPSに対する細胞間受容体として機能することができ、NOD2のロイシンリッチリピートは応答性に必要とされる。

20

30

【0203】

[0217]NOD2コード領域中の3つの一塩基多型での変異が文献に記載されている。R702W(「SNP8」)、G908R(「SNP12」)及び1007fs(「SNP13」)と命名されたこれらの3つのSNPはNOD2遺伝子のカルボキシ末端領域に位置する(Hugotら(上記))。SNP8、SNP12及びSNP13並びに本発明での使用に適したさらなるNOD2遺伝子のSNPのさらなる解説は、例えば、米国特許第6835815号、第6858391号及び第7592437号、並びに米国特許公開第20030190639号、第2005005402号及び第20070072180号に見出すことができる。

40

【0204】

[0218]一部の実施形態では、NOD2バリエントは、NOD2座位のコード領域、例えばカルボキシ末端部分にあるいくつかのロイシンリッチリピートをコードする領域内に位置する。NOD2のロイシンリッチリピート領域に位置するそうしたNOD2バリエントには、これらに限定されないが、R702W(「SNP8」)及びG908R(「SNP12」)が含まれる。本発明で有用なNOD2バリエントは、野生型NOD2ポリペプチドによるNF-κBの活性化と比較して、NF-κBを活性化する能力が低下したNOD2ポリペプチドもコードし得る。非限定例として、NOD2バリエント1007fs(「SNP13」)は、LPS刺激に応答してNF-κBを誘導する能力が低下

50

した、短縮されたNOD2ポリペプチドをもたらす（Oguraら，Nature，411：603～606（2001））。

【0205】

[0219]本発明で有用なNOD2バリエントは、例えばR702W、G908R又は1007fsであり得る。R702W、G908R及び1007fsはNOD2のコード領域内に位置する。一実施形態では、本発明の方法はR702W NOD2バリエントを用いて実施される。本明細書において、用語「R702W」には、NOD2タンパク質のアミノ酸702をコードする三つ組内に生じる、NOD2遺伝子のエクソン4内の一塩基多型が含まれる。野生型NOD2対立遺伝子は、アミノ酸702でアルギニンをコードする三つ組内に存在するシトシン（c）残基をAC007728配列の138991位に含有する。R702W NOD2バリエントは、チミン（t）残基をAC007728配列の138991位に含有し、アルギニン（R）からトリプトファン（W）への置換をNOD2タンパク質のアミノ酸702でもたらす。したがって、このNOD2バリエントは「R702W」又は「702W」で表示され、Hugotら（上記）の以前の番号付けシステムに基づいて「R675W」で表示することもできる。さらに、R702Wバリエントは、「SNP8」対立遺伝子又はSNP8での「2」対立遺伝子としても知られる。R702W又はSNP8に対するNCBIのSNP ID番号はrs2066844である。R702W NOD2バリエント及び他のNOD2バリエントの存在は、好都合にも、例えば対立遺伝子識別アッセイ又は配列解析によって検出することができる。

【0206】

[0220]本発明の方法は、G908R NOD2バリエントを用いて実施することもできる。本明細書において、用語「G908R」には、NOD2タンパク質のアミノ酸908をコードする三つ組内に生じる、NOD2遺伝子のエクソン8内の一塩基多型が含まれる。アミノ酸908はNOD2遺伝子のロイシンリッチリピート領域内に位置する。野生型NOD2対立遺伝子は、アミノ酸908でグリシンをコードする三つ組内に存在するグアニン（g）残基をAC007728配列の128377位に含有する。G908R NOD2バリエントは、シトシン（c）残基をAC007728配列の128377位に含有し、グリシン（G）からアルギニン（R）への置換をNOD2タンパク質のアミノ酸908でもたらす。したがって、このNOD2バリエントは「G908R」又は「908R」で表示され、Hugotら（上記）の以前の番号付けシステムに基づいて、「G881R」で表示することもできる。さらにG908Rバリエントは、「SNP12」対立遺伝子又はSNP12での「2」対立遺伝子としても知られる。G908R SNP12に対するNCBIのSNP ID番号はrs2066845である。

【0207】

[0221]本発明の方法は、1007fs NOD2バリエントを用いて実施することもできる。このバリエントは、NOD2タンパク質の10番目のロイシンリッチリピートにフレームシフトをもたらし、後に未成熟終止コドンが続く単一ヌクレオチドの挿入である。生成されたNOD2タンパク質短縮体は、細菌性リポ多糖に応答するNF-κBの活性化を妨げるように思われる（Oguraら（上記））。本明細書において、用語「1007fs」には、NOD2タンパク質のアミノ酸1007をコードする三つ組に生じる、NOD2遺伝子のエクソン11内の一塩基多型が含まれる。1007fsバリエントは、AC007728配列の121139位に付加されたシトシンを含有し、アミノ酸1007でフレームシフト変異をもたらす。したがって、このNOD2バリエントは、「1007fs」で表示され、Hugotら（上記）の以前の番号付けシステムに基づいて「3020insC」又は「980fs」で表示することもできる。さらに、1007fs NOD2バリエントは、「SNP13」対立遺伝子又はSNP13での「2」対立遺伝子としても知られる。1007fs又はSNP13に対するNCBIのSNP ID番号はrs2066847である。

【0208】

[0222]特定のNOD2バリエント対立遺伝子又は他の多型対立遺伝子は、好都合にも、

例えば P E B i o s y s t e m s (F o s t e r C i t y , C A) から得られる市販の参照 DNA を使用して、ヒト多型研究センター (C E P H) 参照個体 (例えば、1347-02 (D i b ら , N a t u r e , 380:152~154 (1996)) と称される個体) と比較して定義できることを当業者は認識している。さらに SNP に関する具体的な情報は、国立バイオテクノロジー情報センター (N C B I) の d b S N P から得ることができる。

【0209】

[0223] N O D 2 バリエントは、N O D 2 座位の非コード領域に位置し得る。非コード領域には、例えばイントロン配列並びに 5' 及び 3' 非翻訳配列が含まれる。N O D 2 遺伝子の非コード領域に位置する N O D 2 バリエント対立遺伝子の非限定例は、J W 1 バリエントであり、これは、S u g i m u r a ら , A m . J . H u m . G e n e t . , 72:509~518 (2003) 及び米国特許出願公開第 20070072180 号に記載されている。N O D 2 遺伝子の 3' 非翻訳領域に位置する N O D 2 バリエント対立遺伝子の例としては、これらに限定されないが、J W 15 及び J W 16 バリエント対立遺伝子が挙げられ、これらは米国特許出願公開第 20070072180 号に記載されている。N O D 2 遺伝子の 5' 非翻訳領域 (例えばプロモーター領域) に位置する N O D 2 バリエント対立遺伝子の例としては、これらに限定されないが、J W 17 及び J W 18 バリエント対立遺伝子が挙げられ、これらは米国特許出願公開第 20070072180 号に記載されている。

10

【0210】

[0224] 本明細書において、用語「J W 1 バリエント対立遺伝子」には、N O D 2 遺伝子の介在配列 8 (イントロン 8) のヌクレオチド 158 での遺伝的変異が含まれる。A C 007728 配列に関しては、J W 1 バリエント対立遺伝子は 128143 位に位置する。イントロン 8 のヌクレオチド 158 での遺伝的変異は、これらに限定されないが、単一のヌクレオチド置換、多数のヌクレオチド置換、又は 1 つ若しくは複数のヌクレオチドの欠失若しくは挿入であり得る。イントロン 8 の野生型配列は 158 位にシトシンを有する。非限定例としては、J W 1 バリエント対立遺伝子は、シトシン (c) からアデニン (a) 、シトシン (c) からグアニン (g) 、又はシトシン (c) からチミン (t) の置換をイントロン 8 のヌクレオチド 158 に有し得る。一実施形態では、J W 1 バリエント対立遺伝子は、N O D 2 イントロン 8 のヌクレオチド 158 でのシトシン (c) からチミン (t) への変化である。

20

30

【0211】

[0225] 用語「J W 15 バリエント対立遺伝子」には、A C 007728 配列のヌクレオチド位置 118790 位での、N O D 2 の 3' 非翻訳領域の遺伝的変異が含まれる。ヌクレオチド 118790 での遺伝的変異は、これらに限定されないが、単一のヌクレオチド置換、多数のヌクレオチド置換、又は 1 つ若しくは複数のヌクレオチドの欠失若しくは挿入であり得る。野生型配列は 118790 位にアデニン (a) を有する。非限定例としては、J W 15 バリエント対立遺伝子は、アデニン (a) からシトシン (C) 、アデニン (a) からグアニン (g) 、又はアデニン (a) からチミン (t) の置換をヌクレオチド 118790 に有し得る。一実施形態では、J W 15 バリエント対立遺伝子は、ヌクレオチド 118790 でのアデニン (a) からシトシン (c) への変化である。

40

【0212】

[0226] 本明細書において、用語「J W 16 バリエント対立遺伝子」には、A C 007728 配列のヌクレオチド位置 118031 位での、N O D 2 の 3' 非翻訳領域の遺伝的変異が含まれる。ヌクレオチド 118031 での遺伝的変異は、これらに限定されないが、単一のヌクレオチド置換、多数のヌクレオチド置換、又は 1 つ若しくは複数のヌクレオチドの欠失若しくは挿入であり得る。野生型配列は 118031 位にグアニン (g) を有する。非限定例としては、J W 16 バリエント対立遺伝子は、グアニン (g) からシトシン (C) 、グアニン (g) からアデニン、(a) 又はグアニン (g) からチミン (t) の置換をヌクレオチド 118031 に有し得る。一実施形態では、J W 16 バリエント対立遺

50

伝子は、ヌクレオチド 1 1 8 0 3 1 でのグアニン (g) からアデニン (a) への変化である。

【0 2 1 3】

[0227]用語「J W 1 7 バリエント対立遺伝子」には、A C 0 0 7 7 2 8 配列のヌクレオチド位置 1 5 4 6 8 8 位での、N O D 2 の 5 ' 非翻訳領域の遺伝的変異が含まれる。ヌクレオチド 1 5 4 6 8 8 での遺伝的変異は、これらに限定されないが、単一ヌクレオチド置換、多数のヌクレオチド置換、又は 1 つ若しくは複数のヌクレオチドの欠失若しくは挿入であり得る。野生型配列は 1 5 4 6 8 8 位にシトシン (c) を有する。非限定例としては、J W 1 7 バリエント対立遺伝子は、シトシン (c) からグアニン (g)、シトシン (c) からアデニン (a)、又はシトシン (c) からチミン (t) の置換をヌクレオチド 1 5 4 6 8 8 に有し得る。一実施形態では、J W 1 7 バリエント対立遺伝子は、ヌクレオチド 1 5 4 6 8 8 でのシトシン (c) からチミン (t) への変化である。

10

【0 2 1 4】

[0228]本明細書において、用語「J W 1 8 バリエント対立遺伝子」には、A C 0 0 7 7 2 8 配列のヌクレオチド位置 1 5 4 4 7 1 位での、N O D 2 の 5 ' 非翻訳領域の遺伝的変異が含まれる。ヌクレオチド 1 5 4 4 7 1 での遺伝的変異は、これらに限定されないが、単一ヌクレオチド置換、多数のヌクレオチド置換、又は 1 つ若しくは複数のヌクレオチドの欠失若しくは挿入であり得る。野生型配列は 1 5 4 4 7 1 位にシトシン (c) を有する。非限定例としては、J W 1 8 バリエント対立遺伝子は、シトシン (c) からグアニン (g)、シトシン (c) からアデニン (a)、又はシトシン (c) からチミン (t) の置換をヌクレオチド 1 5 4 4 7 1 に有し得る。一実施形態では、J W 1 8 バリエント対立遺伝子は、ヌクレオチド 1 5 4 4 7 1 でのシトシン (c) からチミン (t) への変化である。

20

【0 2 1 5】

[0229]本発明の方法は、N O D 2 座位のコード領域又は非コード領域（例えば、イントロン又はプロモーター領域）に位置するこれら又は他の N O D 2 バリエント対立遺伝子を用いて実施できることが理解される。本発明の方法は、これらに限定されないが S N P 8、S N P 1 2 及び S N P 1 3 対立遺伝子並びに他のコード及び非コード領域のバリエントを含む、1、2、3、4 又はそれより多くの N O D 2 バリエントの存在の決定を含むことができることがさらに理解される。

【0 2 1 6】

L. 他の診断マーカー

本発明での使用に適したさらなる診断マーカーには、これらに限定されないが、ラクトフェリン、抗ラクトフェリン抗体、エラスターゼ、カルプロテクチン、ヘモグロビン、及びそれらの組み合わせが含まれる。

30

【0 2 1 7】

[0231]サンプル中のラクトフェリンの存在又はレベルの決定は本発明で有用であり得る。いくつかの例では、ラクトフェリンの存在又はレベルは、例えばハイブリダイゼーションアッセイ又は増幅ベースのアッセイなどのアッセイを用いて、m R N A 発現レベルで検出される。他のいくつかの例では、ラクトフェリンの存在又はレベルは、例えば免疫アッセイ（例えば E L I S A）又は免疫組織化学的アッセイを使用して、タンパク質発現レベルで検出される。C a l b i o c h e m (S a n D i e g o, C A) から入手可能な E L I S A キットを使用して、血漿、尿、気管支肺胞洗浄又は脳脊髄液サンプル中のヒトラクトフェリンを検出することができる。同様に、U. S. B i o l o g i c a l (S w a m p s c o t t, M A) から入手可能な E L I S A キットを使用して、血漿サンプル中のラクトフェリンレベルを決定することができる。同じく、T E C H L A B, I n c. (B l a c k s b u r g, V A) から入手可能な E L I S A キットを使用して、便サンプル中のラクトフェリンレベルを決定することができる。さらに、米国特許出願公開第 2 0 0 4 0 1 3 7 5 3 6 号は、便サンプル中の上昇したラクトフェリンレベルの存在を決定するための E L I S A アッセイを記載しており、米国特許出願公開第 2 0 0 4 0 0 3 3 5 3 7 号は、便、粘液又は胆汁サンプル中の内在性ラクトフェリンの濃度を決定するための E L I

40

50

S A アッセイについて記載している。一部の実施形態では、次に、例えばラクトフェリタンタンパク質又はその断片を使用して、サンプル中の抗ラクトフェリン抗体の存在又はレベルを検出することができる。

【0218】

[0232]さらに、潜血、便潜血は胃腸病の指標になることが多く、胃腸出血をモニタリングするための種々のキットが開発されている。例えばBeckman Coulterの製品であるHemoCCult SENS Aは、胃腸出血、鉄欠乏、消化性潰瘍、潰瘍性大腸炎に関する診断援助となり、一部の例では、結腸直腸癌のスクリーニングにおける診断援助となる。この特定のアッセイは、過酸化水素によってグアヤクが酸化して青色を生じることに基づく。便サンプル中の血液を検出するための類似の比色アッセイは、Helena Laboratories (Beaumont, TX) から市販品として入手可能である。ヘモグロビン又はヘム活性の存在又はレベルを決定することによって便サンプル中の潜血を検出するための他の方法は、例えば、米国特許第4277250号、第4920045号、第5081040号及び第5310684号に記載されている。

【0219】

[0233]カルプロテクチンは、すべての細胞、組織及び体液に見出されるカルシウム及び亜鉛結合タンパク質である。カルプロテクチンは、好中性顆粒球及びマクロファージの主要なタンパク質であり、これらの細胞のサイトゾル画分の全タンパク質の60%をも占める。したがって、カルプロテクチンは好中球の代謝回転の代用マーカーである。便中のカルプロテクチンの濃度は、腸管粘膜の好中球浸潤の強度及び炎症の重症度と相関する。カルプロテクチンは、少量(50~100mg)の糞便サンプルを用いてELISAにより測定することができる(例えば、Johnera, Scand J Gastroenterol., 36:291~296 (2001)を参照)。

【0220】

V I. アッセイ

I B Dを診断する、及びI B Dの診断を分類する(例えば、I C、C D、U C)ために、当分野で知られている種々のアッセイ、方法、及びキットのうちの任意のものを使用して、サンプル中の1つ又は複数のI B Dマーカーの存在又はレベルを検出又は決定することができる。

【0221】

[0235]本発明は、個体から得たサンプル中の少なくとも1つのマーカーの存在又はレベルの決定に部分的に依拠する。本明細書において、用語「少なくとも1つのマーカーの存在を検出すること」には、当業者に知られている任意の定量的又は定性的アッセイを使用することによって目的の各マーカーの存在を決定することが含まれる。いくつかの例では、特定の形質、変数又は生化学的若しくは血清学的な物質(例えば、タンパク質又は抗体)の有無を決定する定性的なアッセイは目的の各マーカーを検出するのに適する。他のいくつかの例では、R N A、タンパク質、抗体又は活性の有無を決定する定量的なアッセイは目的の各マーカーを検出するのに適する。本明細書において、用語「少なくとも1つのマーカーのレベルを検出すること」には、当業者に知られている任意の直接又は間接的な定量的アッセイを使用することによって目的の各マーカーのレベルを決定することが含まれる。いくつかの例では、例えば相対的又は絶対的なR N A、タンパク質、抗体又は活性の量を決定する定量的なアッセイは目的の各マーカーのレベルを検出するのに適する。マーカーレベルを検出するのに有用な任意のアッセイはマーカーの有無を検出するのにも有用であることを当業者であれば認識するであろう。

【0222】

[0236]本明細書において、用語「抗体」には、ポリクローナル若しくはモノクローナルかつ任意のアイソタイプのものであり得る免疫グロブリン分子の集団、又は免疫グロブリン分子の免疫学的に活性のある断片が含まれる。そのような免疫学的に活性のある断片は、抗原と特異的に結合する抗体分子の部分を構成する、重鎖及び軽鎖可変領域を含有する。例えば、F a b、F a b' 又はF (a b')₂として当技術分野で知られている免疫グ

ロブリン分子の免疫学的に活性な断片は用語「抗体」に含まれる。

【0223】

[0237]フローサイトメトリーを使用して、サンプル中の1つ又は複数のマーカーの存在又はレベルを検出することができる。ビーズに基づく免疫アッセイを含むそのようなフローサイトメトリーアッセイを使用して、例えば、カンジダ・アルビカンズ及びHIVタンパク質に対する血清抗体の検出について記載されたものと同じ様式で、抗体マーカーレベルを決定することができる（例えば、Bishop及びDavis, J. Immunol. Methods, 210:79~87 (1997); McHughら, J. Immunol. Methods, 116:213 (1989); Scillianら, Blood, 73:2041 (1989)を参照）。

10

【0224】

[0238]マーカーに特異的な組換え抗原を発現させるためのファージディスプレイも、サンプル中の1つ又は複数のマーカーの存在又はレベルを検出するために使用することができる。例えば抗体マーカーに特異的な抗原を発現するファージ粒子を、所望の場合は、抗ファージモノクローナル抗体などの抗体を使用してマルチウェルプレートに固定することができる（Felicciら, "Phage-Displayed Peptides as Tools for Characterization of Human Sera", Abelson (編), Methods in Enzymol., 267, San Diego: Academic Press, Inc. (1996)）。

【0225】

[0239]競合的及び非競合的免疫アッセイを含む種々の免疫アッセイ法も、サンプル中の1つ又は複数のマーカーの存在又はレベルを検出するために使用することができる（例えば、Self及びCook, Curr. Opin. Biotechnol., 7:60~65 (1996)を参照）。用語「免疫アッセイ」には、これらに限定されないが、酵素増幅免疫アッセイ法（EMIT）、酵素結合免疫吸着アッセイ（ELISA）、抗原捕捉ELISA、サンドイッチELISA、IgM抗体捕捉ELISA（MAC ELISA）、及び微粒子酵素免疫アッセイ（MEIA）などの酵素免疫アッセイ（EIA）、キャピラリー電気泳動免疫アッセイ（CEIA）、ラジオイムノアッセイ（RIA）、免疫放射線アッセイ（IRMA）、蛍光偏光免疫アッセイ（FPIA）、並びに化学発光アッセイ（CL）を含む方法が包含される。所望の場合は、そのような免疫アッセイは自動であり得る。また、免疫アッセイは、レーザー誘導性蛍光と併せて使用することもできる（例えば、Schmalzing及びNashabeh, Electrophoresis, 18:2184~2193 (1997); Bao, J. Chromatogr. B. Biomed. Sci., 699:463~480 (1997)を参照）。また、フローインジェクションリポソーム免疫アッセイ及びリポソーム免疫センサーなどのリポソーム免疫アッセイも、本発明における使用に適切である（例えば、Rongenら, J. Immunol. Methods, 204:105~133 (1997)を参照）。さらに、マーカー濃度の関数としてピーク速度シグナルに変換される増加した光散乱をタンパク質／抗体複合体の形成をもたらす比濁分析アッセイが、本発明における使用に適切である。比濁分析アッセイはBeckman Coulterから販売されており（Brea, CA; キット#449430）、ベーリング比濁計分析器（Behring Nephelometer Analyzer）を用いて行うことができる（Finkら, J. Clin. Chem. Clin. Biol. Chem., 27:261~276 (1989)）。

20

30

40

【0226】

[0240]抗原捕捉ELISAは、サンプル中の1つ又は複数のマーカーの存在又はレベルを検出するために有用であり得る。例えば、抗原捕捉ELISAでは、目的のマーカーに向けられた抗体を固相に結合させ、マーカーが抗体によって結合されるようにサンプルを加える。洗浄によって未結合のタンパク質を除去した後、結合したマーカーの量を、例えばラジオイムノアッセイを使用して定量することができる（例えば、Harlow及びLane, Antibodies: A Laboratory Manual, Cold

50

Spring Harbor Laboratory, New York, 1988) を参照)。また、サンドイッチ E L I S A も本発明における使用に適切な場合がある。例えば、二抗体のサンドイッチアッセイでは、第 1 の抗体を固体支持体と結合させ、目的のマーカを第 1 の抗体と結合させる。マーカは、マーカと結合する第 2 の抗体の量を測定することによって定量する。抗体は、磁気若しくはクロマトグラフィーマトリックス粒子、アッセイプレート（例えばマイクロタイターウェル）の表面、固体基質材料の小片又は膜（例えば、プラスチック、ナイロン、紙）などの種々の固体支持体上に固定することができる。抗体又は複数の抗体を固体支持体にアレイでコーティングすることによって、アッセイ条片を調製することができる。その後、この条片を試験サンプルに浸し、洗浄及び検出ステップによって素早く処理して、有色スポットなどの測定可能なシグナルを生じさせることができる。

10

【0227】

[0241] 例えばヨウ素-125 (^{125}I) で標識した二次抗体を使用したラジオイムノアッセイ (Harlow 及び Lane (上記)) も、サンプル中の 1 つ又は複数のマーカの存在又はレベルを検出するために適切である。また、化学発光性のマーカで標識した二次抗体も、本発明における使用に適切な場合がある。化学発光性の二次抗体を使用した化学発光アッセイは、マーカレベルの高感度の非放射性検出に適している。そのような二次抗体は種々の供給源、例えば Amersham Lifesciences, Inc. (Arlington Heights, IL) から商業的に入手することができる。

20

【0228】

[0242] 上述の免疫アッセイは、サンプル中の 1 つ又は複数のマーカの存在又はレベルを検出するために特に有用である。非限定的な例として、固定した好中球 E L I S A は、サンプルが A N C A に関して陽性であるかどうかを決定するため、又はサンプル中の A N C A レベルを決定するために有用である。同様に、酵母細胞壁ホスホペプチドマンナンを使用した E L I S A は、サンプルが A S C A - I g A 及び／又は A S C A - I g G に関して陽性であるかどうかを決定するため、又はサンプル中の A S C A - I g A 及び／又は A S C A - I g G レベルを決定するために有用である。Omp C タンパク質又はその断片を使用した E L I S A は、サンプルが抗 Omp C 抗体に関して陽性であるかどうかを決定するため、又はサンプル中の抗 Omp C 抗体レベルを決定するために有用である。I 2 タンパク質又はその断片を使用した E L I S A は、サンプルが抗 I 2 抗体に関して陽性であるかどうかを決定するため、又はサンプル中の抗 I 2 抗体レベルを決定するために有用である。フラジェリタンパク質（例えば、C b i r - 1 フラジェリン、F l a 2 フラジェリン、F l a X フラジェリン）又はその断片を使用した E L I S A は、サンプルが抗フラジェリン抗体に関して陽性であるかどうかを決定するため、又はサンプル中の抗フラジェリン抗体レベルを決定するために有用である。さらに、上述の免疫アッセイは、サンプル中の他のマーカの存在又はレベルを検出するために特に有用である。

30

【0229】

[0243] 抗体と目的のマーカとの特異的な免疫学的結合は直接又は間接的に検出することができる。直接標識には、抗体に付着した蛍光又はルミネセントのタグ、金属、色素、放射性核種などが含まれる。ヨウ素-125 (^{125}I) で標識した抗体を使用して、サンプル中の 1 つ又は複数のマーカのレベルを決定することができる。マーカに特異的な化学発光性抗体を使用した化学発光アッセイは、マーカレベルの高感度の非放射性検出に適している。また、蛍光色素で標識した抗体も、サンプル中の 1 つ又は複数のマーカのレベルを決定するために適している。蛍光色素の例には、これらに限定されないが、D A P I、フルオレセイン、ヘキスト 33258、R-フィコシアニン、B-フィコエリスリン、R-フィコエリスリン、ローダミン、テキサスレッド、及びリサミンが含まれる。蛍光色素に連結した二次抗体は商業的に入手することができ、例えば、ヤギ F (a b')₂ 抗ヒト I g G - F I T C は T a g o I m m u n o l o g i c a l s (B u r l i n g a m e, C A) から入手可能である。

40

50

【0230】

[0244]間接的な標識には、西洋ワサビペルオキシダーゼ (HRP)、アルカリホスファターゼ (AP)、 β -ガラクトシダーゼ、ウレアーゼなどなどの当分野で周知の種々の酵素が含まれる。西洋ワサビペルオキシダーゼ検出システムは、例えば、450 nmで検出可能な可溶性の生成物を過酸化水素の存在下で与える発色基質テトラメチルベンジジン (TMB) を用いて使用することができる。アルカリホスファターゼ検出システムは、例えば405 nmで容易に検出可能な可溶性の生成物を与える発色基質p-ニトロフェニルリン酸を用いて使用することができる。同様に、 β -ガラクトシダーゼ検出システムは、410 nmで検出可能な可溶性の生成物を与える発色基質o-ニトロフェニル- β -D-ガラクトピラノシド (ONPG) を用いて使用することができる。ウレアーゼ検出システムは、尿素-ブロモクレゾールパープル (Sigma Immunochemicals, St. Louis, MO) などの基質を用いて使用することができる。有用な酵素に連結した二次抗体はいくつかの商業的供給源から得ることができ、例えば、ヤギF(ab')₂ 抗ヒトIgG-アルカリホスファターゼは、Jackson ImmunoResearch (West Grove, PA) から購入することができる。

10

【0231】

[0245]直接又は間接的な標識からのシグナルは、例えば、発色基質からの色を検出するための分光光度計、¹²⁵Iを検出するためのガンマカウンターなどの放射線を検出するための放射線カウンター、又は特定の波長の光の存在下で蛍光を検出するための蛍光光度計を使用して分析することができる。酵素連結抗体の検出には、マーカーレベルの量の定量分析を、EMAXマイクロプレートリーダー (EMAX Microplate Reader) (Molecular Devices, Menlo Park, CA) などの分光光度計を用いて製造者の指示に従って行うことができる。所望の場合は、本明細書に記載のアッセイは、自動化する又はロボットで行うことができ、多数のサンプルからのシグナルを同時に検出することができる。

20

【0232】

[0246]定量的ウエスタンブロッティングを使用して、サンプル中の1つ又は複数のマーカーの存在又はレベルを検出又は決定することもできる。ウエスタンブロットは、濃度測定のための走査又はリン光体イメージングなどの周知の方法によって定量することができる。非限定的な例として、タンパク質サンプルを10% SDS-PAGEゲル (Laemmli) ゲル上で電気泳動させる。一次マウスモノクローナル抗体をブロットと反応させ、予備スロットブロット実験を使用して抗体結合が線形であることを確認することができる。ヤギ抗マウス西洋ワサビペルオキシダーゼカップリング抗体 (BioRad) を二次抗体として使用し、シグナル検出は、化学発光を使用して、例えばルネッサンス (Renaissance) 化学発光キット (New England Nuclear, Boston, MA) を用いて製造者の指示に従って行う。走査濃度測定計 (Molecular Dynamics, Sunnyvale, CA) を使用してブロットのオートラジオグラフを分析し、陽性対照に対して正規化する。値は、例えば実測値と陽性対照との比として報告する (濃度測定指数)。そのような方法は、例えばParraら, J. Vasc. Surg., 28:669~675 (1998) に記載のように当分野で周知である。

30

40

【0233】

[0247]あるいは、種々の免疫組織化学的アッセイ法を使用して、サンプル中の1つ又は複数のマーカーの存在又はレベルを検出又は決定することができる。用語「免疫組織化学的アッセイ」には、蛍光顕微鏡観察又は光学顕微鏡観察を使用した、目的のマーカーと反応する抗体とカップリング (すなわちコンジュゲート) させた蛍光色素又は酵素の視覚的検出を利用する方法が包含され、これらに限定されないが、直接蛍光抗体アッセイ、間接蛍光抗体 (IFA) アッセイ、抗補体免疫蛍光、アビジン-ビオチン免疫蛍光、及び免疫ペルオキシダーゼアッセイが含まれる。例えば、IFAアッセイは、サンプルがANCAに関して陽性かどうか、サンプル中のANCAのレベル、サンプルがpANCAに関して陽性かどうか、サンプル中のpANCAのレベル及び/又はANCA染色パターン (例え

50

ば、cANCA、pANCA、NSNA及び／又はSAPPAの染色パターン）を決定するために有用である。サンプル中のANCAの濃度は、例えば、エンドポイント滴定によって、又は既知の参照標準と比較した蛍光の視覚的強度を測定することによって定量することができる。

【0234】

[0248]あるいは、目的のマーカの存在又はレベルは、精製したマーカの量を検出又は定量することによって決定することができる。マーカの精製は、例えば、高圧液体クロマトグラフィー（HPLC）によって単独で、又は質量分析（例えば、MALDI/MS、MALDI-TOF/MS、SELDI-TOF/MS、タンデムMS）と組み合わせて達成することができる。また、目的のマーカの定性的又は定量的な検出は、これら

10

【0235】

[0249]複数のマーカの分析は、別々に又は1つの試験サンプルを用いて同時に実施してもよい。マーカの別々又は連続的なアッセイについて、適切な器具には、エレクシス（Elec Sys）（Roche）、アックスシム（AxSym）（Abbott）、アクセス（Access）（Beckman）、アドビア（ADVIA）（登録商標）、セントール（CENTAUR）（登録商標）（Bayer）、及びニコルスアドバンテージ（NICHOLS ADVANTAGE）（登録商標）（Nichols Institute）免疫アッセイシステムなどの臨床的実験室分析器が含まれる。好ましい器具又はタンパク質チップは、単一の表面上で複数のマーカの同時アッセイを行う。特に有用な物理的様式は、複数の異なるマーカを検出するための複数の異なる、アドレス指定可能な位置を有する表面を含む。そのような様式には、タンパク質マイクロアレイ又は「タンパク質チップ」（例えば、Ngら、J. Cell Mol. Med., 6:329~340（2002）を参照）及び特定のキャピラリー装置（例えば、米国特許第6019944号を参照）が含まれる。これらの実施形態では、異なる表面の位置の各々は、各位置で検出するための1つ又は複数のマーカを固定するための抗体を含んでいてもよい。あるいは、表面は、表面の異なる位置で固定された、検出のために1つ又は複数のマーカを固定するための抗体を含む1つ又は異なる複数の粒子（例えば、微粒子又はナノ粒子）を含んでいてもよい。

20

30

【0236】

[0250]種々の目的のマーカの存在又はレベルを検出するための上述のアッセイに加えて、ノーザン分析、逆転写酵素ポリメラーゼ連鎖反応（RT-PCR）、又はマーカコード配列の一部と相補的な核酸配列とのハイブリダイゼーションに基づく他の任意の方法（例えばスロットブロットハイブリダイゼーション）などの、ルーチン的な方法を用いるマーカ-mRNAレベルの分析も本発明の範囲内にある。適用可能なPCR増幅法は、例えば、Ausubelら、Current Protocols in Molecular Biology, John Wiley & Sons, Inc., New York（1999）、第7章及び補遺47；Theophilusら、“PCR Mutation Detection Protocols”, Humana Press（2002）；及びInnisら、PCR Protocols, San Diego, Academic Press, Inc.（1990）に記載されている。核酸ハイブリダイゼーション法は、Anderson, “Nucleic Acid Hybridization”, BIOS Scientific Publishers, 1999に記載されている。また、複数の転写された核酸配列（例えば、mRNA又はcDNA）の増幅又はハイブリダイゼーションも、マイクロアレイに配置したmRNA又はcDNA配列で行うことができる。マイクロアレイ法は、Hardiman, “Microarrays Methods and Applications: Nuts & Bolts”, DNA Press, 2003；及びBaldiら, “DNA Microarrays an

40

50

d Gene Expression: From Experiments to Data Analysis and Modeling”, Cambridge University Press, 2002に記載されている。

【0237】

[0251]多数のサンプルの効率的な処理のために、いくつかの目的のマーカーを1つの試験へと合わせてもよい。さらに、当業者には、同じ対象からの多数のサンプルを（例えば連続する時点で）試験することの価値が理解されよう。そのような一連のサンプルの試験は、マーカーレベルの経時的な変化の同定を可能にすることができる。マーカーレベルの増加又は減少、及びマーカーレベルの変化の非存在は、IBDの処置を容易にするための有用な予後診断的及び予測的情報を提供することができる。

10

【0238】

[0252]上述のマーカーの1つ又は複数を測定するためのパネルは、IBDを診断するための、IBDの確度の高い経過及び転帰を予測するため及びIBD治療に対する応答の見込みを予測するための本発明のアプローチに関する関連情報を提供するように構築することができる。そうしたパネルは、1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、25、30、35、40又はそれより多くの個々のマーカーの存在又はレベルを検出又は決定するように構築することができる。単一のマーカー又はマーカーのサブセットの分析は、当業者によって種々の臨床設定において実施することもできる。それらには、例えば、通院、応急手当、救命医療、集中治療、モニタリングユニット、入院患者、外来患者、医院、診療所及び集団検診の設定が含まれる。

20

【0239】

[0253]マーカーの分析は種々の物理的様式で行うことができる。多数の試験サンプルの処理を容易にするために、例えばマイクロタイタープレート又は自動化の使用を用いることができる。あるいは、適時の様式での処置、診断及び予後診断容易にするために、単一サンプルの様式を開発することができる。

【0240】

[0254]上述に鑑みて、IBD又はその臨床亜型に関する診断的情報を提供するため、及びIBDの進行の転帰及び経過に関する予後診断的及び予測的情報を提供するため、及びIBDの処置のための適切な治療方式の選択に関する情報を提供するための本発明の方法は、上述の周知のアッセイ又は当技術分野で知られている他のアッセイの1つ又は任意の組み合わせを使用して、（例えば、本明細書に記載するように、1つ又は複数のIBDマーカーの存在又は濃度レベルを決定することによって）実施できることを当業者は認識している。

30

【0241】

VII. ジェノタイピングの方法

種々の手段を使用して、遺伝子又は任意の本明細書に記載の他の遺伝子マーカー中の多型部位において個体をジェノタイピングして、サンプル（例えば核酸サンプル）が特定のバリエーション対立遺伝子又はハプロタイプを含有するかどうかを決定することができる。例えば、個体からの核酸の酵素増幅を好都合に使用して、続く分析のための核酸を得ることができる。また、1つ又は複数の目的の遺伝子マーカーにおける特定のバリエーション対立遺伝子又はハプロタイプの存在又は非存在は、酵素増幅を用いずに個体の核酸から直接決定することもできる。一部の好ましい実施形態では、個体はATG16L1の座位でジェノタイピングされる。他の好ましい実施形態では、個体はNKX2-3座位でジェノタイピングされる。さらに別の好ましい実施形態では、個体はSTAT3座位でジェノタイピングされる。さらなる好ましい実施形態では、個体はECM1座位でジェノタイピングされる。

40

【0242】

[0256]増幅したかどうかにかかわらず、個体からの核酸ジェノタイピングを、種々の方法のうちの任意のものを用いて行うことができる。有用な方法には、これらに限定されな

50

いが、すべて単独で又は組み合わせて使用することができる、ポリメラーゼ連鎖反応（PCR）に基づく分析、配列解析、及び電気泳動分析が含まれる。本明細書において、用語「核酸」は、例えば、ゲノムDNA、cDNA及びmRNAを含む、一本鎖又は二本鎖のDNA又はRNA分子などのポリヌクレオチドを意味する。この用語には、天然及び合成起源のいずれの核酸分子も包含され、また、ネイティブ核酸分子のセンス又はアンチセンス鎖のいずれか又は両方を表す直鎖状、環状又は分岐鎖状の立体配置の分子も包含される。そのような核酸は、精製されていなくても精製されていてもよく、また、例えばビーズ若しくはカラムマトリックスなどの合成材料に結合していてもよいことを理解されたい。

【0243】

[0257]核酸を含有する材料は、個体からルーチンの得られる。そのような材料は、核酸を調製することができる任意の生体物質である。非限定的な例として、材料は、全血、血清、血漿、唾液、頬スワブ、痰、又は核酸を含有する他の体液若しくは組織であり得る。一実施形態では、本発明の方法は、非侵襲性手段によって容易に得てゲノムDNAを調製するために使用することができる全血を用いて実施する。他の実施形態では、ジェノタイピングは、ポリメラーゼ連鎖反応（PCR）を使用した個体の核酸の増幅を含む。核酸を増幅するためのPCRの使用は当分野で周知である（例えば、Mullisら（編）、The Polymerase Chain Reaction, Birkhauser, Boston（1994）を参照）。さらに別の実施形態では、PCR増幅は、1つ又は複数の蛍光標識したプライマーを用いて行われる。さらなる実施形態では、PCR増幅は、DNA副溝結合剤を含有する1つ又は複数の標識又は非標識のプライマーを用いて行われる。

【0244】

[0258]本発明の方法において1つ若しくは複数の遺伝子又は他の遺伝子マーカーにおけるバリエーション対立遺伝子の存在又は非存在を決定するために、種々の異なるプライマーのうちの任意のものを使用して個体の核酸をPCRにより増幅することができる。遺伝子の非限定例としては、ATG16L1、STAT3、ECM1及びNKX2-3が挙げられる。例えば、PCRプライマーを使用して、ATG16L1座位の特定の領域を増幅することができる。当業者には理解されるように、ATG16L1遺伝子又は他の遺伝子マーカー中で目的の多型部位に隣接する配列に基づいて、追加のPCR分析用のプライマーを設計することができる。非限定的な例として、配列プライマーは、ATG16L1遺伝子又は他の遺伝子マーカー中の目的の多型部位の上流又は下流の配列の約15～約30個のヌクレオチドを含有することができる。そのようなプライマーは、一般に、増幅反応において安定なアニーリングステップを可能にする高い融解温度を達成するため十分なグアニン及びシトシンの含有率を有するように設計される。プライマーセレクト（Primer Select）などのいくつかのコンピュータプログラムが、PCRプライマーの設計を援助するために利用可能である。

【0245】

[0259]Applied Biosystemsから入手可能なタックマン（Taqman）（登録商標）対立遺伝子識別アッセイは、多型部位で個体をジェノタイピングし、それによってATG16L1、ECM1、STAT3、NKX2-3遺伝子のうちの1つ若しくは複数又は本明細書に記載の他の遺伝子マーカーにおける特定のバリエーション対立遺伝子又はハプロタイプの存在又は非存在を決定するために有用であり得る。タックマン（登録商標）対立遺伝子識別アッセイでは、各対立遺伝子について特異的に蛍光色素標識したプローブを構築する。プローブは、各対立遺伝子の増幅を識別するためにFAM及びVICなどの異なる蛍光レポーター色素を含有する。さらに、各プローブは、蛍光共鳴エネルギー移動によって蛍光を消光させる消光色素を一端に有する。PCRの間、各プローブは個体からの核酸中の相補的配列と特異的にアニーリングする。Taqポリメラーゼの5'ヌクレアーゼ活性を使用して、対立遺伝子とハイブリダイズするプローブのみを切断する。切断によりレポーター色素が消光色素から分離され、レポーター色素による蛍光の増加がもたらされる。したがって、PCR増幅によって生じる蛍光シグナルは、サンプル中に

どの対立遺伝子が存在するかを示す。プローブと対立遺伝子とのミスマッチは、プローブのハイブリダイゼーション及びTaqポリメラーゼによる切断のいずれの効率も低下させ、その結果、蛍光シグナルはわずかから存在しないものとなる。当業者には、例えばKuttyavinら、Nuc. Acids Research, 28:655~661(2000)に記載されているように、DNA副溝結合剤(MGB)基をDNAプローブとコンジュゲートさせることによって、対立遺伝子識別アッセイにおける改善された特異性を達成できることが理解されよう。副溝結合剤には、これらに限定されないが、ジヒドロシクロピロロインドールトリペプチド(DPI3)などの化合物が含まれる。ATG16L1、STAT3、NKX2-3及びECM1遺伝子の対立遺伝子バリエーションを検出するのに適するタックマン(登録商標)プローブは、例えばApplied Biosystemsから市販品として入手可能である。非限定例としては、カタログ番号:C__9095577__20及びC__11530586__10(ATG16L1); C__29560274__10、C__25803068__10及びC__8690827__20(ECM1); C__3140282__10(STAT3); C__3018644__10、C__31657361__10及びC__3018642__10(NKX2-3)が挙げられる。

【0246】

[0260]一部の実施形態では、ATG16L1のSNP rs2241880を検出するためのプローブは、配列: CCCAGTCCCCAGGACAAATGTGGATACATCATCCTGGTTCTGGTAAAGAAAGT; 及びCCCAGTCCCCAGGACAAATGTGGATGCTCATCCTGGTTCTGGTAAAGAAAGTを含み、両方ともCCCAGTCCCCAGGACAAATGTGGAT [A/G] CTCATCCTGGTTCTGGTAAAGAAAGTに由来し、[A/G]の表記はrs2241880 SNPの位置を表す。さらなる実施形態では、第1のプローブはVIC(商標)色素で標識され、A対立遺伝子を含む、第2のプローブはFAM(商標)で標識され、G対立遺伝子を含む。rs2241880 SNPの存在又は非存在の検出に関して、FAM/FAM(G/G)シグナル又はVIC/VIC(A/A)シグナルはホモ接合性の遺伝子型を示し、VIC/FAMシグナルはヘテロ接合性の変異遺伝子型を示すであろう。

【0247】

[0261]一部の実施形態では、STAT3のSNP rs744166を検出するためのプローブは、配列: CTGTTTGTCTATAAATTACTGTCAAGCTCGATTCCCTCAAGACATTACAG; 及びCTGTTTGTCTATAAATTACTGTCAAGGCTCGATTCCCTCAAGACATTACAGを含み、両方ともCTGTTTGTCTATAAATTACTGTCA [A/G] GCTCGATTCCCTCAAGACATTACAGに由来し、[A/G]の表記はrs744166 SNPの位置を表す。さらなる実施形態では、第1のプローブはVIC(商標)色素で標識され、A対立遺伝子を含む、第2のプローブはFAM(商標)で標識され、G対立遺伝子を含む。rs744166 SNPの存在又は非存在の検出に関して、FAM/FAM(G/G)シグナル又はVIC/VIC(A/A)シグナルはホモ接合性の遺伝子型を示し、VIC/FAMシグナルはヘテロ接合性の変異遺伝子型を示すであろう。

【0248】

[0262]一部の実施形態では、ECM1のSNP rs3737240を検出するためのプローブは、配列: CCCACTGTTTTCCCCATTCCAGGAACGCCAGCTCCATTGTTTGGGACCAAGAGC; 及びCCACTGTTTTCCCCATTCCAGGAATGCCAGCTCCATTGTTTGGGACCAAGAGCを含み、両方ともCCCACTGTTTTCCCCATTCCAGGA [C/T] GCCAGCTCCATTGTTTGGGACCAAGAGCに由来し、[C/T]の表記はrs3737240 SNPの位置を表す。さらなる実施形態では、第1のプローブはVIC(商標)色素で標識され、C対立遺伝子を含む、第2のプローブはFAM(商標)で標識され、T対立遺伝子を含む。rs3737240 SNPの存在又は非存在の検出に関して、FAM/

FAM (T/T) シグナル又はVIC/VIC (C/C) シグナルはホモ接合性の遺伝子型を示し、VIC/FAMシグナルはヘテロ接合性の変異遺伝子型を示すであろう。

【0249】

[0263]一部の実施形態では、ECM1のSNP rs13294を検出するためのプローブは、配列：CCGGGACATCTTGACCATTTGACATCAGTCGAGTCACCCCCCAACCTTCATGGG；及びCCGGGACATCTTGACCATTTGACATCAGTCGAGTCACCCCCCAACCTTCATGGGを含み、両方ともCCGGGACATCTTGACCATTTGACATC [A/G] GTCGAGTCACCCCCCAACCTTCATGGGに由来し、[A/G]の表記はrs13294 SNPの位置を表す。さらなる実施形態では、第1のプローブはVIC (商標) 色素で標識され、A対立遺伝子を含有し、第2のプローブはFAM (商標) で標識され、G対立遺伝子を含有する。rs13294 SNPの存在又は非存在の検出に関して、FAM/FAM (G/G) シグナル又はVIC/VIC (A/A) シグナルはホモ接合性の遺伝子型を示し、VIC/FAMシグナルはヘテロ接合性の変異遺伝子型を示すであろう。

10

【0250】

[0264]一部の実施形態では、NKX2-3のSNP rs10883365を検出するためのプローブは、配列：TTTCGTTCTTCAGACGGTTTGAAGGTATTTGTGTCACAACGTGACCCCCCGGGGA；及びTTTCGTTCTTCAGACGGTTTGAAGGT [A/G] TTTGTGTCACAACGTGACCCCCCGGGGAを含み、両方ともTTTCGTTCTTCAGACGGTTTGAAGGT [A/G] TTTGTGTCACAACGTGACCCCCCGGGGAに由来し、[A/G]の表記はrs10883365 SNPの位置を表す。さらなる実施形態では、第1のプローブはVIC (商標) 色素で標識され、A対立遺伝子を含有し、第2のプローブはFAM (商標) で標識され、G対立遺伝子を含有する。rs10883365 SNPの存在又は非存在の検出に関して、FAM/FAM (G/G) シグナル又はVIC/VIC (A/A) シグナルはホモ接合性の遺伝子型を示し、VIC/FAMシグナルはヘテロ接合性の変異遺伝子型を示すであろう。

20

【0251】

[0265]一部の実施形態では、NKX2-3のSNP rs6584283を検出するためのプローブは、配列：CTGCAGAGCGTCTGTGGGCGTGTAATCGGCGATCAGGCGCTGGAGGGGCGCT；及びCTGCAGAGCGTCTGTGGGCGTGTAATCGGCGCTGGAGGGGCGCTを含み、両方ともCTGCAGAGCGTCTGTGGGCGTGTAAT [C/T] GGC GATCAGGCGCTGGAGGGGCGCTに由来し、[C/T]の表記はrs6584283 SNPの位置を表す。さらなる実施形態では、第1のプローブはVIC (商標) 色素で標識され、C対立遺伝子を含有し、第2のプローブはFAM (商標) で標識され、T対立遺伝子を含有する。rs6584283 SNPの存在又は非存在の検出に関して、FAM/FAM (T/T) シグナル又はVIC/VIC (A/A) シグナルはホモ接合性の遺伝子型を示し、VIC/FAMシグナルはヘテロ接合性の変異遺伝子型を示すであろう。

30

40

【0252】

[0266]一部の実施形態では、ECM1のSNP rs7511649を検出するためのプローブは、配列：TTTCTGACTTCTCCCTGTAAATCTTCTTCTGTATGATTTATTTTGGTAGAT；及びTTTCTGACTTCTCCCTGTAAATCTTCTTCTTCTGATTTATTTTGGTAGATを含み、両方ともTTTCTGACTTCTCCCTGTAAATCTT [C/T] TCTGTATGATTTATTTTGGTAGATに由来し、[C/T]の表記はrs7511649 SNPの位置を表す。さらなる実施形態では、第1のプローブはVIC (商標) 色素で標識され、C対立遺伝子を含有し、第2のプローブはFAM (商標) で標識され、T対立遺伝子を含有する。rs7511649 SNPの存在又は非存在の検出に関して、FAM/FAM (T/T) シグナル又はVIC/VIC (A/A) シグナルはホモ接合性の遺伝子型を示し、VIC/FAMシグナルはヘテロ接合性の変異遺伝子型を示すであろう。

50

M/FAM (T/T) シグナル又はVIC/VIC (A/A) シグナルはホモ接合性の遺伝子型を示し、VIC/FAMシグナルはヘテロ接合性の変異遺伝子型を示すであろう。

【0253】

[0267]他の実施形態では、NKX2-3のSNP rs11190140を検出するためのプローブは、配列：CATTTCAGGCTCCTGATTTCAATAGGCGGGA
AAGAAGGCTGCCAAGGCTGGG；及びCATTTCAGGCTCCTGATTTCAATAGGTTTCAATAGGTGGAAAGAAGGCTGCCAAGGCTGGGを含み、
両方ともCATTTCAGGCTCCTGATTTCAATAGG [C/T] GGAAAA
GAAGGCTGCCAAGGCTGGGに由来し、[C/T]の表記はrs11190140 SNPの位置を表す。さらなる実施形態では、第1のプローブはVIC (商標) 10
色素で標識され、C対立遺伝子を含むし、第2のプローブはFAM (商標) で標識され、
T対立遺伝子を含む。rs11190140 SNPの存在又は非存在の検出に関して、FAM/FAM (T/T) シグナル又はVIC/VIC (C/C) シグナルはホモ接
合性の遺伝子型を示し、VIC/FAMシグナルはヘテロ接合性の変異遺伝子型を示すであ
らう。

【0254】

[0268]配列解析も、本明細書に記載の方法に従って個体をジェノタイピングして、NO
D2遺伝子又は他の遺伝子マーカーにおける特定のバリエーション対立遺伝子又はハプロタイ
プの存在又は非存在を決定するために有用であり得る。当業者には知られているように、
目的のバリエーション対立遺伝子は、ATG16L1、ECM1、STAT3、NKX2-3 20
遺伝子のうちの1つ若しくは複数又は他の遺伝子マーカー中の目的の多型部位に隣接する
配列に基づいて設計された適切なプライマーを用いる配列解析によって検出することがで
きる。追加又は代替の配列プライマーは、ATG16L1、ECM1、STAT3、NK
X2-3遺伝子のうちの1つ若しくは複数又は他の遺伝子マーカー中の目的の多型部位の
約40～約400塩基対上流又は下流の配列に対応する配列の約15～約30個のヌクレ
オチドを含むことができる。そのようなプライマーは、一般に、配列決定反応におい
て安定なアニーリングステップを可能にする高い融解温度を達成するために十分なグアニ
ン及びシトシンの含有率を有するように設計される。

【0255】

[0269]用語「配列解析」には、核酸中のヌクレオチドの順序が決定される任意の手動又は 30
は自動の方法が含まれる。一例として、配列解析を使用してDNAのサンプルのヌクレオ
チド配列を決定することができる。用語「配列解析」には、これらに限定されないが、例
えばマクサムーギルバート及びサンガー配列決定並びにその変形を含むジデオキシ酵素方
法などの、化学的及び酵素的の方法が含まれる。用語「配列解析」には、これらに限定さ
れないが、キャピラリー電気泳動及びレーザー誘導性蛍光検出に依存し、MegaBACE
1000又はABI3700などの機器を用いて行うことができるキャピラリーアレイ
DNA配列決定がさらに含まれる。さらなる非限定的な例として、用語「配列解析」に
は、熱サイクル配列決定 (Searsら, Biotechniques, 13:626～
633 (1992) を参照)、固相配列決定 (Zimmermanら, Methods
Mol. Cell Biol., 3:39～42 (1992) を参照)、及びマトリックス 40
支援レーザー脱離/イオン化飛行時間質量分析などの質量分析を用いた配列決定 (MA
LDI-TOF MS) (Fuら, Nature Biotech., 16:381～3
84 (1998) を参照) が含まれる。用語「配列解析」には、これらに限定されない
が、配列のセグメントを同定するためにすべての可能な短いオリゴヌクレオチドのアレイ
に依存する、ハイブリダイゼーションによる配列決定 (SBH) がさらに含まれる (Chee
ら, Science, 274:610～614 (1996) ; Drmanacら, S
cience, 260:1649～1652 (1993) ; 及びDrmanacら, Na
ture Biotech., 16:54～58 (1998) を参照)。当業者には、こ
れら及びさらなる変形が本明細書中で定義される用語「配列解析」に含まれることが理
解されよう。 50

【0256】

[0270]電気泳動分析も、本発明の方法に従って個体をジェノタイピングして、A T G 1 6 L 1、E C M 1、S T A T 3、N K X 2-3 遺伝子のうちの1つ若しくは複数又は他の遺伝子マーカーにおける特定のバリエーション対立遺伝子又はハプロタイプの存在又は非存在を決定するために有用であり得る。増幅断片などの1つ又は複数の核酸に関連して本明細書で使用される「電気泳動分析」には、荷電分子が電界の影響下で静置培地を通して移動する方法が含まれる。電気泳動遊走は、核酸を主にその荷電に基づいて分離し、荷電は核酸の大きさに比例しており、より小さな分子はより速く遊走する。用語「電気泳動分析」には、これらに限定されないが、アガロース若しくはポリアクリルアミドゲル電気泳動などのスラブゲル電気泳動、又はキャピラリー電気泳動を使用した分析が含まれる。キャピラリー電気泳動分析は、一般に、小さな直径（50～100m）の石英キャピラリー内で、高い（キロボルトレベル）の分離電位の存在下で、数分間の分離時間で起こる。キャピラリー電気泳動分析を使用して、核酸はUV吸収又は蛍光標識によって好都合に検出され、数百塩基対までの断片で一塩基の分解能を得ることができる。そのような電気泳動分析方法及びその変形は、例えばAusubelら、Current Protocols in Molecular Biology, 第2章（補遺45）、John Wiley & Sons, Inc., New York（1999）に記載のように当分野で周知である。

10

【0257】

[0271]制限断片長多型（RFLP）分析も、本発明の方法に従って個体をジェノタイピングして、NOD2 遺伝子又は他の遺伝子マーカーにおける特定のバリエーション対立遺伝子又はハプロタイプの存在又は非存在を決定するために有用な場合がある（Jarchoら、Current Protocols in Human Genetics（Dracopolisら）、第2.7.1～2.7.5ページ、John Wiley & Sons, New York；Innisら（編）、PCR Protocols, San Diego: Academic Press, Inc.（1990）を参照）。本明細書において、「制限断片長多型分析」には、特定の塩基配列、一般には回文構造又は逆方向反復を認識した後に核酸の分解を触媒するエンドヌクレアーゼである制限酵素を使用して多型対立遺伝子を区別するための任意の方法が含まれる。当業者には、RFLP分析の使用は、多型部位でバリエーション対立遺伝子を野生型又は他の対立遺伝子から識別することができる酵素に依存することが理解されよう。

20

30

【0258】

[0272]また、対立遺伝子に特異的なオリゴヌクレオチドハイブリダイゼーションは、本明細書に記載の方法において個体をジェノタイピングして、A T G 1 6 L 1、E C M 1、S T A T 3、N K X 2-3 遺伝子のうちの1つ若しくは複数又は他の遺伝子マーカーにおける特定のバリエーション対立遺伝子又はハプロタイプの存在又は非存在を決定するために有用であり得る。対立遺伝子に特異的なオリゴヌクレオチドハイブリダイゼーションは、例えばバリエーション対立遺伝子を包含する配列に完全に相補的な配列を有する、標識したオリゴヌクレオチドプローブの使用に基づいている。適切な条件下では、バリエーション対立遺伝子に特異的なプローブは、該バリエーション対立遺伝子を含有する核酸とはハイブリダイズするが、プローブに比べて1つ又は複数のヌクレオチドのミスマッチを有する1つ又は複数の他の対立遺伝子とはハイブリダイズしない。所望の場合は、代替（例えば野生型）対立遺伝子に一致する、第2の対立遺伝子に特異的なオリゴヌクレオチドプローブも使用することができる。同様に、対立遺伝子に特異的なオリゴヌクレオチドの増幅法を使用して、例えば、バリエーション対立遺伝子のヌクレオチド配列と完全に相補的であるが、他の対立遺伝子に比べて1つ又は複数のミスマッチを有する、対立遺伝子に特異的なオリゴヌクレオチドプライマーを使用することによって、バリエーション対立遺伝子を選択的に増幅することができる（Mullisら（上記））。当業者には、バリエーション対立遺伝子と他の対立遺伝子を区別する1つ又は複数のヌクレオチドのミスマッチは、多くの場合、対立遺伝子に特異的なオリゴヌクレオチドハイブリダイゼーションにおいて使用する対立遺伝子に特異

40

50

的なオリゴヌクレオチドプライマーの中心に位置することが理解されよう。対照的に、PCR増幅において使用する対立遺伝子に特異的なオリゴヌクレオチドプライマーは、一般に、バリエーションと他の対立遺伝子を区別する1つ又は複数のヌクレオチドのミスマッチをプライマーの3'末端に含有する。

【0259】

[0273]ヘテロ二重鎖移動度アッセイ(HMA)は、本発明の方法においてジェノタイピングを行って、ATG16L1、ECM1、STAT3、NKX2-3遺伝子のうちの1つ若しくは複数又は他の遺伝子マーカーにおける特定のバリエーション対立遺伝子又はハプロタイプの存在又は非存在を決定するために使用することができる他の周知のアッセイである。ミスマッチを保有するDNA二重鎖は、完全に塩基対合した二重鎖の移動度に比べてポリアクリルアミドゲル中での移動度が低下しているため、HMAはバリエーション対立遺伝子の存在を検出するために有用である(Delwartら, Science, 262:1257~1261(1993); Whiteら, Genomics, 12:301~306(1992)を参照)。

【0260】

[0274]一本鎖コンホメーション多型性(SSCP)の方法も、本明細書に記載の方法においてジェノタイピングを行って、ATG16L1、ECM1、STAT3、NKX2-3遺伝子のうちの1つ若しくは複数又は他の遺伝子マーカーにおける特定のバリエーション対立遺伝子又はハプロタイプの存在又は非存在を決定するために有用であり得る(Hayashi, Methods Appl., 1:34~38(1991)を参照)。この方法を使用して、非変性ゲル電気泳動の際に変更された電気泳動移動度を生じる一本鎖DNAの二次構造の差異に基づいてバリエーション対立遺伝子を検出することができる。バリエーション対立遺伝子は、試験断片の電気泳動パターンを既知の対立遺伝子を含有する対応する標準の断片と比較することによって検出される。

【0261】

[0275]変性勾配ゲル電気泳動(DGGE)も、本発明の方法において、ATG16L1、ECM1、STAT3、NKX2-3遺伝子のうちの1つ若しくは複数又は他の遺伝子マーカーにおける特定のバリエーション対立遺伝子又はハプロタイプの存在又は非存在を決定するために有用であり得る。DGGEでは、二本鎖DNAを、漸増濃度の変性剤を含有するゲル中で電気泳動する。ミスマッチした対立遺伝子からなる二本鎖断片は、より迅速に融解するセグメントを有しており、それにより、そのような断片が完全に相補的な配列に比べて異なって移動することが引き起こされる(Sheffieldら, "Identifying DNA Polymorphisms by Denaturing Gradient Gel Electrophoresis", Innisら(上述)、1990を参照)。

【0262】

[0276]個体をジェノタイピングするために有用な他の分子学的方法が当分野で知られており、本発明の方法において有用である。そのような周知のジェノタイピングの方法には、これらに限定されないが、自動配列決定及びRNaseミスマッチ法が含まれる(Winterら, Proc. Natl. Acad. Sci., 82:7575~7579(1985)を参照)。さらに、当業者には、多数のバリエーション対立遺伝子の存在又は非存在を決定する場合は、分子学的方法の任意の組み合わせによって個々のバリエーション対立遺伝子を検出できることが理解されよう。Birrenら(編), Genome Analysis: A Laboratory Manual, 第1巻(Analyzing DNA), New York, Cold Spring Harbor Laboratory Press(1997)を参照されたい。さらに、当業者には、多数のバリエーション対立遺伝子を個々の反応又は単一の反応において検出できることが理解されよう(「多重鎖」アッセイ)。

【0263】

[0277]上述に鑑みて、IBD又はその臨床亜型に関する診断的情報を提供するため、及

びIBDの進行の転帰及び経過に関する予後診断的及び予測的情報を提供するため、及びIBDの処置のための適切な治療方式の選択に関する情報を提供するための本発明の方法は、上述の周知のジェノタイピングアッセイ又は当技術分野で知られている他のアッセイの1つ又は任意の組み合わせを使用して（例えば、ATG16L1、STAT3、ECM1及びNKX2-3などの遺伝子の1つ又は複数のバリエーション対立遺伝子の有無を決定することによって）実施できることを当業者は認識している。

【0264】

VIII. 統計解析

一部の態様において、本発明は、IBD又は非IBDを診断するため、IBDの診断を分類するため（例えば、CD、UC、又はCD若しくはUCに関して不確定）、IBDの亜型をUC、CD又はICとして分類するため、又はUCとCDとを鑑別するための方法及びシステムを提供する。一部の実施形態では、本明細書に記載のアッセイのうちのいずれかによって決定された1つ又は複数のIBDマーカーの存在、レベル及び／又は遺伝子型に1つ又は複数の学習統計的分類システムを適用して、IBDを診断するか、又はその亜型を決定する。他の実施形態では、本明細書に記載のアッセイのうちのいずれかによって決定された1つ又は複数のIBDマーカーの存在、レベル及び／又は遺伝子型に分位分析を適用して、IBDを診断するか、又はその亜型を決定する。本明細書に記載のように、本発明の統計解析は、有利なことに、IBDの診断又はその亜型の決定に関して、改善された感度、特異度、陰性予測値、陽性予測値、及び／又は全体的な精度をもたらす。

【0265】

[0279]用語「統計解析」又は「統計アルゴリズム」又は「統計プロセス」には、変数間の関係性を決定するために使用される種々の統計学的方法及びモデルのうちの任意のものが含まれる。本発明では、変数は、少なくとも1つの目的のマーカーの存在、レベル、又は遺伝子型である。任意の数のマーカーを、本明細書に記載の統計解析を使用して分析することができる。例えば、1個、2個、3個、4個、5個、6個、7個、8個、9個、10個、11個、12個、13個、14個、15個、16個、17個、18個、19個、20個、25個、30個、35個、40個、45個、50個、55個、60個、又はそれより多くのマーカーの存在又はレベルを統計解析に含めることができる。一実施形態では、ロジスティック回帰を使用する。他の実施形態では、直線回帰を使用する。いくつかの実施形態では、本発明の統計解析は、例えば所定の集団内の変数としての、1つ又は複数のマーカーの分位測定を含む。分位とは、データのサンプルを（可能な限り）等しい数の観察値を含有する群へと分ける1セットの「切点」である。例えば、四分位数とは、データのサンプルを4つの（可能な限り）等しい数の観察値を含有する群へと分ける値である。下位四分位数とは、順序づけたデータセットを4分の1上ったデータ値であり、上位四分位数とは、順序づけたデータセットを4分の1下ったデータ値である。五分位数とは、データのサンプルを5つの（可能な限り）等しい数の観察値を含有する群へと分ける値である。本発明には、（連続的な変数と同様に）統計解析における変数としての、マーカーレベルのパーセンタイル範囲（例えば、三分位数、四分位数、五分位数）又はその累積索引（例えば、合計四分位数スコア（QSS）を得るためのマーカーレベルの四分位数の和）の使用も含まれる場合がある。

【0266】

[0280]いくつかの実施形態では、本発明は、四分位数分析を使用して1つ又は複数の目的のマーカーの存在、レベル（例えば規模）、及び／又は遺伝子型を検出又は決定することを含む。この種類の統計解析では、目的のマーカーのレベルは、サンプルの参照データベースに関して第1四分位数（<25%）、第2四分位数（25～50%）、第3四分位数（51%～<75%）、又は第4四分位数（75～100%）にあるとして定義される。これらの四分位数は、それぞれ1、2、3、及び4の四分位数スコアを割り当ててもよい。いくつかの例では、サンプル中で検出されないマーカーは、0又は1の四分位数スコアが割り当てられる一方で、サンプル中で検出される（例えば存在する）マーカー（例えば、サンプルはマーカーに関して陽性である。）は、4の四分位数スコアが割り当てられ

る。一部の実施形態では、四分位数1は最低のマーカーレベルを有するサンプルを表す一方で、四分位数4は最高のマーカーレベルを有するサンプルを表す。他の実施形態では、四分位数1は特定のマーカー遺伝子型（例えば野生型対立遺伝子）を有するサンプルを表す一方で、四分位数4は他の特定のマーカー遺伝子型（例えば対立遺伝子バリエーション）を有するサンプルを表す。サンプルの参照データベースには、広い範囲のIBD（例えば、CD及び／又はUC）患者が含まれる場合がある。そのようなデータベースから、四分位数のカットオフを確立することができる。本発明における使用に適した四分位数分析の非限定的な例は、例えばMowら、Gastroenterology, 126:414~24(2004)に記載されている。

【0267】

[0281]好ましい実施形態では、本発明の統計解析は1つ又は複数の学習統計的分類子システムを含む。本明細書において、用語「学習統計的分類子システム」には、複雑なデータセット（例えば目的のマーカーのパネル）に適応し、そのようなデータセットに基づいて判断を行うことができる機械学習アルゴリズム法が含まれる。一部の実施形態では、判断／分類木（例えば、ランダムフォレスト（RF）又は分類及び回帰木（C&RT））などの単一の学習統計的分類子システムを使用する。他の実施形態では、2個、3個、4個、5個、6個、7個、8個、9個、10個、又はそれより多くの学習統計的分類子システムの組み合わせを、好ましくはタンデムで使用する。学習統計的分類子システムの例には、これらに限定されないが、帰納的学習（例えば、ランダムフォレスト、分類及び回帰木（C&RT）、ブーストツリーなどの判断／分類木）、確率的近似（PAC）学習、コネクション主義学習（例えば、神経回路網（NN）、人工神経回路網（ANN）、ニューロファジー回路網（NFN）、ネットワーク構造、多層パーセプトロンなどのパーセプトロン、多層フィードフォワードネットワーク、神経回路網の適用、信念ネットワークにおけるベイジ学習）、強化学習（例えば、ナイーブ学習、適応性動的学習、及び時間差学習などの既知の環境における受動的学習、未知の環境における受動的学習、未知の環境における能動的学習、学習行動価値関数、強化学習の適用）、並びに遺伝的アルゴリズム及び進化的プログラミングを使用するものが含まれる。他の学習統計的分類子システムには、サポートベクターマシン（例えばカーネル法）、多変量適応型回帰スプライン（MARS）、レーベンバーグマルカルトアルゴリズム、ガウスニュートンアルゴリズム、混合ガウス、勾配降下アルゴリズム、及び学習ベクター量子化（LVQ）が含まれる。

【0268】

[0282]ランダムフォレストとは、Leo Breiman及びAdele Cutlerによって開発されたアルゴリズムを使用して構築された学習統計的分類子システムである。ランダムフォレストでは、多数の個々の決定木を使用し、個々の木によって決定されたクラスの最頻値（すなわち最も頻繁に起こっているもの）を選択することによってクラスを決定する。ランダムフォレスト分析は、例えば、Salford Systems（San Diego, CA）から入手可能なランダムフォレスト（Random Forests）ソフトウェアを用いて行うことができる。ランダムフォレストの説明としては、例えば、Breiman, Machine Learning, 45:5~32(2001)；及びhttp://stat-www.berkeley.edu/users/breiman/RandomForests/cc_home.htmを参照されたい。

【0269】

[0283]分類及び回帰木は、古典的回帰モデルの当てはめのコンピュータ集約的な代替方法を表しており、典型的には、1つ又は複数の予測子に基づいて目的のカテゴリ的又は連続的な応答の最良な可能なモデルを決定するために使用する。分類及び回帰木の分析は、例えば、Salford Systemsから入手可能なC&RTソフトウェア又はStatSoft, Inc.（Tulsa, OK）から入手可能なスタティスティカ（Statistica）データ分析ソフトウェアを用いて行うことができる。分類及び回帰木は、例えば、Breimanら、"Classification and Regre

ssion Trees”, Chapman and Hall, New York (1984); 及び Steinberg ら, “CART: Tree-Structured Non-Parametric Data Analysis”, Salford Systems, San Diego (1995) に説明されている。

【0270】

[0284] 神経回路網とは、計算のコネクショニストアプローチに基づく情報処理に数学的又は計算的なモデルを使用する、相互接続された人工ニューロンの群である。典型的には、神経回路網は、ネットワーク内を流れる外部又は内部情報に基づいてその構造を変化させる適応性システムである。神経回路網の具体的な例には、パーセプトロン、単層パーセプトロン、多層パーセプトロン、逆伝搬ネットワーク、アダリン (ADALINE) ネットワーク、マダリン (MADALINE) ネットワーク、ラーンマトリックス (Learnmatrix) ネットワーク、ラジアル基底関数 (RBF) ネットワーク、及び自己組織化マップ又はコホーネン自己組織化ネットワークなどのフィードフォワード神経回路網、単純再帰ネットワーク及びホップフィールドネットワークなどの再帰神経回路網、ボルツマン機械などの確率的神経回路網、機械及び連想神経回路網の委員会などのモジュラー神経回路網、並びに瞬時訓練神経回路網、スパイキング神経回路網、動的神経回路網、及びカスケード神経回路網などの他の種類のネットワークが含まれる。神経回路網の分析は、例えば StatSoft, Inc. から入手可能なスタティスティカデータ分析ソフトウェアを用いて行うことができる。神経回路網の説明には、例えば、Freeman ら, “Neural Networks: Algorithms, Applications and Programming Techniques”, Addison-Wesley Publishing Company (1991); Zadeh, Information and Control, 8: 338~353 (1965); Zadeh, “IEEE Trans. on Systems, Man and Cybernetics”, 3: 28~44 (1973); Gersho ら, “Vector Quantization and Signal Compression”, Kluwer Academic Publishers, Boston, Dordrecht, London (1992); 及び Hassoun, “Fundamentals of Artificial Neural Networks”, MIT Press, Cambridge, Massachusetts, London (1995) を参照されたい。

【0271】

[0285] サポートベクターマシーンとは、分類及び回帰に使用される 1 セットの関連する教師ありの学習法であり、例えば Cristianini ら, “An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods”, Cambridge University Press (2000) に記載されている。サポートベクターマシーン分析は、例えば、Thorsten Joachims (カーネル大学) によって開発された SVM^{light} ソフトウェアを使用して、又は Chih-Chung Chang 及び Chih-Jen Lin (台湾国立大学) によって開発された LIBSVM ソフトウェアを用いて行うことができる。

【0272】

[0286] 本明細書に記載の種々の統計学的方法及びモデルは、健常な個体並びに IBD (例えば、CD 及び／又は UC) 患者からのサンプル (例えば、血清学的及び／又は遺伝的サンプル) のコホートを使用して訓練及び試験することができる。例えば、医師、好ましくは胃腸科専門医によって、例えば米国特許第 6218129 号に記載のように生検、大腸内視鏡検査、又は免疫アッセイを使用して IBD 又はその臨床亜型に罹患しているとして診断された患者からのサンプルが、本発明の統計学的方法及びモデルの訓練及び試験における使用に適している。また、IBD が診断された患者からのサンプルは、例えば、米国特許第 5750355 号及び第 5830675 号に記載の免疫アッセイを用いてクローン病又は潰瘍性大腸炎にも階層化することができる。健常な個体からのサンプルには、I

B Dサンプルとして同定されなかったものを含めることができる。当業者は、本発明の統計学的方法及びモデルを訓練及び試験するために使用することができる患者サンプルのコホートを得るためのさらなる方法及び診断基準を知っているであろう。

【0273】

[0287]本明細書において、用語「感度」とは、サンプルが陽性である場合に、本発明の診断、又は予測方法、システム又はコードが陽性結果を与える、例えば、予測された診断を有する確率をいう。感度は、真の陽性結果の数を真陽性及び偽陰性の和で除算することによって計算する。感度は、本質的には、本発明がどれだけ良好に予測された診断を有する者を、予測された診断を有さない者から正しく同定するかの測度である。統計学的方法及びモデルは、感度が少なくとも約60%となるように選択することができ、例えば、少

10

【0274】

[0288]用語「特異度」とは、サンプルが陽性でない場合に、本発明の診断又は予測方法、システム又はコードが陰性結果を与える、例えば、予測された診断を有さない確率をいう。特異度は、真の陰性結果の数を真陰性及び偽陽性の和で除算することによって計算する。特異度は、本質的には、本発明がどれだけ良好に予測された診断を有さない者を予測された診断を有する者から排除するかの測度である。統計学的方法及びモデルは、特異度が少なくとも約60%となるように選択することができ、例えば、少なくとも約65%、

20

【0275】

[0289]本明細書において、用語「陰性予測値」すなわち「NPV」とは、予測された診断を有さないとして同定された個体が実際には予測された診断を有さない確率をいう。陰性予測値は、真陰性の数を真陰性及び偽陰性の和で除算することによって計算することができる。陰性予測値は、診断又は予後診断方法、システム又はコードの特徴、及び分析した集団における疾患の有病率によって決定する。統計学的方法及びモデルは、疾患有病率を有する集団中の陰性予測値が約70%～約99%の範囲内となるように選択することが

30

【0276】

[0290]用語「陽性予測値」すなわち「PPV」とは、予測された診断を有するとして同定された個体が実際に予測された診断を有する確率をいう。陽性予測値は、真陽性の数を真陽性及び偽陽性の和で除算することによって計算することができる。陽性予測値は、診断又は予後診断方法、システム又はコードの特徴、及び分析した集団における疾患の有病率によって決定する。統計学的方法及びモデルは、疾患有病率を有する集団中の陽性予測値が約70%～約99%の範囲内となるように選択することができ、例えば、少なくとも

40

【0277】

[0291]陰性及び陽性の予測値を含む予測値は、分析した集団における疾患の有病率によって影響を受ける。本発明では、統計学的方法及びモデルは、特定のIBDの有病率を有する臨床集団について所望の臨床的パラメータを生成するように選択することができる。例えば、統計学的方法及びモデルは、例えば胃腸科専門医院又は一般開業医院などの診療所において見ることができる、約1%、2%、3%、4%、5%、6%、7%、8%、9

50

%、10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、又は70%までのIBDの有病率について選択することができる。

【0278】

[0292]本明細書において、用語「全体的な一致」又は「全体的な精度」とは、本発明の診断又は予後診断方法、システム又はコードがIBDを診断又は予測する精度をいう。全体的な精度は、真陽性及び真陰性の和をサンプルの結果の合計数で除算したものとして計算し、分析した集団における疾患の有病率によって影響を受ける。例えば、本発明の統計学的方法及びモデルは、疾患有病率を有する患者集団における全体的な精度が少なくとも約40%であるように選択することができ、例えば、少なくとも約40%、41%、42%、43%、44%、45%、46%、47%、48%、49%、50%、51%、52%、53%、54%、55%、56%、57%、58%、59%、60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、又は99%であり得る。

10

【0279】

[0293]特定の実施形態では、IBDと関連する診断マーカーの予測モデリングのための本明細書に記載の統計学的方法及びモデルは、モデル構築のランダムフォレスト法を用いる。ランダムフォレストは、一般的には、各木を構築する無作為成分を有する多くの決定木から成り立つ分類子である。各決定木は、同じ分類問題に対処するが、例の異なる集合及び例のデータセットから無作為に選択される特徴の異なるサブセットが提供される。

20

【0280】

[0294]一部の実施形態では、1本の木は、利用可能なサンプル（例えば訓練セット）の無作為の2/3を、木の各ノードで分割される決定を行うために選択される特徴の無作為の2/3と共に用いて構築することができる。訓練の間にフォレストが構築されたら、すべての決定木にわたって採決することによって新しい例を分類する。2クラスのランダムフォレスト分類子に関する最も単純な事例では、最も投票数が多いクラスが勝つ。他の例では、訓練中に確立された、投票の勝利数のカットオフを予め設定して、性能評価尺度を最適化する。一部の例では、偽陽性が偽陰性よりもコストが高い場合、カットオフをより高く設定することができる。

30

【0281】

[0295]前記方法は非常に正確であるため、この種類の統計学的データをモデリングするのに非常に強固である。各決定木はデータの異なるサブセットから構築されるため、それは他の機械学習アプローチよりもデータの過剰適合の傾向が少ない。前記方法はまた、そのアンサンブル性質のため過剰訓練に対して本質的に抵抗性である。典型的には、多数の木を訓練データのランダムサブセットに成長させ、訓練し、個々の木の結果を平均することによって予測を行う。個々の木の過剰適合又は性能不足は、アンサンブル平均が計算された際に取消される傾向がある。ランダムフォレストモデリングの利点としては、取消しに起因する過剰訓練の最小化、予測空間における任意形状領域の作出、任意非線形マッピングの作出、変数重要性、及び異常値が挙げられる。前記方法は、変数（マーカー）重要性をランク付けする能力を有し、リサンプリングの際の計算効率を有する。

40

【0282】

[0296]一部の態様では、本発明のランダムフォレストモデリングはアンサンブル方法を含み、フォレストは数千の決定木を含む。決定木は、一連の質問（例えばスプリット）を生成することによって予測に使用することができる。配列中の所与のノードにおいて、尋ねられる質問（例えばスプリット）は以前の質問に対する回答に依存する。木は、それがクラスをどのように良く分離することができるかに基づいて、各ノードでの最良のスプリットを選択する。これらの木は、訓練データの異なる無作為なサブサンプルを用いることによって生成される。他の例では、木は、各ノードでのマーカーの異なる無作為なサブセットを用いて生成される。予測モードの間、各クラスの確率は、それを予測するアルゴリ

50

ズムにおける木の割合である。

【0283】

[0297] 一部の実施形態では、ランダムフォレストの一般的ステップは、(1) 訓練に利用可能なデータの大きさに2/3から従来の再帰分割木を構築し、その後の枝刈りを用いないステップ；(2) 各ノードで利用可能な変数の小さく異なるサブセットであり、変数の数の平方根以下である前記サブセットのみを考慮するステップ；(3) 複数の木（フォレスト）を作出するためにステップ1及び2を複数回反復するステップ；及び(4) 予測モードの間に、所与のクラスを予測するフォレストにおける木の割合であるスコアを報告するステップを含む。次いで、カットオフを用いて、クラス予測（例えば、IBD、非IBD、IC、UC若しくはCD診断、あるいはCD及びUCに関して不確定とサンプルを分類すること）をもたらすことができる。

10

【0284】

[0298] いくつかの実施形態では、IBD診断アルゴリズムは、血清-遺伝子-炎症マーカー（例えば実施例）の既知の存在、非存在及び/又はレベルを有する予測された診断の後ろ向きコホートを用いて確立される。一部の実施形態では、ランダムフォレストを構築するのに用いられる例（例えば訓練及び試験セット）は、本明細書に記載の方法を用いる研究に由来するものである。

【0285】

[0299] 特定の実施形態では、血清-遺伝子-炎症マーカーの存在、非存在及び/又はレベルのテキストベースの測定値並びに既知の疾患診断は、決定されたルールに基づいてバイナリー値を割り当てられる。一部の実施形態では、検出されないか、又はDNase感受性（細胞質性）であると決定されたpANCA測定値は0の値を割り当てられる。他の実施形態では、DNase感受性1+P、2+P、3+P又は4+Pであると決定されたpANCA測定値は1の値を割り当てられる。一部の実施形態では、検出されない、DNase感受性（細胞質性）であるか、又はDNase感受性1+Pであると決定されたpANCA 2測定値は0の値を割り当てられる。他の実施形態では、DNase感受性2+P、3+P又は4+Pであると決定されたpANCA 2測定値は1の値を割り当てられる。

20

【0286】

[0300] いくつかの実施形態では、IBD診断アルゴリズムのいくつかの部分において用いられる変数（例えば、遺伝子マーカー、血清学的マーカー及び既知の疾患診断）のすべての可能な対の組み合わせに関する相関及び関係を算出した。連続変数対のため、及び連続バイナリー変数対のために、Pearsonの相関係数及びそのp値を算出することができる。バイナリー-バイナリー変数対のためには、カイ二乗分析を用いて関係を評価することができる。すべての計算を、ソフトウェアプログラムMatlabを用いて実施することもできる。

30

【0287】

[0301] 一部の実施形態では、IBD診断アルゴリズムは、決定された基準を満たすサンプルが訓練データから除去され、UC対CDを予測するためのランダムフォレストがICではない既知のIBD患者のみに由来する訓練データを使用し、決定木又はルールのセットが「分類不能」ルールに基づく、サンプル除去後のすべての訓練データを用いてIBD対非IBDを予測するためのランダムフォレストモデルを含む。いくつかの実施形態では、IBD診断アルゴリズムが予測モードにある場合、それは第1のモデル（例えば、IBD対非IBDに関するランダムフォレスト）を用いてIBD対非IBDを予測することができる。非IBDが予測される場合、アルゴリズムは終了する。そうでなければ、アルゴリズムは、決定木又はルールのセットにおいて「分類不能」ルールを適用する。「分類不能」ルールに基づいてCD及びUCに関して不確定と分類されないサンプルは、第2のモデル（例えば、UC対CDに関するランダムフォレスト）を適用することにより処理される。IBD診断アルゴリズムは、個体を非IBD、CD若しくはUC患者と分類するか、又はIBDを有する個体をCD及びUCに関して不確定と分類することができる。

40

50

【0288】

[0302]いくつかの実施形態では、「分類不能」ルール又はルールのセットは、 $ANCA \geq Q3$ 、 $pANCA$ 陽性、及び抗 $Cbir1$ 抗体、抗 $Fla2$ 抗体若しくは抗 $FlaX$ 抗体のいずれかが $\geq Q3$ であり得る。他の実施形態では、「分類不能」ルール又はルールのセットは、 $pANCA$ 陽性、及び血清抗フラジェリン抗体（例えば、抗 $Cbir1$ 抗体、抗 $Fla2$ 抗体、抗 $FlaX$ 抗体）のうちのいずれか2つのマーカーが $\geq Q3$ であり得る。

【0289】

[0303]一部の実施形態では、IBD診断アルゴリズムは、訓練及び試験の両方の反復（例えばリサンプリング）が繰り返される内部検証の方法を含む。いくつかの例では、リサンプリングは、統計値を得るために約100回繰り返される。内部検証は、訓練セットを分割して、2/3をモデル訓練セットにし、1/3を試験セットにするステップ；モデル訓練セットを用いてアルゴリズム（例えば、一対のランダムフォレストモデル）を構築するステップ；及び試験セットを用いて性能を計算するステップを含む。他の実施形態では、内部検証は、訓練セットに適用された検証セットと同じサンプル除去ルール及び「分類不能」ルールを適用するステップ並びにアルゴリズムの性能を算出するステップを含む。いくつかの例では、同じ選抜除去ルールを適用できない場合、2%以上の性能低下を生じ得る。

【0290】

[0304]一部の実施形態では、サンプルが、最後に経験した症状が医師への訪問の6か月より前であった患者に由来する場合；サンプルが、最後の症状が「利用できない」と記述され、特定の日付の前に試験に登録した患者に由来する場合；患者が試験対象患者基準／試験除外基準を満たさなかった場合；及び患者の疾患持続期間に関する情報がない場合、IBD対非IBD又はUC対CDのいずれかを予測するためのランダムフォレストモデルからサンプルが除去される。

【0291】

[0305]一部の実施形態では、サンプルが、最後に経験した症状が医師への訪問の6か月より前であった患者に由来する場合；サンプルが、最後の症状が「利用できない」と記述され、特定の日付の前に試験に登録した患者に由来する場合；患者が試験対象患者基準／試験除外基準を満たさなかった場合；及び患者の疾患持続期間に関する情報がない場合、サンプルが除去される。他の実施形態では、UC対CD診断のためのランダムフォレストモデルからサンプルが除去される。一部の例では、次の特徴を有するサンプルは、UC対CDランダムフォレストモデルを構築するのに用いられる訓練データセットから除去される。その特徴としては、（1）UC診断；（2）四分位スコア（Q）3未満のANCAレベル、 $pANCA$ 陰性；及び（3） $Q \geq 3$ である2つ以上のマーカー（例えば、 $ASCA-A$ 、 $ASCA-G$ 、抗 $OmpC$ 、抗 $Cbir1$ 、抗 $Fla2$ 、抗 $FlaX$ ）が挙げられる。典型的には、これらのサンプルの90%がCD患者に由来する。一部の例では、次の基準を満たすサンプルは排除される。（1）既知のCD診断；（2） $ANCA < 16.8 EU/ml$ 、 $pANCA$ 陰性；及び（3）次の条件のうちの2つ以上が満たされる： $ASCA-A \geq 10.675 EU/ml$ 、 $ASCA-G \geq 14.175 EU/ml$ 、抗 $OmpC \geq 9.4 EU/ml$ 、抗 $Cbir1 \geq 29.475 EU/ml$ 、抗 $Fla2 \geq 34.5 EU/ml$ 、又は抗 $FlaX \geq 28.875 EU/ml$ 。他の例では、（1）対象がUC診断を有し；（2） $ANCA \geq 16.8 EU/ml$ ； $pANCA$ 陽性であり；（3）次の条件： $ASCA-A < 20.675 EU/ml$ 、 $ASCA-G < 14.175 EU/ml$ 、抗 $OmpC < 9.4 EU/ml$ 、抗 $Cbir1 < 29.475 EU/ml$ 、抗 $Fla2 < 34.5 EU/ml$ 及び抗 $FlaX < 28.875 EU/ml$ のすべてが満たされる場合、UC対CDランダムフォレストモデルを構築するために用いられる訓練データセットから対象のサンプルが除去される。これらの対象の90%は典型的にはUC患者に由来する。さらに別の例では、サンプルが次の基準のうちの1つを満たす場合、そのサンプルは訓練セット（例えば訓練及び検証セット）から除外される。（1） $pANCA =$

10

20

30

40

50

1であり、 $ANCA < 11.9$ である；(2) $pANCA = 0$ であり、 $ANCA \geq 11.9$ である；(3) 明らかな誤差の影響を受ける；及び(4) 再試験されるのが通常である。

【0292】

[0306] 一部の実施形態では、本発明の疾患分類アルゴリズムは、1つ又は複数のランダムフォレストモデルを構築するために訓練セット全体を使用する。それは、リサンプリングフェーズの平均カットオフから疾患を分類するためのカットオフを確立する。アルゴリズムは、訓練セットを、検証セットのための $1/3$ とモデル訓練及び試験セットの両方のための $2/3$ とに繰り返し分割すること；平均カットオフ値を用いて、モデル訓練セット及び試験セット上で1つのアルゴリズムを作出すること；検証セット上での性能を計算し、それを試験セット上での平均性能と比較すること；並びに試験セット上での平均確率を検証セット上での最終確率と比較することによって内部的に検証される。さらに、試験セットの性能は、モデル訓練／試験セット（これは、訓練セット全体のサブセットである。）のサブセットであったモデル訓練セット及び試験セットから作出された一対のランダムフォレストモデルに基づくアルゴリズムから計算される。

【0293】

[0307] いくつかの実施形態では、本発明のIBD診断アルゴリズムは、17種の血清－遺伝子－炎症生物学的マーカー（例えば、ANCA、ASCA-A、ASCA-G、pANCA、抗FlaX、SAA、抗Fla2、ICAM、抗OmpC、抗Cbir1、VCAM、CRP、NKX2-3、ATG16L1、STAT3、ECM1、VEGF、及びそれらの組み合わせ）に由来する測定値を用いて、IBD対非IBDを予測するための第1のランダムフォレストモデルに基づくモデルスコアを計算する。第1のランダムモデルは、患者がIBDを有するかどうかを決定する。スコアがIBD対非IBDカットオフ未満（例えば < 0.64 ）である場合、サンプルはIBDを有する患者に由来するものと予測される。そうでなければ、サンプルは、非IBDを有する患者に由来するものと予測される。IBDを有すると予測されたサンプルは、マーカーパターンに基づいて不確定を除外するように設計された決定木又はルールのセットであるアルゴリズムの次のステップに進む。サンプルが「分類不能」ルールのいずれかのパターンに一致する場合、アルゴリズムはサンプルがIBDを有すると予測するが、CD及びUCに関しては不確定とサンプルを分類する。そうでなければ、サンプルは、UC対CDを予測するための第2のランダムフォレストモデルであるアルゴリズムの次のステップに進む。次いで、診断アルゴリズムは、11種の血清－遺伝子－炎症生物学的マーカー（例えば、ANCA、ASCA-A、ASCA-G、抗FlaX、抗Fla2、pANCA、抗OmpC、抗Cbir1、ECM1、STAT3、VEGF、及びそれらの組み合わせ）に由来する測定値を使用して、UC対CDを予測するための第2のランダムフォレストモデルに基づくモデルスコアを計算する。スコアがUC対CDカットオフ（例えば 0.35 ）未満である場合、アルゴリズムはサンプルがCDを有すると予測する。そうでなければ、アルゴリズムはサンプルがUCを有すると予測する。

【0294】

[0308] 一部の実施形態では、IBD対非IBDモデルにおいて使用される血清－遺伝子－炎症マーカーは重要度によってランク付けされる。訓練セットにおいて重要性の高い順にマーカーを列挙すると、ANCA、ASCA-A、ASCA-G、pANCA、抗FlaX、SAA、抗Fla2、ICAM、抗OmpC、抗Cbir1、VCAM、CRP、NKX2-3、ATG16L1、STAT3、ECM1、及びVEGFとなる。他の実施形態では、UC対CDモデルにおいて使用される血清－遺伝子－炎症マーカーは重要度によってランク付けされ、訓練セットにおいて重要性の高い順に次のように列挙される。ANCA、ASCA-A、ASCA-G、抗FlaX、抗Fla2、pANCA、抗OmpC、抗Cbir1、ECM1、STAT3、及びVEGF。マーカーの重要度は本明細書に記載されており、実施例3を参照されたい。

【0295】

[0309]他のいくつかの実施形態では、本発明の I B D 診断アルゴリズムは、サンプルが p A N C A 2 陰性であるかどうかを決定して、I B D 対非 I B D を予測するための第 1 のランダムフォレストモデル又は別々の p A N C A 決定木分析のいずれかにサンプルを向かわせる初期ステップを含む（図 2 及び実施例 9 参照）。いくつかの例では、陰性（0 若しくは「検出されず」）又は弱い（+1）p A N C A 決定のいずれかを有する p A N C A 2 サンプルを 2 ステップのランダムフォレストアルゴリズムキューから除去し、I B D の臨床予測を行うための決定行列又はルールセットにかける。p A N C A 決定木分析後に I B D を予測するための決定行列の非限定例は実施例 9 に記載される。すなわち、いくつかの態様では、p A N C A 決定木を、p A N C A 2（-）アッセイ結果を有するサンプルにおいて予測的 I B D 診断を行うための代替経路として実装することができる。

10

【0296】

I X. 疾患分類システム

[0310]図 3 は、本発明の一実施形態による疾患分類システム（D C S）（300）を例示する。そこに示される通り、D C S は、プロセッサ（315）及び記憶モジュール（310）を有するコンピュータなどの D C S 知能モジュール（305）を含む。知能モジュールはまた、1つ若しくは複数の直接接続（例えば、U S B、F i r e w i r e、又は他のインターフェース）及び 1つ若しくは複数のネットワーク接続（例えば、モデム又は他のネットワークインターフェースデバイス）上の情報を送信及び受信するための通信モジュール（示さず）も含む。記憶モジュールは、内部記憶デバイス及び 1つ若しくは複数の外部記憶デバイスを含んでもよい。知能モジュールはまた、モニター又はプリンターなどのディスプレイモジュール（325）も含む。一態様では、知能モジュールは、直接接続を介するか、又はネットワーク（340）上で、試験システム（350）などのデータ獲得モジュールから、患者の試験結果などのデータを受信する。例えば、試験システムを、1つ又は複数の患者サンプル（355）上で多分析物試験を実施し、試験結果を知能モジュールに自動的に提供するように構成することができる。また、データを、ユーザによる直接入力により知能モジュールに提供するか、又はそれをコンパクトディスク（C D）若しくはデジタル多用途ディスク（D V D）などの携帯用媒体からダウンロードすることもできる。試験システムを、知能モジュールと統合する、知能モジュールに直接連結する、又はネットワーク上で知能モジュールと遠隔的に連結することができる。知能モジュールはまた、周知のようにネットワーク上の 1つ又は複数のクライアントシステム（330）に、及びそれから、データを通信することもできる。例えば、要求中の医師又は医療機関は、クライアントシステム（330）を使用して、実験室又は病院にあってよい知能モジュールからの報告を取得し、見ることができる。

20

30

【0297】

[0311]ネットワークは、L A N（ローカルエリアネットワーク）、W A N（ワイドエリアネットワーク）、無線ネットワーク、ポイントツーポイントネットワーク、スター型ネットワーク、トークンリングネットワーク、ハブネットワーク、又は他の機器構成であってもよい。現在使用されている最も一般的な型のネットワークは、T C P / I P（伝送制御プロトコル及びインターネットプロトコル）ネットワーク、例えば、大文字「I」を用いて「インターネット」と呼ばれることが多いネットワークのグローバルインターネットであり、本明細書の実施例の多くにおいて用いられるが、T C P / I P は現在好ましいプロトコルであるが、本発明が使用するネットワークはそれに限定されないことが理解されるべきである。

40

【0298】

[0312]図 3 に示されるシステム中のいくつかの要素は、本明細書で詳細に説明する必要がない従来の周知の要素を含んでもよい。例えば、知能モジュールを、デスクトップパーソナルコンピュータ、ワークステーション、メインフレーム、ラップトップなどとして実装することができる。各クライアントシステムは、デスクトップパーソナルコンピュータ、ワークステーション、ラップトップ、P D A、携帯電話、又はインターネット若しくは他のネットワーク接続に直接若しくは間接にインターフェース接続することができる任意

50

のWAP対応デバイス若しくは他の任意のコンピュータデバイスを含んでもよい。クライアントシステムは、典型的には、クライアントシステムのユーザが、ネットワーク上の知能モジュールから入手可能な情報及びページにアクセスする、それを処理する、及び見ることが可能にする、HTTPクライアント、例えば、Microsoftのインターネットエクスプローラ（Internet Explorer）（商標）ブラウザ、Netscapeのナビゲーター（Navigator）（商標）ブラウザ、Operaのブラウザ、又は携帯電話、PDA若しくは他の無線デバイスの場合、WAP対応ブラウザなどのブラウジングプログラムを実行する。各クライアントシステムはまた、典型的には、知能モジュールによって提供されるページ、書式、及び他の情報と共に、ディスプレイ（例えば、モニタースクリーン、LCDディスプレイ）上でブラウザによって提供されるグラフィカルユーザインターフェース（GUI）と相互作用するための、キーボード、マウス、タッチスクリーン、ペンなどの1つ又は複数のユーザインターフェースデバイス（335）も含む。上述の通り、本発明は、ネットワークの特異的グローバルインターネットワークを指すインターネットと共に使用するのに好適である。しかしながら、イントラネット、エクストラネット、バーチャルプライベートネットワーク（VPN）、非TCP/IPに基づくネットワーク、任意のLAN又はWANなどの、他のネットワークをインターネットの代わりに使用することができることが理解されるべきである。

【0299】

[0313]一実施形態によれば、各クライアントシステム及びその成分のすべては、インテル（Intel）（登録商標）ペンティアム（Pentium）（登録商標）プロセッサなどの中央演算ユニットを用いるコンピュータコードの実行を含むブラウザなどのアプリケーションを用いて操作者により設定可能である。同様に、知能モジュール及びその成分のすべては、インテルペンティアムプロセッサなどの中央演算ユニット（315）又はマルチプロセッサユニットを用いるコンピュータコードの実行を含むアプリケーションを用いて操作者により設定可能であってもよい。本明細書に記載のデータ及び試験結果を処理するための知能モジュールを操作及び設定するためのコンピュータコードは、ハードディスク上にダウンロード及び保存されるのが好ましいが、全プログラムコード、又はその一部を、周知である他の任意の揮発性若しくは非揮発性記憶媒体若しくはデバイス、例えば、ROM若しくはRAM中に保存するか、又はコンパクトディスク（CD）媒体、デジタル多用途ディスク（DVD）媒体、フロッピーディスク、ROM、RAMなどの、プログラムコードを保存することができる他の任意のコンピュータ可読媒体（360）上に提供することもできる。

【0300】

[0314]本発明の種々の態様及び実施形態を実施するためのコンピュータコードを、例えば、R、C、C++、C#、HTML、Java、JavaScriptなどのコンピュータシステム上で実行することができる任意のプログラミング言語、又はVBScriptなどの他の任意のスクリプト言語中に実装することができる。Rプログラミング言語が好ましい。さらに、全プログラムコード、又はその一部を、周知のような任意の通信媒体及びプロトコル（例えば、TCP/IP、HTTP、HTTPS、イーサネット）を使用して、インターネット上で、又は周知のような他の任意の従来のネットワーク接続（例えば、エクストラネット、VPN、LAN）上でソフトウェアソース（例えばサーバー）から伝送及びダウンロードすることができるキャリア信号として具体化することができる。

【0301】

[0315]一実施形態によれば、知能モジュールは、患者の試験結果を分析してIBD又はCD若しくはUCなどのその臨床垂型の診断を決定するための疾患分類プロセスを実装する。このデータを、メモリー（310）中の1つ又は複数のデータ表又は他の論理データ構造物中に、あるいは知能モジュールと連結した他の記憶装置又はデータベースシステム中に保存することができる。典型的には、1つ又は複数の統計解析又はプロセスが、特定の患者に関する試験データを含むデータセットに適用される。例えば、試験データは、患

者に由来するサンプル中の少なくとも1つ又は一団のマーカの存在、レベル、及び／又は遺伝子型を示すデータを含む診断マーカープロファイルを含んでもよい。一実施形態では、ランダムフォレストなどの統計解析は特定の患者に関する試験データに適用され、試験データは、患者に由来するサンプル中で決定された少なくとも1つ又は一団のマーカの存在、レベル、及び／又は遺伝子型を含む。統計的に誘導された決定を、知能モジュールに結合若しくは連結したディスプレイデバイス上に表示させるか、又は決定を他のシステム、例えばクライアントシステム(330)に提供し、表示させることができる。特定の実施形態では、統計的に誘導された決定を、レポート又はプリントアウトの形態で表示してもよく、医師が、表示された結果を比較し、解釈して、合理的なIBD診断又は予測を行うことができるような参照表、チャート、グラフ又はモデルを含んでもよい。

10

【0302】

X. 療法及び治療モニタリング

[0316] IBDの診断が本明細書に記載の方法に従って分類又は予測されたら、本発明は、分類又は予測に基づいて療法の経過を推奨することをさらに含んでもよい。いくつかの例では、本発明は、IBD、CD、UC、IC、又は不確定IBDと関連する1つ又は複数の症状を処置するのに有用な治療有効量のIBD治療剤を個体に投与することをさらに含んでもよい。治療適用のためには、IBD治療剤を単独で投与するか、あるいは1つ若しくは複数のさらなるIBD治療剤及び／又はIBD治療剤と関連する副作用を軽減させる1つ若しくは複数の薬剤と共に同時投与することができる。IBD治療剤の例としては、限定されるものではないが、生物製剤、従来の薬剤、及びそれらの組み合わせが挙げられる。すなわち、本発明は、有利なことに、正しい薬剤が正しい時間に正しい患者に与えられるように、処置決定を指導し、IBDのための療法選択を通知することにより、医師が「オーダーメイド医療」を実施できるようにする。

20

【0303】

[0317] IBD治療剤は、必要に応じて好適な医薬賦形剤と共に投与することができ、許容される投与様式のいずれかによって実施することができる。すなわち、投与は、例えば、静脈内、局所、皮下、経皮膚、経皮、筋肉内、経口、頬、舌下、歯肉、口蓋、関節内、非経口、細動脈内、皮内、心室内、頭蓋内、腹腔内、病変内、鼻内、直腸、経膈、又は吸入によるものであってもよい。「同時投与する」とは、IBD治療剤が、第2の薬剤(例えば、他のIBD治療剤、IBD治療剤の副作用を軽減するのに有用な薬剤)の投与と同時に、その直前又は直後に投与されることを意味する。

30

【0304】

[0318] IBD治療剤の治療有効量を、例えば、少なくとも2、3、4、5、6、7、8回以上反復投与するか、又は用量を連続注入により投与することができる。用量は、好ましくは、正確な用量の単純な投与にとって好適な単位剤形中の、例えば、錠剤、ピル、ペレット、カプセル、粉末、溶液、懸濁液、乳濁液、坐剤、停留浣腸、クリーム、軟膏、ローション、ゲル、エアロゾル、気泡などの、固体、半固体、凍結乾燥粉末、又は液体剤形の形態を取ってもよい。

【0305】

[0319] 本明細書において、用語「単位剤形」は、各単位が、好適な医薬賦形剤(例えば、アンプル)と共に、所望の開始、許容性、及び／又は治療効果をもたらすように算出されたIBD治療剤の所定量を含有する、ヒト対象及び他の哺乳動物のための単一の用量として好適な物理的に個別の単位を含む。さらに、より濃縮された剤形を調製した後、それからより希薄な単位剤形を生成することができる。すなわち、より濃縮された剤形は、IBD治療剤の量よりも実質的に多い、例えば、その少なくとも1、2、3、4、5、6、7、8、9、10倍以上を含む。

40

【0306】

[0320] そのような剤形を調製するための方法は当業者に公知である(例えば、Remington's Pharmaceutical Sciences, 第18版, Mack Publishing Co., Easton, PA (1990)を参照)。剤形は

50

、従来の医薬担体又は賦形剤を典型的に含み、他の薬剤、担体、アジュバント、希釈剤、組織浸透促進剤、可溶化剤などをさらにも含む。好適な賦形剤を、当業界で周知の方法により、特定の剤形及び投与経路に合わせることができる（例えば、Remington's Pharmaceutical Sciences（上記）を参照）。

【0307】

[0321]好適な賦形剤の例としては、限定されるものではないが、ラクトース、デキストロース、スクロース、ソルビトール、マンニトール、デンプン、アカシアゴム、リン酸カルシウム、アルギネート、トラガカント、ゼラチン、ケイ酸カルシウム、微結晶性セルロース、ポリビニルピロリドン、セルロース、水、塩水、シロップ、メチルセルロース、エチルセルロース、ヒドロキシプロピルメチルセルロース、及びポリアクリル酸、例えば、カーボポール（Carbopol）、例えば、カーボポール941、カーボポール980、カーボポール981などが挙げられる。剤形は、タルク、ステアリン酸マグネシウム、及び鉱油などの潤滑剤；湿潤剤；乳化剤；懸濁剤；メチルー、エチルー、及びプロピルーヒドロキシ安息香酸（すなわちパラベン）などの保存剤；無機及び有機の酸及び塩基などのpH調整剤；甘味料；並びに香料をさらにも含む。剤形はまた、生分解性ポリマー、デキストラン、及びシクロデキストリン含有複合体を含んでもよい。

10

【0308】

[0322]経口投与のために、治療有効用量は、錠剤、カプセル、乳濁液、懸濁液、溶液、シロップ、スプレー、ロゼンジ剤、粉末、及び持続放出製剤の形態を取ってもよい。経口投与のための好適な賦形剤としては、医薬等級のマンニトール、ラクトース、デンプン、ステアリン酸マグネシウム、サッカリンナトリウム、タルカム、セルロース、グルコース、ゼラチン、スクロース、炭酸マグネシウムなどが挙げられる。

20

【0309】

[0323]一部の実施形態では、治療有効用量はピル、錠剤、又はカプセルの形態を取り、剤形は、IBD治療剤と共に次のいずれかを含んでもよい。ラクトース、スクロース、リン酸二カルシウムなどの希釈剤；デンプン又はその誘導体などの崩壊剤；ステアリン酸マグネシウムなどの潤滑剤；並びにデンプン、アカシアゴム、ポリビニルピロリドン、ゼラチン、セルロース及びその誘導体などの結合剤。IBD治療剤を、例えば、ポリエチレングリコール（PEG）担体中に配置される坐剤に製剤化することもできる。

30

【0310】

[0324]IBD治療剤及び任意選択で1つ又は複数の薬学的に許容されるアジュバントを、例えば、水性塩水（例えば0.9%w/v塩化ナトリウム）、水性デキストロース、グリセロール、エタノールなどの担体中に溶解又は分散させて、例えば、経口、局所、又は静脈内投与のための溶液又は懸濁液を形成させることによって液体剤形を調製することができる。IBD治療剤を滞留浣腸に製剤化することもできる。

【0311】

[0325]局所投与のために、治療有効用量は、乳濁液、ローション、ゲル、気泡、クリーム、ゼリー、溶液、懸濁液、軟膏、及び経皮パッチの形態を取ってもよい。吸入による投与のために、IBD治療剤を、乾燥粉末として、又はネブライザーを介して液体形態で送達することができる。非経口投与のために、治療有効用量は滅菌注射可能溶液及び滅菌包装粉末の形態を取ってもよい。注射可能溶液は、約4.5～約7.5のpHで製剤化されるのが好ましい。

40

【0312】

[0326]治療有効用量を凍結乾燥形態で提供することもできる。そのような剤形は、例えば、投与前の再構成のためのバッファー（例えば重炭酸塩）を含んでもよく、又はバッファーを、例えば、水を用いる再構成のために凍結乾燥剤形中に含有させてもよい。凍結乾燥剤形は、好適な血管収縮剤、例えば、エピネフリンをさらにも含む。凍結乾燥剤形は、再構成された剤形を個体にすぐに投与できるように、シリンジ（任意選択で、再構成のためのバッファーと共に包装される。）中に提供してもよい。

【0313】

50

[0327] I B D又はその臨床サブタイプの処置のための治療的使用において、I B D治療剤は、1日約0.001mg/kg～約1000mg/kgの初期用量で投与することができる。約0.01mg/kg～約500mg/kg、約0.1mg/kg～約200mg/kg、約1mg/kg～約100mg/kg、又は約10mg/kg～約50mg/kgの日用量範囲を用いることができる。しかしながら、用量は、個体の要件、I B D症状の重篤度、及び用いられるI B D治療剤に応じて変化してもよい。例えば、用量は、本明細書に記載の方法に従って特定の臨床サブタイプのI B Dを有すると分類された個体におけるI B D症状の型及び重症度を考慮して経験的に決定することができる。本発明の文脈における、個体に投与される用量は、時間と共に個体における有益な治療応答に影響するのに十分なものであるべきである。用量のサイズは、個体における特定のI B D治療剤の投与に伴う任意の有害な副作用の存在、性質及び程度により決定することもできる。特定の状況のための適切な用量の決定は医師の知見の範囲内にある。一般に、処置はI B D治療剤の最適用量より少ない、より小さい用量から開始される。その後、環境下での最適効果に達するまで用量を少しずつ増加させる。便宜的に総日用量を分割し、必要に応じて1日の間に一部ずつ投与することができる。

10

【0314】

[0328]本明細書において、用語「I B D治療剤」は、I B Dと関連する1つ又は複数の症状を処置するのに有用であるあらゆる薬学的に許容される形態の薬剤を含む。例えば、I B D治療剤は、ラセミ混合物又は異性体混合物、イオン交換樹脂に結合した固体複合体などの中にあってもよい。さらに、I B D治療剤は溶媒和形態を取ってもよい。この用語はまた、記載されるI B D治療剤のあらゆる薬学的に許容される塩、誘導体、及び類似体、並びにそれらの組み合わせを含むことが意図される。例えば、I B D治療剤の薬学的に許容される塩としては、限定されるものではないが、その酒石酸塩、コハク酸塩、酒石酸塩、重酒石酸塩、二塩酸塩、サリチル酸塩、ヘミコハク酸塩、クエン酸塩、マレイン酸塩、塩酸塩、カルバミン酸塩、硫酸塩、硝酸塩、及び安息香酸塩形態、並びにそれらの組み合わせが挙げられる。例えば、I B D治療剤の薬学的に許容される塩、I B D治療剤の遊離塩基、又はその混合物などの、任意の形態のI B D治療剤が、本発明の方法における使用にとって好適である。好適なI B D治療剤の例としては、限定されるものではないが、生物製剤、従来の薬剤、及びそれらの組み合わせが挙げられる。

20

【0315】

[0329]生物製剤としては、例えば、抗サイトカイン及びケモカイン抗体、例えば、抗腫瘍壊死因子アルファ(TNF α)抗体が挙げられる。抗TNF α 抗体の非限定例としては、キメラIgG1抗TNF α モノクローナル抗体であるインフリキシマブ(レミケード(Remicade)(登録商標))(Centocor, Inc, Horsham, PA)などのキメラモノクローナル抗体；CDP571及びPEG化CDP870などのヒト化モノクローナル抗体；アダリムマブ(ヒュミラ(Humira)(登録商標))(Abbott Laboratories, Abbott Park, IL)などの完全ヒトモノクローナル抗体；エタネルセプト(エンブレル(Enbrel)(登録商標))(Amgen, Thousand Oaks, CA；Wyeth Pharmaceuticals Inc, Collegeville, PA)などのp75融合タンパク質、小分子(例えばMAPキナーゼ阻害剤)；及びそれらの組み合わせが挙げられる。Ghosh, Novartis Found Symp., 263:193～205(2004)を参照されたい。

30

40

【0316】

[0330]他の生物製剤としては、例えば、細胞接着分子 α 4-インテグリンに対するヒト化モノクローナル抗体であるナタリズマブ(タイサブリ(Tysabri)(登録商標))(Elan Pharmaceuticals, Inc., Dublin, Ireland；Biogen Idec, Cambridge, MA)、及びヒト化IgG1抗 α 4 β 7インテグリンモノクローナル抗体であるMLN-02(Millennium Pharmaceuticals, Cambridge, MA)などの抗細胞接着抗体；抗

50

T細胞剤；ヒト化IgG2M3抗CD3モノクローナル抗体である、ビシリズマブ（ヌビオン（Nuviion）（登録商標））（PDL BioPharma, Inc. Incline Village, NV）などの抗CD3抗体；キメラ抗CD4モノクローナル抗体であるプリリキシマブ（cM-T412）（Centocor, Inc., Horsham, PA）などの抗CD4抗体；ヒト化IgG1抗CD25モノクローナル抗体であるダクリズマブ（ゼナパックス（Zenapax）（登録商標））（PDL BioPharma, Inc. Incline Village, NV；Roche, Nutley, NJ）及びキメラIgG1抗CD25モノクローナル抗体であるバシリキシマブ（シムレクト（Simullect）（登録商標））（Novartis, Basel, Switzerland）などの抗IL-2受容体アルファ（CD25）抗体；及びそれらの組み合わせが挙げられる。

10

【0317】

[0331]前記生物製剤に加えて、miR又はmiRの阻害剤が本発明において有用である。すなわち、いくつかの実施形態では、本発明は、有効量のi) miRNA阻害剤分子又はii) miRNA分子を、IBDを有する患者に導入するか、又は提供することによるIBDの治療又は予防を提供する。

【0318】

[0332]miRの送達のための1つの有用な製剤は、リボソームである。リボソーム及び乳濁液は、本発明の核酸を送達するのに用いることができる送達ビヒクルの周知の例である。本発明の核酸を、担体又は脂質と共に投与して、細胞による取込みを増加させることができる。例えば、オリゴヌクレオチドを、陽イオン脂質と共に投与することができる。陽イオン脂質の例としては、限定されるものではないが、リポフェクチン、DOTMA、DOPE、及びDOTAPが挙げられる。国際公開第00/71096号（参照により組み入れられる。）には、遺伝子療法のために有効に用いることができるDOTAP：コレステロール又はコレステロール誘導体製剤などの種々の製剤が記載されている。他の開示ではまた、ナノ粒子を含む種々の脂質又はリボソーム製剤及び投与方法が考察されている。そのようなものとしては、限定されるものではないが、米国特許出願公開第2003/0203865号、第2002/0150626号、第2003/0032615号、及び第2004/0048787号（これらは、それらが製剤並びに核酸の投与及び送達の他の関連する態様を開示する程度まで参照により組み入れられる。）が挙げられる。粒子を形成させるために用いられる方法は、米国特許第5844107号、第5877302号、第6008336号、第6077835号、第5972901号、第6200801号及び第5972900号（これらは、それらの態様に関して参照により組み入れられる。）にも開示されている。核酸を、ポリ（L-リシン）などの陽イオン性アミンと共に投与することもできる。

20

30

【0319】

[0333]従来の薬剤の例としては、限定されるものではないが、アミノサリチレート（例えば、メサラジン、スルファサラジン）、コルチコステロイド（例えばプレドニゾン）、チオプリン（例えば、アザチオプリン、6-メルカプトプリン）、メトトレキサート、その遊離塩基、薬学的に許容される塩、誘導体、類似体、及びそれらの組み合わせが挙げられる。

40

【0320】

[0334]当業者であれば、本発明における使用に好適なさらなるIBD治療剤が分かるであろう（例えば、Sands, Surg. Clin. North Am., 86:1045-1064（2006）；Daneseら, Mini Rev. Med. Chem., 6:771-784（2006）；Domenech, Digestion, 73（Suppl. 1）:67-76（2006）；Nakamuraら, World J. Gastroenterol., 12:4628-4635（2006）；及びGionchettiら, World J. Gastroenterol., 12:3306-3313（2006）を参照）。

50

【0321】

[0335] 診断及び／又は予測情報が個体のサンプルから得られたら、個体を定期的な時間隔でモニタリングして特定の治療レジメンの有効性を評価することもできる。例えば、特定のマーカーの存在又はレベルは、薬剤などの処置の治療効果に基づいて変化してもよい。いくつかの実施形態では、患者をモニタリングして、応答を評価し、個別化されたアプローチにおける特定の薬剤又は処置の効果を理解することができる。さらに、患者は薬剤に応答しなくてもよいが、マーカーは変化してもよく、これは、これらの患者がそのマーカーレベルによって同定することができる特殊な集団（非応答性）に属することを示唆する。これらの患者にはその現在の療法を中止させ、代替処置を処方することができる。

【0322】

X I. 実施例

[0336] 本発明を具体的な実施例によりさらに詳細に説明する。以下の実施例は説明のために提供されるものであり、本発明を限定することを意図したものではない。当業者であれば、変更又は改変しても本質的に同じ結果が得られる種々の重要でないパラメータを容易に理解するであろう。

【0323】

実施例 1（血清－遺伝子－炎症（s g i）マーカーを用いる I B D 診断のランダムフォレストモデリング）：

[0337] 本実施例は、1つ又は複数のランダムフォレストモデルを含む学習統計分類システムを訓練及び検証する方法を説明する。一部の実施形態では、疾患を診断するためのランダムフォレストモデルを作出する方法は、血清－遺伝子－炎症（s g i）マーカーのテキストベースの測定値のバイナリー変数への変換方法を含む。他の実施形態では、疾患を診断するためのランダムフォレストモデルを作出する方法は、リサンプリングによる不完全サンプル再利用及び性能誤差推定を含む段階的変数消去の方法を含む。本実施例はまた、モデルの予測性に対する種々の s g i マーカーの重要性を決定することを含む、疾患を診断するためのランダムフォレストモデルを作出する方法を説明する。

【0324】

[0338] 一部の実施形態では、I B D と関連する診断マーカーの予測モデリングのための本明細書に記載の統計的方法及びモデルは、モデル構築のランダムフォレスト方法を用いる。ランダムフォレストは、各木を構築する無作為成分を含む多くの決定木から作られる分類子である。各決定木は同じ分類問題に対処するが、異なる例の集合及び例のデータセットから無作為に選択される特徴の異なるサブセットが提供される。例えば、1本の木を、利用可能な例の無作為の 2 / 3 を、木の各ノードで分割される決定を行うために選択される特徴の無作為の 2 / 3 と共に用いて構築することができる。訓練の間にフォレストが構築されたら、すべての決定木にわたって採決することによって新しい例を分類する。2クラスのランダムフォレスト分類子に関する最も単純な事例では、最も投票数の多いクラスが勝つ。他の例では、訓練の間に確立された投票の勝利数に関するカットオフを予め設定して、性能尺度を最適化する。一部の例では、偽陽性が偽陰性よりもコストが高い場合、カットオフをより高く設定することができる。

【0325】

[0339] ランダムフォレスト分析は、非常に正確であるため、この種類の統計データをモデリングするのに非常に強固である。各決定木はデータの異なるサブセットから構築されるため、ランダムフォレストは他の機械学習アプローチよりもデータの過剰適合の傾向が少ない。個々の木による任意の過剰適合又は標準以下のものは、アンサンブル平均が計算された時に取消される傾向がある。

【0326】

[0340] いくつかの実施形態では、I B D 診断アルゴリズムは、血清－遺伝子－炎症（s g i）マーカーの既知の存在、非存在及び／又はレベルを有する予測された診断の後ろ向きコホートを用いて確立される。一部の実施形態では、ランダムフォレストを構築するために用いられるサンプル（例えば訓練及び検証データセット）は、本明細書に記載の方法

10

20

30

40

50

を用いる研究に由来するものである。簡単に述べると、訓練データセットはアルゴリズムを訓練するために用いられ、検証データセットは訓練されたアルゴリズムの性能を決定するために用いられる。ランダムフォレストモデリングの方法はマーカーの比較を可能にし、したがって、診断を予測するマーカーの重要性を決定することが可能となる。

【0327】

A. 血清－遺伝子－炎症マーカーのテキストベースの測定値のバイナリー変数への変換 [0341]一部の態様では、バイナリー値はサンプルのデータセットに基づいて定義される（表1参照）。いくつかの実施形態では、データセットは、IBD、CD、UC又は過敏性腸症候群（IBS）を有すると予測された対象、及び健常対照対象のデータを含む。

【0328】

表1：データセットに加えられる定義されたバイナリー列

【表 1】

新しい列名	バイナリー値	定義
DxIBD	1	診断列が IBD、CD 又は UC である場合
	0	健常、HC 又は IBS の場合

DxUC	1	診断が UC である場合
	0	CD の場合
pANCA	1	「ANCA IFA M3 (pANCA)」列が「検出されず」又は「DNase 耐性」である場合
	0	その他の場合(ブランクを除く)
GLI1	1	「GLI1-G933D-rs2228224」が TT である場合
	0	両方又は FAM の場合
	(ブランク)	不明の場合
MDR1	1	「MDR1 三対立遺伝子-rs2032582」が TT である場合
	0	AA、GA、GG、GT 又は TA である場合
	(ブランク)	不明の場合
ATG	1	「ATG16L1-T300A-rs2241880」が FAM である場合
	0	両方又は VIC の場合
	(ブランク)	不明の場合
ECM1	1	「ECM1-rs3737240」が FAM である場合
	0	両方又は VIC の場合
	(ブランク)	不明の場合
IRGM89	1	「IRGM-rs13361189」が VIC である場合
	0	両方又は FAM の場合
	(ブランク)	不明の場合
IRGM47	1	「IRGM-rs4958847」が VIC である場合
	0	両方又は FAM の場合
	(ブランク)	不明の場合
NKX2	1	「NKX2-3-rs10883365」が FAM である場合
	0	両方又は VIC の場合
	(ブランク)	不明の場合
MAGI2	1	「MAGI2 rs2160322」が VIC である場合
	0	両方又は FAM の場合
	(ブランク)	不明の場合
STAT3	1	「STAT3 rs744166」が VIC である場合
	0	両方又は FAM の場合
	(ブランク)	不明の場合

【0329】

B. リサンプリングによる不完全サンプル再利用及び性能誤差推定を含む段階的変数消去

[0342]他の実施形態では、疾患を診断するためのランダムフォレストモデルを作出する方法は、リサンプリングによる不完全サンプル再利用及び性能誤差推定を含む段階的変数消去の方法を含む。いくつかの実施形態では、マーカーが診断予測を改善することができるかどうかを決定するための方法が用いられる。一部の例では、すべての項目が入力されたデータが全マーカーに関して限られている場合、全マーカーが訓練中に使用される。そ

10

20

30

40

50

の後、あまり重要でないマーカーを消去した後、あまり重要でないマーカーを失っているサンプルを再利用し、訓練データセット中で使用する。

【0330】

[0343]一部の態様では、以下のステップの「変数消去プロトコール」が使用される。

- 1) データの不完全な行を一時的に消去する。
 - 2) データを訓練データセット（行の2／3）と試験データセット（行の1／3）に無作為に分割する（注：これを数回繰り返す）。
 - 3) 訓練データセットからランダムフォレストモデルを作出する。
 - 4) データセット上での予測を用いて性能を決定する。
 - a. ROCスコア
 - b. 目的関数を最大化するカットオフを用いる感度／特異度、ここで、Wは、0.3に設定された調整可能な重みである。
- $$W(1 - \text{感度})^2 + (1 - W)(1 - \text{特異度})^2$$
- 5) 所与のフォレストに関する各変数の重要度を計算する。
 - a. 所与のフォレストに関する1つの変数の値を無作為に順序変更する。
 - b. 結果として生じる正確度の低下を決定する。
 - c. すべての変数について繰り返す。
 - 6) これまでに実施されたデータの異なる無作為分割から標準誤差を計算する。
 - 7) 標準誤差が十分に小さいか、又は許容可能な最大反復回数を超えた場合、1つの変数を消去する。
 - a. 複数のデータ分割にわたる平均重要性が最小である変数を発見する。
 - b. 不完全な行の最後の消去の前のデータセットに戻る。
 - c. 「重要でない」変数を削除する。
 - d. 変数が残っている場合、プロトコールの始めに戻る。

【0331】

[0344]上記プロトコールは、どの血清－遺伝子－炎症（s g i）マーカーがIBD対非IBD診断の予測を改善することができるかを決定するために使用される。表2は、ASCA-A、IRGM47、IRGM89、MDR1、ECM1、VEGF、CRP、NKX2-3、STAT3、ATG16L1、SAA、GLI1、MAGI2、抗OmpA、ICAM、抗Fla2、ASCA-G、ANCA、抗FlaX、VCAM-1、抗Cbir1、抗OmpC、及びpANCAなどの23種の血清－遺伝子－炎症マーカーを用いるIBD対非IBD診断に関する変数重要性を例示する。IBD診断のランダムフォレストモデリングのこの例示的説明においては次の基準が要求された。別の試験データ上のROCスコアにおける最小リサンプリング数5、最大20、及び標準誤差目標0.01。用語「Least_IMP」とは、最も小さい平均重要性を有する変数を指す。用語「OOB_ROC」とは、訓練中のout-of-bag予測上で計算されたROCスコアを指す。用語「TEST_Roc」とは、フォレストを作出するのに使用されなかった、別の試験セット上でのROCスコアを指す。用語「SENS」及び「SPEC」とは、それぞれ、試験セット上での感度及び特異度を指す。用語「NTRAIN0」とは、0に等しいDxIBDを有する（すなわち非IBD）訓練サンプルの数を指す。用語「NTRAIN1」とは、1に等しいDxIBDを有する（すなわちIBD）訓練サンプルの数を指す。用語「NTEST0」及び「NTEST1」は試験サンプルの数である。注目すべきことに、以下の表中のすべての数は、複数の再分割にわたる平均であり、エラーバーは標準誤差（すなわち、平均における誤差の推定値）である。「NVAR」とは、IBD診断に対する血清－遺伝子－炎症マーカーの決定された重要性を指す。表2は、アルゴリズムにおいてpANCAが最も重要であり、ASCA-Aがあまり重要でないことを示す。

【0332】

表2：血清－遺伝子－炎症マーカーを使用するIBD対非IBDのマーカー重要性

【表 2】

	NVAR	OOB_ ROC	TEST_ ROC	SENS	SPEC	LEAST_ IMP	NTRAIN 0	NTRAIN 1	NTEST 0	NTEST 1
1	23	0.919+- 0.007	0.937+- 0.010	0.897+- 0.018	0.913+- 0.014	ASCAA	59.7+- 1.4	53.3+- 1.4	30.3+- 1.4	26.7+- 1.4
2	22	0.903+- 0.006	0.943+- 0.010	0.846+- 0.018	0.951+- 0.020	IRGM47	59.8+- 1.9	53.2+- 1.9	30.2+- 1.9	26.8+- 1.9
3	21	0.918+- 0.005	0.914+- 0.011	0.855+- 0.017	0.889+- 0.012	IRGM89	60.5+- 0.7	54.5+- 0.7	29.5+- 0.7	28.5+- 0.7
4	20	0.908+- 0.004	0.955+- 0.006	0.910+- 0.007	0.934+- 0.019	MDR1	60.0+- 1.4	57.0+- 1.4	30.0+- 1.4	29.0+- 1.4
5	19	0.917+- 0.006	0.941+- 0.010	0.904+- 0.015	0.942+- 0.012	ECM1	60.1+- 0.6	56.9+- 0.6	29.9+- 0.6	29.1+- 0.6
6	18	0.903+- 0.004	0.908+- 0.008	0.774+- 0.018	0.914+- 0.016	VEGF	87.2+- 2.2	281.8+- 2.2	41.8+- 2.2	143.2+- 2.2
7	17	0.893+- 0.003	0.903+- 0.009	0.770+- 0.011	0.906+- 0.009	CRP	132.2+- 1.4	283.8+- 1.4	61.8+- 1.4	146.2+- 1.4
8	16	0.899+- 0.004	0.900+- 0.006	0.807+- 0.020	0.877+- 0.011	NKX2	130.4+- 2.8	285.6+- 2.8	63.6+- 2.8	144.4+- 2.8
9	15	0.890+- 0.007	0.918+- 0.010	0.807+- 0.019	0.885+- 0.012	STAT3	133.4+- 2.0	283.6+- 2.0	60.6+- 2.0	148.4+- 2.0
10	14	0.897+- 0.004	0.910+- 0.006	0.760+- 0.014	0.923+- 0.013	ATG16L1	132.4+- 1.6	286.6+- 1.6	61.6+- 1.6	148.4+- 1.6
11	13	0.904+- 0.003	0.905+- 0.005	0.778+- 0.019	0.918+- 0.012	SAA	132.0+- 2.2	292.0+- 2.2	63.0+- 2.2	149.0+- 2.2
12	12	0.899+- 0.002	0.893+- 0.005	0.779+- 0.029	0.879+- 0.021	GLI1	125.6+- 1.6	306.4+- 1.6	69.4+- 1.6	146.6+- 1.6
13	11	0.902+- 0.006	0.906+- 0.006	0.778+- 0.018	0.910+- 0.012	MAGI2	129.4+- 2.4	307.6+- 2.4	70.6+- 2.4	148.4+- 2.4
14	10	0.905+- 0.004	0.907+- 0.004	0.758+- 0.014	0.899+- 0.017	OmpA	139.6+- 3.8	316.4+- 3.8	70.4+- 3.8	157.6+- 3.8
15	9	0.890+- 0.003	0.876+- 0.007	0.740+- 0.022	0.872+- 0.017	ICAM	162.8+- 1.6	323.2+- 1.6	85.2+- 1.6	158.8+- 1.6
16	8	0.886+- 0.004	0.890+- 0.007	0.763+- 0.024	0.894+- 0.025	Fla2	169.6+- 2.1	325.4+- 2.1	86.4+- 2.1	161.6+- 2.1
17	7	0.879+- 0.005	0.880+- 0.009	0.716+- 0.028	0.907+- 0.007	ASCAG	173.8+- 1.5	324.2+- 1.5	86.2+- 1.5	163.8+- 1.5
18	6	0.878+- 0.005	0.887+- 0.006	0.773+- 0.019	0.878+- 0.015	ANCA	173.4+- 1.8	324.6+- 1.8	86.6+- 1.8	163.4+- 1.8
19	5	0.868+- 0.005	0.893+- 0.007	0.739+- 0.014	0.917+- 0.008	Flax	170.2+- 2.0	327.8+- 2.0	89.8+- 2.0	160.2+- 2.0
20	4	0.868+- 0.006	0.869+- 0.007	0.741+- 0.013	0.898+- 0.013	VCAM	172.8+- 4.0	331.2+- 4.0	89.2+- 4.0	162.8+- 4.0
21	3	0.820+- 0.006	0.809+- 0.005	0.666+- 0.023	0.865+- 0.017	CBir1	254.2+- 1.0	442.8+- 1.0	120.8+- 1.0	228.2+- 1.0
22	2	0.786+- 0.004	0.773+- 0.009	0.592+- 0.016	0.861+- 0.008	OmpC	251.9+- 4.1	445.1+- 4.1	123.1+- 4.1	225.9+- 4.1
23	1	0.507+- 0.003	0.731+- 0.006	0.5036+- 0.014	0.913+- 0.004	pANCA	242.4+- 3.0	454.6+- 3.0	132.6+- 3.0	216.4+- 3.0

【0333】

[0345]一部の例では、あまり重要でない変数を削除し、行を回復させたので、訓練及び試験データセットのサイズが増大した。他の例では、試験データセット上のROCスコアは0.95という高さであった。いくつかの例では、2つのマーカー（すなわち、VCAM及び抗Flax）は、最も重要なマーカーの1つとして2つの血清学的マーカー（すなわち、ASCA-G及びASCA-A）に取って代わった。表3は、血清学的マーカーのみを用いる場合、pANCAがIBD対非IBDを予測するための最も重要なマーカーであり、このコホートではASCA-Gはあまり重要でなかったことを示す。いくつかの例

10

20

30

40

50

では、最良の試験ROCスコアは0.8であったが、これは、試験したマーカーがかなり誤差を低下させることができることを示している。

【0334】

表3：血清学的マーカーのみを使用するIBD対非IBDのマーカー重要性

【表3】

	NVAR	OOB_ ROC	TEST_ ROC	SENS	SPEC	LEAST_ IMP	NTRAIN 0	NTRAIN 1	NTEST 0	NTEST 1
1	6	0.846+- 0.007	0.850+- 0.010	0.699+- 0.019	0.893+- 0.010	ASCAG	248.3+- 3.1	448.7+- 3.1	126.7+- 3.1	222.3+- 3.1
2	5	0.845+- 0.005	0.843+- 0.008	0.707+- 0.013	0.853+- 0.010	ANCA	254.4+- 3.9	442.6+- 3.9	120.6+- 3.9	228.4+- 3.9
3	4	0.843+- 0.005	0.852+- 0.010	0.714+- 0.010	0.877+- 0.013	ASCAA	246.2+- 3.8	450.8+- 3.8	128.8+- 3.8	220.2+- 3.8
4	3	0.817+- 0.006	0.815+- 0.008	0.669+- 0.014	0.870+- 0.007	CBir1	248.8+- 1.9	448.2+- 1.9	126.2+- 1.9	222.8+- 1.9
5	2	0.776+- 0.009	0.784+- 0.010	0.618+- 0.014	0.876+- 0.014	OmpC	248.7+- 2.0	448.3+- 2.0	126.3+- 2.0	222.7+- 2.0
6	1	0.511+- 0.002	0.742+- 0.005	1.000+- 0.000	0.976+- 0.005	pANCA	248.0+- 3.0	449.0+- 3.0	127.0+- 3.0	222.0+- 3.0

10

【0335】

[0346]上記プロトコルを、UC対CDの予測を改善するのに有用であり得る血清一遺伝子一炎症マーカーを決定するために使用することもできる。表2は、ASCA-A、IRGM47、IRGM89、MDR1、ECM1、VEGF、CRP、NKX2-3、STAT3、ATG16L1、SAA、GLI1、MAGI2、抗OmpA、ICAM、抗Fla2、ASCA-G、ANCA、抗FlaX、VCAM-1、抗CBir1、抗OmpC、及びpANCAなどの23種の血清一遺伝子一炎症マーカーを用いるUC対CD診断のためのマーカー重要性を例示する。一部の例では、データが限られており、コホートは23種すべてのマーカー及びUCの診断を有するサンプルを含まない。いくつかの例では、各マーカーを、コホート中の診断数と比較した後、変数（例えばマーカー）を消去することができるかどうかを決定することができる（表4参照）。

20

【0336】

表4：各マーカーに関するUC診断数（NUC）及びCD診断数（NCD）

30

【表 4】

	マーカー	NUC	NCD
1	ANCA	435	236
2	pANCA	435	236
3	ASCAA	435	236
4	ASCAG	435	236
5	OmpC	435	236
6	CBir1	435	236
7	GLI1	433	243
8	MDR1	432	242
9	ATG16L1	430	243
10	ECM1	108	171
11	IRGM89	427	231
12	IRGM47	425	229
13	NKX2	402	202
14	MAGI2	401	200
15	STAT3	401	200
16	Flax	405	231
17	Fla2	405	232
18	OmpA	397	231
19	CRP	327	176
20	SAA	315	174
21	ICAM	322	177
22	VCAM	326	168
23	VEGF	427	236

10

20

【0337】

[0347]この例では、マーカー ECM1 は最も少ない数の UC 診断と関連し、それによって段階的変数消去プロトコールから消去された。本発明の方法及び UC 対 CD を予測するための 22 種の血清－遺伝子－炎症マーカーのデータセットを用いて、ランダムフォレスト分類子を構築することができる（表 5 参照）。

30

【0338】

表 5：ECM1 以外の 22 種のマーカーに関する UC 対 CD 予測のマーカー重要性

【表 5】

	NVAR	OOB_ ROC	TEST_ ROC	SENS	SPEC	LEAST_ IMP	NTRAIN 0	NTRAIN 1	NTEST 0	NTEST 1
1	22	0.932+- 0.005	0.944+- 0.007	0.877+- 0.013	0.913+- 0.012	STAT3	84.8+- 2.7	191.2+- 2.7	47.2+- 2.7	91.8+- 2.7
2	21	0.940+- 0.004	0.919+- 0.010	0.850+- 0.015	0.869+- 0.022	MDR1	88.3+- 0.8	189.7+- 0.8	46.7+- 0.8	93.3+- 0.8
3	20	0.939+- 0.004	0.928+- 0.009	0.838+- 0.012	0.901+- 0.011	GLI1	89.0+- 1.1	189.0+- 1.1	46.0+- 1.1	94.0+- 1.1
4	19	0.941+- 0.003	0.930+- 0.003	0.865+- 0.010	0.895+- 0.016	MAGI2	89.6+- 1.5	189.4+- 1.5	45.4+- 1.5	94.6+- 1.5
5	18	0.938+- 0.004	0.951+- 0.005	0.876+- 0.016	0.921+- 0.011	NKX2	90.2+- 2.1	189.8+- 2.1	46.8+- 2.1	94.2+- 2.1
6	17	0.940+- 0.004	0.954+- 0.004	0.868+- 0.017	0.930+- 0.006	IRGM47	100.2+- 2.3	187.8+- 2.3	48.8+- 2.3	96.2+- 2.3
7	16	0.938+- 0.003	0.937+- 0.007	0.882+- 0.013	0.903+- 0.010	VCAM	105.0+- 1.5	190.0+- 1.5	52.0+- 1.5	96.0+- 1.5
8	15	0.947+- 0.001	0.942+- 0.008	0.871+- 0.029	0.910+- 0.024	CRP	114.6+- 1.5	185.4+- 1.5	50.4+- 1.5	100.6+- 1.5
9	14	0.946+- 0.003	0.938+- 0.010	0.834+- 0.029	0.920+- 0.010	IRGM89	109.1+- 1.8	190.9+- 1.8	55.9+- 1.8	95.1+- 1.8
10	13	0.942+- 0.003	0.944+- 0.005	0.853+- 0.010	0.927+- 0.009	ATG16L1	110.4+- 1.6	195.6+- 1.6	60.6+- 1.6	93.4+- 1.6
11	12	0.944+- 0.003	0.944+- 0.009	0.888+- 0.009	0.894+- 0.014	VEGF	107.8+- 1.8	202.2+- 1.8	63.2+- 1.8	92.8+- 1.8
12	11	0.944+- 0.002	0.936+- 0.004	0.838+- 0.014	0.907+- 0.013	pANCA	115.4+- 3.1	198.6+- 3.1	55.6+- 3.1	101.4+- 3.1
13	10	0.943+- 0.005	0.947+- 0.006	0.869+- 0.027	0.916+- 0.008	CBir1	111.8+- 0.8	202.2+- 0.8	59.2+- 0.8	97.8+- 0.8
14	9	0.943+- 0.002	0.939+- 0.005	0.855+- 0.020	0.923+- 0.017	Fla2	116.8+- 1.6	197.2+- 1.6	54.2+- 1.6	102.8+- 1.6
15	8	0.937+- 0.004	0.944+- 0.010	0.836+- 0.018	0.929+- 0.010	OmpA	115.6+- 2.6	198.4+- 2.6	55.4+- 2.6	102.6+- 2.6
16	7	0.938+- 0.003	0.953+- 0.004	0.849+- 0.016	0.947+- 0.016	SAA	114.2+- 0.7	205.8+- 0.7	56.8+- 0.7	103.2+- 0.7
17	6	0.942+- 0.004	0.942+- 0.008	0.838+- 0.016	0.897+- 0.008	ANCA	112.7+- 2.5	215.3+- 2.5	61.3+- 2.5	103.7+- 2.5
18	5	0.938+- 0.004	0.946+- 0.006	0.859+- 0.008	0.900+- 0.013	Flax	116.8+- 1.9	211.2+- 1.9	57.2+- 1.9	107.8+- 1.9
19	4	0.946+- 0.004	0.934+- 0.007	0.892+- 0.023	0.889+- 0.013	ASCAA	117.0+- 2.7	215.0+- 2.7	60.0+- 2.7	107.0+- 2.7
20	3	0.913+- 0.005	0.917+- 0.008	0.820+- 0.020	0.899+- 0.007	ICAM	113.4+- 1.7	218.6+- 1.7	63.6+- 1.7	103.4+- 1.7
21	2	0.801+- 0.006	0.813+- 0.008	0.632+- 0.019	0.829+- 0.019	OmpC	158.2+- 2.6	288.8+- 2.6	77.8+- 2.6	146.2+- 2.6
22	1	0.754+- 0.007	0.748+- 0.007	0.584+- 0.014	0.773+- 0.009	ASCAG	157.6+- 1.7	289.4+- 1.7	78.4+- 1.7	145.6+- 1.7

【0 3 3 9】

【0348】表 5 は、A S C A - G が U C 対 C D を予測するための最も重要な血清 - 遺伝子 - 炎症マーカーであり、S T A T 3 はコホートのためのランダムフォレストモデルを用いた場合、あまり重要ではなかったことを示す。この例では、試験セット上の R O C スコアは 0.95 の高さであり、7 つのマーカーが最も高いマーカー重要性を有すると決定された（例えば、A S C A - G、I C A M、A S C A - A、抗 F l a X、A N C A 及び S A A）。注目すべきことに、上から 6 つの最も重要な変数における 2 つのマーカー（すなわち、V C A M 及び抗 F l a X）は 2 つの血清学的マーカーを置き換えた。

【0 3 4 0】

10

20

30

40

50

[0349] データは 22 種の全マーカーに関して限られているため、サンプルを訓練中に再利用し、あまり重要でない変数は削除した。血清学的マーカー（例えば、ASCA-G、pANCA、抗OmpC、ANCA、ASCA-A 及び抗CBir1）のみを用いる変数の重要性を、本明細書に記載の方法を用いて決定した（表 6 参照）。最良の試験 ROC スコアは 0.89 であったが、これは上から 7 つのマーカーについて見られた 0.95 のスコアよりも有意に低い。

【0341】

表 6：血清学的マーカーのみを使用する UC 対 CD 予測のマーカー重要性

【表 6】

	NVAR	OOB_ ROC	TEST_ ROC	SENS	SPEC	LEAST_ IMP	NTRAIN 0	NTRAIN 1	NTEST 0	NTEST 1
1	6	0.880+- 0.005	0.879+- 0.009	0.764+- 0.016	0.850+- 0.015	CBir1	156.0+- 1.9	291.0+- 1.9	80.0+- 1.9	144.0+- 1.9
2	5	0.878+- 0.006	0.890+- 0.009	0.768+- 0.028	0.860+- 0.014	ASCAA	157.0+- 2.0	290.0+- 2.0	79.0+- 2.0	145.0+- 2.0
3	4	0.842+- 0.006	0.870+- 0.009	0.771+- 0.014	0.821+- 0.013	ANCA	158.2+- 2.1	288.8+- 2.1	77.8+- 2.1	146.2+- 2.1
4	3	0.852+- 0.004	0.848+- 0.004	0.748+- 0.010	0.823+- 0.003	OmpC	161.6+- 2.8	285.4+- 2.8	74.4+- 2.8	149.6+- 2.8
5	2	0.801+- 0.006	0.805+- 0.007	0.628+- 0.024	0.822+- 0.023	pANCA	159.0+- 2.5	288.0+- 2.5	77.0+- 2.5	147.0+- 2.5
6	1	0.735+- 0.007	0.742+- 0.009	0.595+- 0.022	0.755+- 0.019	ASCAG	157.0+- 2.8	290.0+- 2.8	79.0+- 2.8	145.0+- 2.8

【0342】

[0350] sgi アルゴリズムのいくつかの部分において用いられた 17 種のマーカーのすべての可能な対の組み合わせに関する相関及び関係を算出した。連続変数対について、及び連続バイナリー変数対について、Pearson 相関係数及びその p 値を算出した。バイナリー-バイナリー変数対については、カイ二乗分析を用いて関連を評価した。すべての算出を、ソフトウェアプログラム Matlab を用いて行った。

【0343】

[0351] 結果を表 7 及び表 8 に示す。表 7 には、連続変数間及び連続変数とバイナリー変数と間のペアワイズ相関が、その p 値と共に列挙されている。テキストベースのデータから変換されたバイナリー変数間の関係はここでは示されないことに留意されたい。表 8 は、統計的に有意なペアワイズ相関を列挙する。42 の有意な相関が見出された。遺伝子マーカー、pANCA、又は pANCA 2 の間には関係はなかった。1 つを除くすべて（ANCA 及び ASCA-G）が正の相関を示した。

【0344】

C. sgi アルゴリズムにおいて使用されたマーカー間のペアワイズ相関

[0352] 連続変数、及び連続変数とバイナリー変数とのすべての対の組み合わせを用いて相関を算出した。Pearson 相関係数及び各相関係数に関する p 値を示す。p 値は、0 の相関係数と比較した統計的有意性を表す。0.0000 の p 値は 10^{-4} 未満である。

【0345】

表 7：すべてのユニークな対の組み合わせに関する相関

【表 7 - 1】

マーカー1	マーカー2	相関係数	P 値
ASCAA	ASCAG	0.7472	0.0000
ASCAA	ANCA	-0.0915	0.0004
ASCAA	OmpC	0.2191	0.0000
ASCAA	CBir1	0.2878	0.0000
ASCAA	Fla2	0.3742	0.0000
ASCAA	FlaX	0.4015	0.0000
ASCAA	pANCA	-0.0204	0.4270
ASCAA	VEGF	0.1004	0.0001
ASCAA	CRP	0.1160	0.0000
ASCAA	SAA	0.1451	0.0000
ASCAA	ICAM	0.0721	0.0049
ASCAA	VCAM	0.0714	0.0054
ASCAA	ATG16L1	0.0340	0.1847
ASCAA	NKX23	-0.0050	0.8443
ASCAA	ECM1	-0.0357	0.1641
ASCAA	STAT3	-0.0150	0.5601
ASCAA	pANCA2	-0.0798	0.0018
ASCAG	ANCA	-0.0987	0.0001
ASCAG	OmpC	0.1837	0.0000
ASCAG	CBir1	0.2932	0.0000
ASCAG	Fla2	0.3646	0.0000
ASCAG	FlaX	0.3882	0.0000
ASCAG	pANCA	-0.0448	0.0806
ASCAG	VEGF	0.0970	0.0002
ASCAG	CRP	0.1029	0.0001
ASCAG	SAA	0.1222	0.0000
ASCAG	ICAM	0.0543	0.0343
ASCAG	VCAM	0.0640	0.0125
ASCAG	ATG16L1	0.0259	0.3133
ASCAG	NKX23	0.0150	0.5595
ASCAG	ECM1	-0.0682	0.0078
ASCAG	STAT3	-0.0069	0.7868
ASCAG	pANCA2	-0.0858	0.0008
ANCA	OmpC	0.0177	0.4895
ANCA	CBir1	-0.0514	0.0451
ANCA	Fla2	-0.0555	0.0304
ANCA	FlaX	-0.0600	0.0193
ANCA	pANCA	0.6071	0.0000
ANCA	VEGF	0.0511	0.0462
ANCA	CRP	0.0654	0.0108
ANCA	SAA	0.0670	0.0090
ANCA	ICAM	0.0980	0.0001
ANCA	VCAM	0.0760	0.0030
ANCA	ATG16L1	0.0389	0.1291
ANCA	NKX23	0.0408	0.1117
ANCA	ECM1	0.0032	0.9022
ANCA	STAT3	-0.0383	0.1358
ANCA	pANCA2	0.8180	0.0000

10

20

30

40

【表 7 - 2】

マーカー1	マーカー2	相関係数	P 値
OmpC	CBir1	0.0798	0.0019
OmpC	Fla2	0.1332	0.0000
OmpC	FlaX	0.1451	0.0000
OmpC	pANCA	0.0398	0.1211
OmpC	VEGF	0.0763	0.0029
OmpC	CRP	0.0697	0.0066
OmpC	SAA	0.0827	0.0013
OmpC	ICAM	0.0784	0.0022
OmpC	VCAM	0.1000	0.0001
OmpC	ATG16L1	0.0198	0.4408
OmpC	NKX23	0.0133	0.6030
OmpC	ECM1	-0.0308	0.2308
OmpC	STAT3	-0.0189	0.4622
OmpC	pANCA2	0.0077	0.7634
CBir1	Fla2	0.7023	0.0000
CBir1	FlaX	0.7190	0.0000
CBir1	pANCA	-0.0419	0.1025
CBir1	VEGF	0.0819	0.0014
CBir1	CRP	0.0724	0.0048
CBir1	SAA	0.0771	0.0026
CBir1	ICAM	0.0556	0.0302
CBir1	VCAM	0.0119	0.6417
CBir1	ATG16L1	0.0260	0.3120
CBir1	NKX23	0.0327	0.2026
CBir1	ECM1	0.0121	0.6374
CBir1	STAT3	0.0093	0.7186
CBir1	pANCA2	-0.0598	0.0198
Fla2	FlaX	0.9530	0.0000
Fla2	pANCA	-0.0136	0.5975
Fla2	VEGF	0.1232	0.0000
Fla2	CRP	0.1097	0.0000
Fla2	SAA	0.1167	0.0000
Fla2	ICAM	0.0365	0.1545
Fla2	VCAM	0.0249	0.3328
Fla2	ATG16L1	0.0392	0.1269
Fla2	NKX23	0.0210	0.4139
Fla2	ECM1	-0.0107	0.6779
Fla2	STAT3	0.0064	0.8023
Fla2	pANCA2	-0.0482	0.0602
FlaX	pANCA	-0.0147	0.5659
FlaX	VEGF	0.1380	0.0000
FlaX	CRP	0.1305	0.0000
FlaX	SAA	0.1321	0.0000
FlaX	ICAM	0.0342	0.1827
FlaX	VCAM	0.0186	0.4697
FlaX	ATG16L1	0.0336	0.1898
FlaX	NKX23	0.0236	0.3569
FlaX	ECM1	-0.0119	0.6433
FlaX	STAT3	0.0084	0.7427
FlaX	pANCA2	-0.0562	0.0284

10

20

30

40

【表 7 - 3】

マーカー1	マーカー2	相関係数	P 値
pANCA	VEGF	0.0259	0.3132
pANCA	CRP	0.0659	0.0101
pANCA	SAA	0.0584	0.0227
pANCA	ICAM	0.0729	0.0045
pANCA	VCAM	0.0573	0.0254
VEGF	CRP	0.3037	0.0000
VEGF	SAA	0.2628	0.0000
VEGF	ICAM	0.1413	0.0000
VEGF	VCAM	0.1017	0.0001
VEGF	ATG16L1	0.0519	0.0429
VEGF	NKX23	0.0442	0.0853
VEGF	ECM1	0.0434	0.0904
VEGF	STAT3	-0.0102	0.6899
VEGF	pANCA2	0.0484	0.0591
CRP	SAA	0.8056	0.0000
CRP	ICAM	0.2810	0.0000
CRP	VCAM	0.2625	0.0000
CRP	ATG16L1	0.0023	0.9289
CRP	NKX23	0.0311	0.2255
CRP	ECM1	0.0253	0.3251
CRP	STAT3	0.0410	0.1102
CRP	pANCA2	0.0640	0.0126
SAA	ICAM	0.2487	0.0000
SAA	VCAM	0.2611	0.0000
SAA	ATG16L1	0.0347	0.1763
SAA	NKX23	0.0064	0.8046
SAA	ECM1	0.0090	0.7259
SAA	STAT3	0.0455	0.0765
SAA	pANCA2	0.0681	0.0079
ICAM	VCAM	0.6830	0.0000
ICAM	ATG16L1	0.0033	0.8979
ICAM	NKX23	0.0052	0.8410
ICAM	ECM1	0.0034	0.8945
ICAM	STAT3	0.0149	0.5607
ICAM	pANCA2	0.0741	0.0039
VCAM	ATG16L1	0.0087	0.7337
VCAM	NKX23	-0.0053	0.8370
VCAM	ECM1	0.0222	0.3873
VCAM	STAT3	-0.0003	0.9903
VCAM	pANCA2	0.0613	0.0169

【 0 3 4 6 】

[0353] 4 2 の統計的に有意な相関が、s g i アルゴリズムのいくつかの部分において用いられた 1 8 種の変数のすべての可能なユニークな対の組み合わせのサブセット間で同定された（表 8 参照）。相関係数及びその p 値を示す。p 値は、0 の相関係数と比較した統計的有意性を表す。0. 0 0 E + 0 0 の p 値は、M A T L A B ソフトウェアの場合、限界以下である。B o n f e r r o n i 相関を適用して、真の $p < 0. 0 5$ で統計的有意性について試験した。

【 0 3 4 7 】

表 8 : すべてのユニークな対の組み合わせの統計的に有意な相関

10

20

30

40

【表 8】

マーカー1	マーカー2	相関係数	P 値
Fla2	FlaX	0.9530	0.00E+00
pANCA2	ANCA	0.8180	0.00E+00
SAA	CRP	0.8056	0.00E+00
ASCAG	ASCAA	0.7472	1.35E-271
pANCA	pANCA2	0.7332	1.53E-256
CBir1	FlaX	0.7190	3.55E-242
Fla2	CBir1	0.7023	2.39E-226
VCAM	ICAM	0.6830	2.42E-209
ANCA	pANCA	0.6071	8.89E-154
ASCAA	FlaX	0.4015	5.59E-60
ASCAG	FlaX	0.3882	7.49E-56
ASCAA	Fla2	0.3742	1.01E-51
ASCAG	Fla2	0.3646	5.53E-49
VEGF	CRP	0.3037	8.51E-34
ASCAG	CBir1	0.2932	1.62E-31
ASCAA	CBir1	0.2878	2.25E-30
CRP	ICAM	0.2810	5.54E-29
SAA	VEGF	0.2628	1.98E-25
CRP	VCAM	0.2625	2.26E-25
SAA	VCAM	0.2611	4.14E-25
SAA	ICAM	0.2487	7.35E-23
ASCAA	OmpC	0.2191	5.60E-18
ASCAG	OmpC	0.1837	5.22E-13
OmpC	FlaX	0.1451	1.34E-08
ASCAA	SAA	0.1451	1.34E-08
VEGF	ICAM	0.1413	3.16E-08
VEGF	FlaX	0.1380	6.62E-08
OmpC	Fla2	0.1332	1.87E-07
FlaX	SAA	0.1321	2.37E-07
FlaX	CRP	0.1305	3.33E-07
VEGF	Fla2	0.1232	1.44E-06
ASCAG	SAA	0.1222	1.76E-06
Fla2	SAA	0.1167	5.07E-06
ASCAA	CRP	0.1160	5.77E-06
CRP	Fla2	0.1097	1.83E-05
ASCAG	CRP	0.1029	5.88E-05
VEGF	VCAM	0.1017	7.09E-05
VEGF	ASCAA	0.1004	8.78E-05
OmpC	VCAM	0.1000	9.46E-05
ANCA	ICAM	0.0980	1.30E-04
VEGF	ASCAG	0.0970	1.53E-04
ANCA	ASCAG	-0.0987	1.16E-04

【 0 3 4 8 】

【0354】本実施例は、本発明によるランダムフォレストモデリングが、I B D 対非 I B D を診断し、U C と C D とを鑑別するのに特に有用であることを実証する。

【 0 3 4 9 】

実施例 2 (I B D 診断のための 3 クラスランダムフォレストモデリング) :

10

20

30

40

50

[0355] 本実施例は、訓練セットが健常対照からのサンプルを含まない場合の I B D 診断のランダムフォレストモデリングの方法を例示する。本実施例は、健常対照サンプルを含まないデータセットを用いてアルゴリズムを訓練することを例示する。本実施例はまた、訓練コホートが I B D 及び過敏性腸症候群 (I B S) 患者を含む、 I B D 対非 I B D を予測するための 7 種の血清学的マーカーのランダムフォレストモデリングも例示する。本実施例は、非 I B D、C D 及び U C などの、3 つの疾患分類を予測するためのランダムフォレストモデリングをさらに例示する。本実施例はまた、 I B S 対健常対象を予測するためのランダムフォレストモデルも例示する。

【 0 3 5 0 】

A. 健常対照を含まないランダムフォレストモデリング

[0356] 一部の実施形態では、複数のランダムフォレストモデルの I B D 診断アルゴリズムを、血清－遺伝子－炎症マーカーの既知の存在、非存在及び／又はレベルを含む予測された診断の後ろ向きコホートに由来するデータセットを用いて訓練及び検証する。いくつかの例では、データセットは、 I B D、C D、U C 又は過敏性腸症候群 (I B S) を有すると予測された対象のデータを含む (表 9 参照)。いくつかの例では、健常対照患者に由来するデータはランダムフォレストモデリングから除外される。一部の例では、データセットは、これらに限定されないが、A S C A－A、A S C A－G、I R G M 4 7、I R G M 8 9、M A G I 2、M D R 1、抗 O m p A、抗 O m p C、抗 C B i r 1、抗 F l a 2、抗 F l a X、V E G F、C R P、S A A、I C A M、V C A M、A N C A、p A N C A、p A N C A 2、A T G 1 6 L 1、E C M 1、S T A T 3、及び N K X 2－3 などの 2 3 種のマーカーを含む。表 9 は、 I B D 及び I B S 患者のみ (健常対照を含まない) を用いる 2 3 種の血清－遺伝子－炎症マーカーの I B D 対非 I B D 診断に関する変数 (例えばマーカー) 重要性を例示する。 I B D 診断のランダムフォレストモデリングのこの例示的説明においては次の基準が要求された。別の試験データ上での R O C スコアにおける最小リサンプリング数 5、最大 2 0、及び標準誤差目標 0. 0 1。用語「L e a s t _ I M P」とは、最も小さい平均重要性を有する変数を指す。用語「O O B _ R O C」とは、訓練中の o u t－o f－b a g 予測上で計算された R O C スコアを指す。用語「T E S T _ R o c」とは、フォレストを作出するのに使用されなかった、別の試験セット上での R O C スコアを指す。用語「S E N S」及び「S P E C」とは、それぞれ、試験セット上での感度及び特異度を指す。用語「N T R A I N 0」とは、0 に等しい D x I B D を有する (すなわち非 I B D) 訓練サンプルの数を指す。用語「N T R A I N 1」とは、1 に等しい D x I B D を有する (すなわち I B D) 訓練サンプルの数を指す。用語「N T E S T 0」及び「N T E S T 1」は試験サンプルの数である。注目すべきことに、以下の表中のすべての数は、複数の再分割にわたる平均であり、エラーバーは標準誤差 (すなわち、平均における誤差の推定値) である。表 9 は、ランダムフォレストモデルをコホートに由来するサンプルを用いずに構築する場合、 I B D 対非 I B D を予測するための最も重要な血清－遺伝子－炎症マーカーは p A N C A であり、あまり重要ではないマーカーは A S C A－A であることを示す。

【 0 3 5 1 】

表 9： I B D 及び I B S 患者のみを使用する 2 3 種の血清－遺伝子－炎症マーカーに関する I B D 対非 I B D のためのモデリング

【表 9】

	NV AR	OOB_ ROC	TEST_ ROC	SENS	SPEC	LEAST_ IMP	NTRAIN 0	NTRAIN 1	NTEST 0	NTEST 1	NRE SAMPLE	RAN SEED
1	23	0.919+- 0.007	0.937+- 0.010	0.897+- 0.018	0.913+- 0.014	ASCAA	59.7+- 1.4	53.3+- 1.4	30.3+- 1.4	26.7+- 1.4	7	1008
2	22	0.903+- 0.006	0.943+- 0.010	0.846+- 0.018	0.951+- 0.020	IRGM47	59.8+- 1.9	53.2+- 1.9	30.2+- 1.9	26.8+- 1.9	5	1013
3	21	0.918+- 0.005	0.914+- 0.011	0.855+- 0.017	0.889+- 0.012	IRGM89	60.5+- 0.7	54.5+- 0.7	29.5+- 0.7	28.5+- 0.7	20	1033
4	20	0.908+- 0.004	0.955+- 0.006	0.910+- 0.007	0.934+- 0.019	MDR1	60.0+- 1.4	57.0+- 1.4	30.0+- 1.4	29.0+- 1.4	5	1038
5	19	0.917+- 0.006	0.941+- 0.010	0.904+- 0.015	0.942+- 0.012	ECM1	60.1+- 0.6	56.9+- 0.6	29.9+- 0.6	29.1+- 0.6	8	1046
6	18	0.903+- 0.004	0.908+- 0.008	0.774+- 0.018	0.914+- 0.016	VEGF	87.2+- 2.2	281.8+- 2.2	41.8+- 2.2	143.2+- 2.2	5	1051
7	17	0.904+- 0.005	0.902+- 0.010	0.799+- 0.025	0.887+- 0.013	NKX2	89.0+- 2.1	283.0+- 2.1	40.0+- 2.1	147.0+- 2.1	6	1057
8	16	0.897+- 0.004	0.909+- 0.008	0.789+- 0.007	0.914+- 0.012	CRP	83.2+- 1.5	290.8+- 1.5	45.8+- 1.5	141.2+- 1.5	5	1062
9	15	0.899+- 0.004	0.917+- 0.009	0.797+- 0.030	0.892+- 0.011	STAT3	86.6+- 2.0	287.4+- 2.0	42.4+- 2.0	144.6+- 2.0	5	1067
10	14	0.912+- 0.003	0.896+- 0.008	0.766+- 0.007	0.892+- 0.013	SAA	89.2+- 1.9	286.8+- 1.9	39.8+- 1.9	148.2+- 1.9	5	1072
11	13	0.901+- 0.005	0.920+- 0.004	0.801+- 0.025	0.902+- 0.027	GLI1	85.4+- 1.9	298.6+- 1.9	43.6+- 1.9	148.4+- 1.9	5	1077
12	12	0.903+- 0.002	0.909+- 0.009	0.800+- 0.022	0.896+- 0.016	ATG16L1	85.8+- 1.7	300.2+- 1.7	43.2+- 1.7	149.8+- 1.7	6	1083
13	11	0.909+- 0.003	0.903+- 0.009	0.810+- 0.011	0.876+- 0.014	MAGI2	85.3+- 2.2	304.7+- 2.2	44.7+- 2.2	151.3+- 2.2	10	1093
14	10	0.906+- 0.005	0.912+- 0.008	0.816+- 0.017	0.883+- 0.012	ICAM	83.6+- 1.4	318.4+- 1.4	46.4+- 1.4	155.6+- 1.4	5	1098
15	9	0.896+- 0.005	0.914+- 0.006	0.774+- 0.022	0.921+- 0.008	Flax	88.0+- 4.0	317.0+- 4.0	42.0+- 4.0	161.0+- 4.0	5	1103
16	8	0.909+- 0.003	0.899+- 0.009	0.775+- 0.018	0.902+- 0.012	VCAM	86.7+- 1.2	319.3+- 1.2	43.3+- 1.2	160.7+- 1.2	10	1113
17	7	0.904+- 0.002	0.910+- 0.006	0.792+- 0.017	0.886+- 0.012	ANCA	88.2+- 1.5	416.8+- 1.5	43.8+- 1.5	209.2+- 1.5	5	1118
18	6	0.892+- 0.005	0.905+- 0.005	0.785+- 0.022	0.902+- 0.011	ASCAG	86.8+- 1.2	418.2+- 1.2	45.2+- 1.2	207.8+- 1.2	5	1123
19	5	0.884+- 0.006	0.899+- 0.009	0.797+- 0.017	0.888+- 0.016	OmpA	86.7+- 1.9	418.3+- 1.9	45.3+- 1.9	207.7+- 1.9	6	1129
20	4	0.853+- 0.005	0.844+- 0.004	0.678+- 0.013	0.873+- 0.012	OmpC	116.4+- 0.4	424.6+- 0.4	58.6+- 0.4	212.4+- 0.4	5	1134
21	3	0.802+- 0.009	0.827+- 0.009	0.672+- 0.025	0.855+- 0.023	Fla2	114.2+- 3.6	426.8+- 3.6	60.8+- 3.6	210.2+- 3.6	5	1139
22	2	0.746+- 0.007	0.754+- 0.008	0.621+- 0.039	0.846+- 0.025	CBir1	117.8+- 2.0	448.2+- 2.0	60.2+- 2.0	222.8+- 2.0	5	1144
23	1	0.630+- 0.004	0.735+- 0.006	0.502+- 0.018	0.874+- 0.022	pANCA	120.0+- 3.0	446.0+- 3.0	58.0+- 3.0	225.0+- 3.0	5	1149

【0352】

【0357】他の例では、IBD対非IBDのためのランダムフォレストモデルを、IBD及びIBS患者に由来するサンプルのみを用いて血清学的マーカーのみについて決定した。表10は、血清学的マーカーのIBD対非IBD診断に関する変数重要性を例示する。IBD対非IBDを予測するためには、pANCAが最も重要であり、ASCA-Aはあまり重要ではない。

【0353】

表10：血清学的マーカーのみ並びにIBD及びIBS患者のみを使用するIBD対非IBDのためのモデリング

10

20

30

40

50

【表 10】

	NV AR	OOB_ ROC	TEST_ ROC	SENS	SPEC	LEAST_ IMP	NTRAIN 0	NTRAIN 1	NTEST 0	NTEST 1	NRE SAMPLE	RAN SEED
1	6	0.842+- 0.006	0.864+- 0.004	0.766+- 0.011	0.872+- 0.009	ASCAA	119.0+- 4.1	447.0+- 4.1	59.0+- 4.1	224.0+- 4.1	5	1006
2	5	0.852+- 0.005	0.827+- 0.010	0.669+- 0.034	0.870+- 0.006	ANCA	120.3+- 2.5	445.7+- 2.5	57.7+- 2.5	225.3+- 2.5	6	1012
3	4	0.828+- 0.006	0.844+- 0.007	0.701+- 0.014	0.868+- 0.017	ASCAG	114.6+- 3.1	451.4+- 3.1	63.4+- 3.1	219.6+- 3.1	5	1017
4	3	0.818+- 0.004	0.825+- 0.008	0.686+- 0.026	0.871+- 0.018	CBir1	120.4+- 3.4	445.6+- 3.4	57.6+- 3.4	225.4+- 3.4	5	1022
5	2	0.767+- 0.010	0.781+- 0.009	0.633+- 0.022	0.855+- 0.006	OmpC	116.7+- 2.5	449.3+- 2.5	61.3+- 2.5	221.7+- 2.5	6	1028
6	1	0.629+- 0.006	0.731+- 0.009	0.508+- 0.021	0.871+- 0.019	pANCA	121.8+- 2.8	444.2+- 2.8	56.2+- 2.8	226.8+- 2.8	5	1033

【0354】

B. 疾患の3つの分類のランダムフォレストモデリング

[0358] アルゴリズムの性能を試験するために、3クラス予測（例えば、非IBD、CD及びUC）を用いる単一モデルを開発した。「変数消去プロトコル（Variable Elimination Protocol）」に記載の変数消去及びリサンプリングの方法を、予測のための2つのカットオフ値が確立されるように調整した。一方のカットオフ値をIBD対非IBDに対して使用し、IBDが決定された場合は他方をUC対CDに対して使用した。いくつかの例では、完全な試験データセットをIBD対非IBDモデルのために用いて、非IBD確率カットオフを見出すことができ、それより上は非IBDと予測され、それより下はアルゴリズムが第2のカットオフに続く。第1のカットオフ値を使用して、IBD対非IBDに関するROC値を決定することができる。第2のカットオフを決定するために、既知のIBD患者の試験データセットを使用して、 $p(UC) / (p(CD) + (UC))$ （式中、 $p(UC)$ 及び $p(CD)$ はそれぞれ、UC及びCDに関するモデル確率である。）を見出すことができる。第2のカットオフ値を使用して、UC対CDに関するROC値を決定することができる。2つのカットオフに基づく3クラス（例えば、非IBD、CD、UC）の単一混同マトリックスを使用して、各試験サンプルの診断を予測した。

【0355】

[0359] 全診断ドメインについて「最小重要」変数が選択、消去される単一の3クラスモデル上で段階的変数消去を実施した。3クラスランダムフォレストモデルを、各サンプルについて非IBD、UC及びCDを予測するために7種の血清学的マーカーから構築した（表11参照）。表11において、1行目は7種全部のマーカーを含み、2行目は最小重要マーカーを除くすべてなどを含む。用語「IBD__TEST__ROC」とは、非IBD試験対／IBDカットオフ（例えば第1のカットオフ）及び全試験データセットを用いることを指す。用語「UC__TEST__ROC」とは、UC対CDカットオフ及びIBDであることが知られる試験データセットのサンプルのみを用いることを指す。一部の例では、残りの感度及び選択性は全試験データセット上で決定される。特定の例では、表11の最終行におけるIBD性能は、単一バイナリー変数上のランダムフォレスト方法に関する統計プログラムRにおけるアーチファクトであり、さらに考慮すべきではない。

【0356】

表11： 7種の血清－遺伝子－炎症マーカーに関する非IBD、UC又はCDを予測するための3クラスランダムフォレストモデル

【表 1 1】

	NV ARTEST_ ROC	IBD_ TEST_ ROC	UC_ TEST_ ROC	IBD_ SENS	IBD_ SPEC	CD_ SENS	CD_ SPEC	UC_ SENS	UC_ SPEC	LEAST_ IMP	NTRAIN_ 0	NTRAIN_ 1	NTRAIN_ 2	NTEST_ 0	NTEST_ 1	NTEST_ 2
1	6	0.850+- 0.008	0.872+- 0.010	0.754+- 0.018	0.836+- 0.009	0.601+- 0.031	0.871+- 0.007	0.616+- 0.026	0.912+- 0.007	ANCA	249.4+- 2.9	155.3+- 2.6	292.3+- 4.3	125.6+- 2.9	80.7+- 2.6	142.7+- 4.3
2	5	0.843+- 0.007	0.883+- 0.009	0.721+- 0.012	0.859+- 0.014	0.608+- 0.020	0.909+- 0.016	0.612+- 0.015	0.911+- 0.008	ASCAA	250.7+- 3.6	156.1+- 3.1	290.1+- 3.8	124.3+- 3.6	79.9+- 3.1	144.9+- 3.8
3	4	0.833+- 0.009	0.852+- 0.007	0.726+- 0.012	0.869+- 0.007	0.606+- 0.027	0.905+- 0.008	0.608+- 0.015	0.915+- 0.008	CBir1	248.2+- 2.9	161.7+- 2.2	287.2+- 5.0	126.8+- 2.9	74.3+- 2.2	147.8+- 5.0
4	3	0.784+- 0.005	0.864+- 0.006	0.669+- 0.014	0.848+- 0.018	0.536+- 0.017	0.893+- 0.009	0.567+- 0.020	0.925+- 0.005	OmpC	249.6+- 2.3	157.8+- 1.9	289.6+- 2.9	125.4+- 2.3	78.2+- 1.9	145.4+- 2.9
5	2	0.763+- 0.006	0.785+- 0.010	0.613+- 0.010	0.831+- 0.012	0.434+- 0.022	0.862+- 0.008	0.495+- 0.015	0.928+- 0.008	ASCAG	247.7+- 2.2	159.7+- 1.6	289.7+- 2.7	127.3+- 2.2	76.3+- 1.6	145.3+- 2.7
6	1	0.746+- 0.006	0.691+- 0.010	1.000+- 0.000	0.000+- 0.000	0.356+- 0.135	0.685+- 0.119	0.835+- 0.063	0.439+- 0.166	pANCA	250.1+- 2.5	158.1+- 2.3	288.8+- 2.6	124.9+- 2.5	77.9+- 2.3	146.2+- 2.6

10

【0 3 5 7】

[0360]また、3クラスランダムフォレストモデルを、各サンプルについて非IBD、UC及びCDを予測するために22種の血清－遺伝子－炎症マーカーを用いて構築した。表12は、このモデルにおいて、最も重要なマーカーはpANCAであり、あまり重要でないマーカーはMDR1であることを示す。

【0 3 5 8】

表12： 22種の血清－遺伝子－炎症マーカーに関する非IBD、UC又はCDを予測するための3クラスランダムフォレストモデル

20

【表 1 2】

	NV ART TEST ROC	IBD TEST ROC	UC TEST ROC	IBD SENS	IBD SPEC	CD SENS	CD SPEC	UC SENS	UC SPEC	LEAST IMP	NTRAIN 0	NTRAIN 1	NTRAIN 2	NTEST 0	NTEST 1	NTEST 2
1	22	0.922+- 0.006	0.933+- 0.006	0.830+- 0.027	0.894+- 0.017	0.814+- 0.029	0.903+- 0.021	0.691+- 0.016	0.936+- 0.010	MDR1	84.2+- 1.9	90.6+- 2.2	186.2+- 2.0	42.8+- 1.9	41.4+- 2.2	96.8+- 2.0
2	21	0.920+- 0.009	0.928+- 0.009	0.847+- 0.018	0.877+- 0.011	0.814+- 0.010	0.930+- 0.013	0.723+- 0.022	0.899+- 0.003	CRP	83.4+- 0.7	88.8+- 1.8	188.8+- 2.2	43.6+- 0.7	43.2+- 1.8	94.2+- 2.2
3	20	0.909+- 0.010	0.924+- 0.009	0.823+- 0.016	0.865+- 0.016	0.792+- 0.016	0.888+- 0.009	0.666+- 0.022	0.922+- 0.011	STAT3	84.4+- 1.5	87.9+- 1.3	188.7+- 1.5	42.6+- 1.5	44.1+- 1.3	94.3+- 1.5
4	19	0.911+- 0.006	0.936+- 0.007	0.818+- 0.014	0.866+- 0.009	0.813+- 0.029	0.904+- 0.012	0.675+- 0.025	0.920+- 0.012	GLI1	88.4+- 1.7	92.0+- 2.2	182.6+- 1.7	38.6+- 1.7	43.0+- 2.2	100.4+- 1.7
5	18	0.915+- 0.009	0.932+- 0.007	0.810+- 0.015	0.902+- 0.013	0.804+- 0.031	0.914+- 0.005	0.665+- 0.013	0.926+- 0.008	NKX2	85.5+- 2.4	86.8+- 2.5	191.7+- 3.9	41.5+- 2.4	48.2+- 2.5	92.3+- 3.9
6	17	0.916+- 0.007	0.937+- 0.010	0.818+- 0.026	0.894+- 0.015	0.808+- 0.017	0.907+- 0.014	0.667+- 0.018	0.925+- 0.007	ATG16 L1	85.6+- 2.5	90.7+- 2.1	187.7+- 1.3	41.4+- 2.5	44.3+- 2.1	96.3+- 1.3
7	16	0.933+- 0.009	0.933+- 0.009	0.855+- 0.013	0.903+- 0.024	0.819+- 0.017	0.914+- 0.011	0.715+- 0.022	0.913+- 0.012	IRGM47	86.2+- 2.1	90.0+- 2.2	189.8+- 2.1	41.8+- 2.1	45.0+- 2.2	97.2+- 2.1
8	15	0.918+- 0.008	0.920+- 0.009	0.819+- 0.018	0.895+- 0.009	0.800+- 0.031	0.915+- 0.014	0.680+- 0.015	0.917+- 0.014	MAGI2	86.7+- 1.6	93.7+- 2.4	190.7+- 3.1	41.3+- 1.6	45.3+- 2.4	99.3+- 3.1
9	14	0.909+- 0.009	0.935+- 0.009	0.819+- 0.015	0.880+- 0.018	0.794+- 0.024	0.918+- 0.009	0.685+- 0.017	0.916+- 0.009	IRGM89	84.6+- 2.2	104.5+- 1.7	193.9+- 2.1	43.4+- 2.2	52.5+- 1.7	96.1+- 2.1
10	13	0.911+- 0.010	0.933+- 0.005	0.830+- 0.015	0.890+- 0.015	0.792+- 0.018	0.920+- 0.007	0.688+- 0.016	0.903+- 0.010	VEGF	87.9+- 1.6	106.8+- 1.5	196.3+- 2.3	42.1+- 1.6	55.2+- 1.5	98.7+- 2.3
11	12	0.907+- 0.003	0.930+- 0.005	0.794+- 0.017	0.888+- 0.013	0.748+- 0.024	0.916+- 0.010	0.675+- 0.025	0.937+- 0.007	SAA	138.8+- 2.7	106.8+- 1.9	202.4+- 1.5	71.2+- 2.7	55.2+- 1.9	97.6+- 1.5
12	11	0.911+- 0.008	0.929+- 0.010	0.813+- 0.019	0.895+- 0.010	0.758+- 0.020	0.918+- 0.010	0.684+- 0.028	0.923+- 0.009	Fla2	138.7+- 2.6	111.0+- 2.0	206.3+- 2.7	71.3+- 2.6	54.0+- 2.0	102.7+- 2.7
13	10	0.910+- 0.003	0.923+- 0.009	0.825+- 0.007	0.884+- 0.006	0.779+- 0.015	0.907+- 0.009	0.691+- 0.013	0.933+- 0.010	OmpA	141.7+- 3.3	110.3+- 2.6	206.0+- 2.9	71.3+- 3.3	54.7+- 2.6	104.0+- 2.9
14	9	0.902+- 0.005	0.933+- 0.009	0.802+- 0.010	0.862+- 0.013	0.777+- 0.020	0.916+- 0.011	0.667+- 0.029	0.918+- 0.005	ANCA	166.8+- 1.2	111.2+- 0.7	212.0+- 1.8	85.2+- 1.2	53.8+- 0.7	106.0+- 1.8
15	8	0.903+- 0.008	0.947+- 0.009	0.798+- 0.010	0.896+- 0.007	0.788+- 0.027	0.929+- 0.006	0.677+- 0.016	0.936+- 0.005	VCAM	167.8+- 2.2	110.8+- 1.9	211.4+- 2.9	84.2+- 2.2	54.2+- 1.9	106.6+- 2.9
16	7	0.889+- 0.006	0.933+- 0.009	0.793+- 0.017	0.872+- 0.016	0.787+- 0.019	0.933+- 0.008	0.660+- 0.023	0.911+- 0.010	Flax	171.1+- 1.3	114.9+- 1.8	213.0+- 1.9	84.9+- 1.3	59.1+- 1.8	106.0+- 1.9
17	6	0.877+- 0.005	0.931+- 0.007	0.747+- 0.018	0.892+- 0.008	0.732+- 0.016	0.938+- 0.003	0.642+- 0.026	0.932+- 0.010	ICAM	173.8+- 1.2	117.6+- 4.6	213.6+- 4.4	85.2+- 1.2	59.4+- 4.6	108.4+- 4.4
18	5	0.836+- 0.005	0.866+- 0.010	0.722+- 0.006	0.846+- 0.011	0.579+- 0.028	0.894+- 0.011	0.602+- 0.022	0.906+- 0.004	ASCAA	247.6+- 5.3	155.0+- 4.3	294.4+- 4.0	127.4+- 5.3	81.0+- 4.3	140.6+- 4.0
19	4	0.825+- 0.009	0.842+- 0.007	0.715+- 0.017	0.841+- 0.008	0.592+- 0.015	0.886+- 0.008	0.576+- 0.022	0.907+- 0.009	CBir1	248.6+- 2.9	158.1+- 1.8	290.3+- 2.4	126.4+- 2.9	77.9+- 1.8	144.7+- 2.4
20	3	0.781+- 0.006	0.835+- 0.008	0.642+- 0.012	0.846+- 0.020	0.457+- 0.016	0.890+- 0.016	0.555+- 0.012	0.917+- 0.006	OmpC	252.4+- 1.3	157.4+- 2.1	287.2+- 2.5	122.6+- 1.3	78.6+- 2.1	147.8+- 2.5
21	2	0.753+- 0.006	0.795+- 0.010	0.613+- 0.018	0.821+- 0.021	0.456+- 0.028	0.865+- 0.015	0.495+- 0.020	0.924+- 0.009	ASCAG	252.5+- 3.6	158.5+- 2.5	286.0+- 3.5	122.5+- 3.6	77.5+- 2.5	149.0+- 3.5
22	1	0.731+- 0.006	0.693+- 0.010	1.000+- 0.000	0.000+- 0.000	0.000+- 0.000	1.000+- 0.000	1.000+- 0.000	0.000+- 0.000	pANCA	242.4+- 3.0	158.6+- 3.3	296.0+- 4.3	132.6+- 3.0	77.4+- 3.3	139.0+- 4.3

【0 3 5 9】

C. IBS 対健常個体を診断するためのランダムフォレストモデリング

[0361] 健常な対象と IBS を有すると予測される患者とを鑑別するために、別のランダムフォレストモデルを作出した。上述の 7 種の血清学的マーカーをこのモデルにおいて使用した（表 1 3 参照）。

【0 3 6 0】

表 1 3：IBS 対健常対照のためのランダムフォレストモデル

10

20

30

40

【表 1 3】

	NVAR	OOB_ ROC	TEST_ ROC	SENS	SPEC	LEAST_ IMP	NTRAIN 0	NTRAIN 1	NTEST 0	NTEST 1
1	6	0.776+- 0.005	0.775+- 0.010	0.639+- 0.020	0.794+- 0.007	pANCA	130.9+- 1.1	119.1+- 1.1	66.1+- 1.1	58.9+- 1.1
2	5	0.785+- 0.006	0.760+- 0.010	0.630+- 0.026	0.782+- 0.023	CBir1	133.6+- 1.6	116.4+- 1.6	63.4+- 1.6	61.6+- 1.6
3	4	0.782+- 0.007	0.791+- 0.010	0.704+- 0.016	0.777+- 0.011	ANCA	130.4+- 1.0	119.6+- 1.0	66.6+- 1.0	58.4+- 1.0
4	3	0.758+- 0.006	0.766+- 0.010	0.643+- 0.017	0.820+- 0.018	ASCAA	131.1+- 1.0	118.9+- 1.0	65.9+- 1.0	59.1+- 1.0
5	2	0.736+- 0.009	0.711+- 0.009	0.595+- 0.016	0.761+- 0.012	ASCAG	130.9+- 1.2	119.1+- 1.2	66.1+- 1.2	58.9+- 1.2
6	1	0.597+- 0.009	0.624+- 0.011	0.488+- 0.022	0.745+- 0.014	OmpC	130.6+- 0.7	119.5+- 0.7	66.5+- 0.7	58.5+- 0.7

10

【0 3 6 1】

[0362]表 1 4 は、I B S 対健常対照に関するランダムフォレストモデルを改変することによるさらに別のランダムフォレストモデルを示す。この実施形態では、最も重要なマーカーから開始して、あるマーカーをモデルに順に組み込んだ。クラスのうちの 1 つについて最早完全なデータが存在しない場合、モデルは完全なものであると決定した。表 1 4 は、1 7 種の血清－遺伝子－炎症マーカーを用いて I B S 対健常対象を予測するためのランダムフォレストモデルの予測を示す。

20

【0 3 6 2】

表 1 4： 1 7 種の血清－遺伝子－炎症マーカーを使用する I B S 対健常対照のためのランダムフォレストモデル

【表 1 4】

	NVAR	OOB_ ROC	TEST_ ROC	SENS	SPEC	LEAST_ IMP	NTRAIN 0	NTRAIN 1	NTEST 0	NTEST 1
1	17	0.883+- 0.006	0.877+- 0.010	0.775+- 0.024	0.882+- 0.020	CRP	43.5+- 0.8	86.5+- 0.8	22.5+- 0.8	43.5+- 0.8
2	16	0.874+- 0.008	0.900+- 0.010	0.827+- 0.019	0.878+- 0.011	STAT3	43.4+- 0.9	87.6+- 0.9	22.6+- 0.9	43.4+- 0.9
3	15	0.884+- 0.005	0.900+- 0.011	0.818+- 0.014	0.879+- 0.014	GLI1	45.0+- 0.7	86.0+- 0.7	21.0+- 0.7	45.0+- 0.7
4	14	0.896+- 0.004	0.901+- 0.010	0.805+- 0.022	0.883+- 0.010	MAGI2	45.6+- 0.8	87.4+- 0.8	23.4+- 0.8	43.6+- 0.8
5	13	0.885+- 0.005	0.901+- 0.010	0.804+- 0.020	0.886+- 0.015	NKX2	46.9+- 0.7	87.1+- 0.7	23.1+- 0.7	43.9+- 0.7
6	12	0.889+- 0.005	0.922+- 0.010	0.845+- 0.015	0.889+- 0.012	ATG16L1	45.7+- 1.2	88.3+- 1.2	24.3+- 1.2	42.7+- 1.2
7	11	0.888+- 0.008	0.901+- 0.010	0.836+- 0.021	0.867+- 0.012	pANCA	55.3+- 1.2	86.7+- 1.2	25.7+- 1.2	45.3+- 1.2
8	10	0.887+- 0.004	0.897+- 0.010	0.807+- 0.020	0.878+- 0.019	ASCAA	52.0+- 0.8	90.0+- 0.8	29.0+- 0.8	42.0+- 0.8
9	9	0.888+- 0.009	0.900+- 0.010	0.780+- 0.020	0.910+- 0.014	CBir1	54.0+- 1.5	88.0+- 1.5	27.0+- 1.5	44.0+- 1.5
10	8	0.879+- 0.009	0.913+- 0.010	0.802+- 0.019	0.905+- 0.008	SAA	54.2+- 1.5	87.8+- 1.5	26.8+- 1.5	44.2+- 1.5
11	7	0.880+- 0.006	0.888+- 0.009	0.802+- 0.022	0.860+- 0.012	ANCA	54.6+- 1.0	87.4+- 1.0	26.4+- 1.0	44.6+- 1.0
12	6	0.879+- 0.005	0.878+- 0.010	0.768+- 0.021	0.864+- 0.010	Fla2	53.4+- 0.7	88.6+- 0.7	27.6+- 0.7	43.4+- 0.7
13	5	0.869+- 0.005	0.900+- 0.009	0.826+- 0.035	0.866+- 0.012	ASCAG	55.4+- 1.9	88.6+- 1.9	28.6+- 1.9	43.4+- 1.9
14	4	0.867+- 0.003	0.853+- 0.006	0.772+- 0.033	0.822+- 0.024	Flax	56.8+- 2.0	87.2+- 2.0	27.2+- 2.0	44.8+- 2.0
15	3	0.839+- 0.008	0.845+- 0.010	0.730+- 0.015	0.830+- 0.016	OmpA	55.4+- 0.8	88.6+- 0.8	28.6+- 0.8	43.4+- 0.8
16	2	0.767+- 0.008	0.784+- 0.010	0.577+- 0.029	0.861+- 0.019	OmpC	57.4+- 1.2	114.6+- 1.2	27.6+- 1.2	59.4+- 1.2
17	1	0.688+- 0.010	0.718+- 0.015	0.619+- 0.025	0.732+- 0.011	ICAM	56.8+- 0.8	115.2+- 0.8	28.2+- 0.8	58.8+- 0.8

10

20

30

【0 3 6 3】

【0363】用いた 1 7 種のマーカーのうち、アルゴリズムは、I B S 対健常患者を予測するための最も重要なマーカーが I C A M、抗 O m p C、抗 O m p A、及び抗 F l a Xであることを示す。いくつかの態様では、本発明によるランダムフォレストモデリングは、I B S を有する患者対健常対照患者を予測することにより I B S を診断するのに有用であり得る。

【0 3 6 4】

実施例 3 (I B D 血清－遺伝子－炎症 (s g i) 診断アルゴリズム) :

40

【0364】本実施例は、I B D 血清－遺伝子－炎症 (s g i) 診断アルゴリズムの開発及び検証を例示する。いくつかの態様では、診断アルゴリズムは、炎症性腸疾患 (I B D) 対非炎症性腸疾患 (非 I B D) 及び／又はそのサブタイプ (例えば、潰瘍性大腸炎 (U C) 、クローン病 (C D) 、分類不能大腸炎 (I C) 、並びに C D 及び U C に関して不確定の I B D) を診断するのに用いることができる I B D s g i 診断試験の成分である。本実施例はまた、本明細書に記載のアルゴリズムを用いて I B D 、非 I B D 、U C 、C D 、I C 、及び／又は不確定 I B D を予測するための方法も例示する。本実施例は、診断アルゴリズムを作出するための方法をさらに説明する。

【0 3 6 5】

【0365】I B D s g i 診断アルゴリズムは 1 7 種のマーカーの測定値を利用する (p A

50

N C A は 2 回測定されるため値は 1 8 個)。表 1 5 は、マーカー及びマーカーを測定するためのアッセイ方法の例を記載する。

【 0 3 6 6 】

表 1 5 : I B D 診断アルゴリズムの血清－遺伝子－炎症 (s g i) マーカー
【表 1 5】

	マーカー	アッセイフォーマット
1	ASCA-IgA	ELISA
2	ASCA-IgG	ELISA
3	抗 OmpC	ELISA
4	抗 CBir1	ELISA
5	pANCA	間接免疫蛍光
6	pANCA2	間接免疫蛍光
7	ANCA	ELISA
8	抗 Fla2	ELISA
9	抗 FlaX	ELISA
10	CRP	ELISA
11	SAA	ELISA
12	ICAM	ELISA
13	VCAM	ELISA
14	VEGF	ELISA
15	ATG16L1	ジェノタイピング PCR
16	ECM1	ジェノタイピング PCR
17	NKX2-3	ジェノタイピング PCR
18	STAT3	ジェノタイピング PCR

10

20

【 0 3 6 7 】

[0366]いくつかの実施形態では、I B D s g i 診断試験はデータ分析計算アルゴリズムを含む。いくつかの実施形態では、最終的な試験結果は、患者の I B D 状態の予測を反映する確率スコアである。

30

【 0 3 6 8 】

[0367]本発明の I B D s g i 診断試験は、患者が I B D 及び／又はそのサブタイプ（例えば、U C、C D、I C、及び／又は C D 及び U C に関して不確定である I B D）を有するかどうかを診断するため又は医師が診断するのを助けるために、有利なことに血清、遺伝子及び炎症データを使用する。いくつかの実施形態では、診断試験は、患者のリスクプロファイルを評価するために、1 3 種の血清学的及び炎症バイオマーカー並びに 4 種の遺伝子マーカーを用いる。他のいくつかの実施形態では、試験は生物学的マーカーの測定値を取り、患者の I B D 状態を予測するための計算アルゴリズムを用いる。特に、試験は、患者が I B D 又は非 I B D を有するかどうかを予測することができる。試験はまた、I B D を有すると予測され、分類不能大腸炎 (I C) 及び／又は C D 及び U C に関して不確定の I B D を有するものから除外された患者に由来するマーカー測定値を検査し、患者が U C 又は C D を有するかどうかを予測することもできる。診断試験は、医師が、（任意選択でさらなる臨床知見を併用して、）最大限の情報に基づいて患者の管理のための決断を行うのを助ける包括的結果を提供する。

40

【 0 3 6 9 】

A. 血清－遺伝子－炎症 (s g i) マーカーの測定

[0368] I B D s g i 診断試験は、複数の血清－遺伝子－炎症 (s g i) マーカーの測定値を取り、患者の I B D 状態を予測するための計算アルゴリズムを用いる。血清－遺伝子－炎症 (s g i) マーカーを測定する方法は、例えば、米国特許第 5 7 5 0 4 7 9 号、

50

第5830675号、第7873479号、米国特許出願公開第2011/0045476号、及び国際出願PCT/US2011/039174（これらは、参照によりすべての目的のためにそのまま本明細書に組み入れられる。）に記載されている。

【0370】

B. IBD sgi診断アルゴリズム

[0369] IBD sgi診断アルゴリズムは、患者のIBD状態を予測するための計算アルゴリズム（血清－遺伝子－炎症（sgi）アルゴリズムと呼ばれる。）を利用する。このアルゴリズムの一実施形態のステップを図1に例示する。第1に、アルゴリズムは、17種の生物学的マーカー（血清、遺伝子、及び炎症マーカー）の測定データに基づくランダムフォレストモデルを使用し、患者がIBDを有するか、又はIBDを有さない（非IBD）かを予測する。次に、決定木又はルールのセットを用いて、IBDを有すると予測された患者が不確定診断（例えば、CD対UCに関して不確定との診断）に一致するバイオマーカーパターンを有するかどうかを決定する。次いで、11種のバイオマーカー測定値を用いる別のランダムフォレストに基づくアルゴリズムは、IBDを有すると予測された患者がUC又はCDのいずれかを有するかどうかを予測する。すなわち、本発明のIBD sgi診断アルゴリズムは、非IBD、UC、CD、あるいはCD及びUCに関して不確定であるIBDを有するものとして患者を分類する。

【0371】

C. モデリング戦略及び訓練セット

[0370] sgiアルゴリズムの成分は、各木を構築する無作為成分を含む多くの決定木から作られるランダムフォレストモデルとして知られるある型の計算分類子を用いる。sgiアルゴリズムについては、10,000本の木を有するフォレストを、IBD対非IBD及びUC対CD分類の各々について構築した。フォレスト中の各決定木は同じ分類問題に対処するが、例の異なる集合及び例のデータセットから無作為に選択される特徴の異なるサブセットが提供される。例えば、1本の木は、利用可能な例の無作為の2/3を、木のそれぞれのノードで分割される決定を行うために選択される特徴の無作為の2/3と共に用いて構築することができる。フォレストが構築されたら、すべての決定木にわたって新しく採決することによって新しい例を分類する。2クラスの分類子に関する最も単純な事例では、最も投票数が多いクラスが勝つ。一部の事例では、投票の勝利数のカットオフを予め設定して、性能評価尺度を最適化する。例えば、偽陽性が偽陰性よりもコストが高い例では、カットオフをより高い値に設定する。

【0372】

[0371] sgiアルゴリズムのランダムフォレスト分類子を、既知の非IBD、CD及びUC患者に由来する生物学的データに基づいてサンプルの集合（既知の診断を有する患者のコホートに由来するサンプルのデータセット）である訓練データから構築した。サンプルを、各例に関する真の分類と共に、分類に関連することが疑われる特徴（例えば生物学的マーカー）に関するデータを含む構造化された形式で提供した。この計算分類子アプローチは次の理由から「管理機械学習」と呼ばれる。（1）例の真のクラスが事前に知られている（管理）；（2）コンピュータが処理を行う（機械）；及び（3）アルゴリズムが異なるクラスを識別する特徴パターンを発見し、認識する（学習）。特徴パターンが学習されたら、分類子は新しい例の分類を予測することができる。

【0373】

[0372] IBD対非IBD診断を予測するための2クラスランダムフォレスト分類子を、既知の非IBD、CD及びUC患者に由来する血清－遺伝子－炎症生物学的マーカーのデータの集合から構築した。26種のマーカーに関する測定値を含む2420個のサンプルが元々考慮された。血清－遺伝子－炎症生物学的マーカーとしては、これらに限定されるものではないが、抗サッカロマイセス・セレピシエ抗体（例えば、ASCA-IgA、ASCA-IgG）、抗好中球抗体（例えば、ANCA、pANCA、例えば、IFA核周囲パターン及びDNase感受性）、抗微生物及び抗フラジェリン抗体（例えば、OmpA、OmpC、CBir1、A4-Fla2、FlaXに対する抗体）、増殖因子（例え

ば V E G F)、急性期タンパク質(例えば C R P)、アポリポタンパク質(例えば S A A)、細胞接着分子(例えば、I C A M-1、V C A M-1)、及び S N P(例えば、A T G 1 6 L 1 1 1 5 5 (A>G)、G L I 1 2 8 7 6 (G>A)、I L 1 0 (A>G)、I R G M (A>G)、I R G M (C>T)、M A G I 2 (C>G)、T L 1 A (C>T) X B P 1 (A>G)、M D R 1 2 6 7 7 (G>T/A)、E C M 1 5 8 8 (C>T)、N K X 2-3 (A>G)、S T A T 3 (A>G))が挙げられる。

【0374】

[0373]モデリング戦略は、アルゴリズムを作出、試験及び検証するための既知の診断のサンプル(患者)から入手可能なデータを用いる。最初に、入手可能なデータを、訓練データセットと検証データセットとに分割した。得られるアルゴリズムが新しいサンプルを予測する能力について公正に説明するために、基本的なモデリング、変数選択、及びモデル構造に関するすべての実験は、訓練データセットの使用にのみ限定された。異なるアプローチの有効性の評価の間に、訓練データセットを、アルゴリズムを作出するために使用される2/3のモデル訓練、及び性能を測定するために使用される1/3の試験データセットに無作為にさらに分割した。訓練及び試験データセットを無作為に再分割し(「リサンプリング」と呼ばれるプロセス)、平均値を比較し、確実に平均の標準誤差が性能差よりも小さくなるようにすることによって、アルゴリズムのモデリングの間のすべての実験を約50~100回繰り返した。

【0375】

[0374]入手可能なサンプルは、26種の生物学的マーカーのデータを含んでいた。マーカーのサブセットがより高い予測値を有するかどうかを決定するために、変数消去プロトコールを行った。簡単に述べると、このプロトコールは、リサンプリングにより各マーカーの平均及び標準誤差重要性を決定すること、最小重要マーカーを削除すること、並びにすべてのマーカーがその標準誤差よりも高い正の重要性を有するようになるまで繰り返すことを含む。変数重要性、アッセイ安定性、及び利用可能性を考慮することにより、I B D対非I B Dモデルを構築するのに用いられるマーカーの最終的なセット(例えば17種のマーカー)を選択した。これらのマーカーのすべてがU C対C Dモデルについて有意であったわけではなく、そのため、マーカーのあるサブセットのみ(例えば11種のマーカー)をこのモデルのために用いた。

【0376】

[0375]すべての訓練基準及びマーカーを決めた後、I B D対非I B Dモデルを、すべての訓練データを用いて作出した。次いで、U C対C Dモデルを、このデータの適切なサブセット上で作出した。

【0377】

[0376]訓練データセット中の既知のC D及びU Cサンプル(例えば、C D又はU Cを有すると以前に診断された患者に由来するサンプル)を次のように階層化した。診断(C DかU Cか)、民族/人種(ユダヤ系かそれ以外か)、サンプル供給源(地域社会か研究機関か)、診断からの時間(5年未満か5年以上か)。肝炎、G E R D又は他のすべての診断、民族/人種、及びサンプル供給源により、疾患対照サンプルを階層化した。さらに、他に2つの制約を適用した。いくつかの例では、訓練又は検証セットのいずれにおいても、アフリカ系アメリカ人又は小児サンプルは使用しなかった。そして、特定の cohorts に由来する174個のC Dサンプルは別のプロトコールの一部として使用されたため、検証セットにおいて使用しなかった。

【0378】

D. 個別の数値マーカー尺度へのテキストの転換

[0377]s g i アルゴリズムには数値が必要であるため、0又は1の値をルールセットと相関させることにより、任意の生物学的マーカーに関するテキストベースのデータをバイナリー変数に転換する(表16参照)。

【0379】

表16: 6種の血清-遺伝子-炎症生物学的マーカーに関するテキストベースのデ

10

20

30

40

50

ータの離散数値データ（例えばバイナリー変数）への転換のためのルール

【表 1 6】

バイオマーカー	0 の値に割り当てられるデータ	1 の値に割り当てられるデータ
pANCA	検出されず DNase 感受性 (細胞質)	DNase 感受性 1+P DNase 感受性 2+P DNase 感受性 3+P DNase 感受性 4+P
pANCA2	検出されず DNase 感受性 (細胞質) DNase 感受性 1+P	DNase 感受性 2+P DNase 感受性 3+P DNase 感受性 4+P
ATG1L1	VIC 両方	FAM
ECM1	VIC 両方	FAM
NKX2-3	VIC 両方	FAM
STAT3	両方 FAM	VIC

10

20

【0380】

[0378] pANCA 及び pANCA2 データは、患者サンプル中で検出された ANCA の蛍光標識の量及び位置を指す。用語「検出されず」は、標識が検出されず、それ故、サンプルが陰性であると考えられ、0 の値を割り当てられたことを示す。「DNase 耐性」と記載されたサンプルは、標識が DNase 処理後に消失することができず、「検出されず」と考えられたことを示す。この標識パターンは本実施例において使用された sg アルゴリズムを確立するために使用されたデータサンプルセットにおいては観察されなかったことに留意されたい。しかしながら、「DNase 耐性」サンプルは通常、サンプルセットの 2 % 中に生じる。用語「DNase 感受性 (細胞質)」は、DNase 処理後に、標識が細胞の細胞質に分布し、サンプルが陰性であると考えられたことを示す。それはまた、0 の値を割り当てられた。用語「DNase 感受性 1 + P」、「DNase 感受性 2 + P」、「DNase 感受性 3 + P」、「DNase 感受性 4 + P」は、標識が、1 + P から 4 + P までの見られた標識の強度及び量の増大と共に細胞核の周囲であるという視覚的評価（核周囲パターン）により観察されることを示したが、ここで 4 + P は最大量の標識と相関する。pANCA については、これらの標識は、検出され、陽性であると考えられ、1 の値を割り当てられる。pANCA2 については、2 + P から 4 + P までの標識のみが陽性であると考えられ、1 の値を割り当てられる。注目すべきことに、pANCA 及び pANCA2 は単一のマーカーの 2 つの独立に決定される測定値である。図 4 は、核周囲パターン、DNase 感受性細胞質パターン、及び DNase 耐性パターンなどの異なる pANCA 染色パターンの例を示す。

30

40

【0381】

[0379] ATG16L1、ECM1、NKX2-3 及び STAT3 などの遺伝子マーカーについて、VIC 及び FAM とは、各遺伝子の異なる対立遺伝子を標識するのに用いられた蛍光リポーター染料を指す。VIC 又は FAM 染料のいずれかを用いて各遺伝子の疾患対立遺伝子を標識する。テキストデータ「VIC」、「FAM」又は「両方」は、正常対立遺伝子、疾患対立遺伝子、又は両方の対立遺伝子がサンプル中に見られるかどうかを示す。一部の実施形態では、正常対立遺伝子のための染料のみ（遺伝子に応じて「VIC」又は「FAM」）が検出される場合、患者は正常対立遺伝子に関してホモ接合性であり、マーカーデータは 0 の値を割り当てられる。いくつかの例では、疾患対立遺伝子のための染料のみ（遺伝子に応じて「VIC」又は「FAM」）が検出される場合、これは患者が

50

疾患対立遺伝子に関してホモ接合性であることを示し、1の値を割り当てられる。他の例では両方の染料が検出され、「両方」と記載される場合、これは、患者が疾患対立遺伝子に関してヘテロ接合性であることを示し、0の値を割り当てられる。

【0382】

E. サンプル除去

[0380] s g i アルゴリズムを構築するために、サンプルに関する利用可能なデータを再検討し、サンプルが規定の基準に合致する場合、それらをデータセットから除去した。規定の基準の例としては、サンプルが最後に経験した症状が医師への訪問の6か月より前であった患者に由来する場合、又はサンプルが、最後の症状が「利用できない」と記述され、以前に試験に登録した患者に由来する場合が挙げられる。除外される患者の他の例は、試験対象患者基準及び試験除外基準を満たすことができなかった患者並びに疾患の長さに関する情報がない患者である。p A N C A 2 陽性及び A N C A < 1 1 . 9 E U / m l を有する患者サンプルは、データから除去された。同様に、p A N C A 2 陰性及び A N C A > 1 1 . 9 E U / m l を有する患者サンプルも除外された。これにより、訓練セットについては1, 083個のサンプル及び検証セットについては437個のサンプルが残された。

【0383】

F. I B D 対非 I B D 診断を予測するための第1のランダムフォレストモデリング

[0381] 残りのサンプルの全部を使用し、RからのR a n d o m F o r e s t パッケージなどの管理機械学習アプローチを使用して、I B D 対非 I B D 診断を予測する2クラスランダムフォレスト分類子を構築した。抗サッカロミセス・セレビジア抗体（例えば、A S C A - I g A、A S C A - I g G）、抗好中球抗体（例えば、p A N C A 及び A N C A）、抗微生物及び抗フラジェリン抗体（例えば、抗 O m p C、抗 C B i r 1、抗 F l a 2 及び抗 F l a X）、増殖因子（例えば V E G F）、急性期タンパク質（C R P）、アポリポタンパク質（例えば S A A）、細胞接着分子（例えば、I C A M 及び V C A M）、及び S N P（例えば、A T G 1 6 L 1、E C M 1、N K X 2 - 3、S T A T 3）などの17種の血清-遺伝子-炎症（s g i）マーカーを使用した。感度と特異度の間の所望のトレードオフを達成するために、スコアカットオフを0.64に設定した。

【0384】

[0382] s g i アルゴリズムは、診断試験における新しいサンプルの最初のモデルスコアを計算するために、17種の血清-遺伝子-炎症生物学的マーカーからのスコアを用いる。スコアが0.64未満である場合、サンプルはI B D を有する患者に由来すると予測される。そうでなければ、サンプルは非 I B D を有する患者に由来すると予測される。I B D を有すると予測されたサンプルはアルゴリズムの次のステップに進み、決定木又はルールのセットを用いて、不確定サンプルを予測する。

【0385】

G. サンプルが不確定であるかどうかを決定するための決定木又はルールのセット

[0383] いくつかの実施形態では、決定木又はルールのセットを用いて、I B D を有すると第1のランダムフォレストモデルにおいて分類されたサンプルが、不確定として分類することができるような不明瞭なマーカーパターンも有するかどうかを決定する。A N C A が 1 6 . 8 E U / m l 以上であり、p A N C A 2 スコアが陽性であり、以下の条件：C B i r 1 が 2 9 . 4 7 5 E U / m l 以上である、F l a 2 が 3 4 . 5 E U / m l 以上であるか、又は F l a X が 2 8 . 8 7 5 E U / m l 以上である、のうちの1つ又は複数が当てはまる場合、サンプルは「分類不能」と決定され、患者はC D 及び U C に関して「不確定」である。同様に、p A N C A 2 スコアが陽性であり、これらの条件：C B i r 1 が 2 9 . 4 7 5 E U / m l 以上である、F l a 2 が 3 4 . 5 E U / m l 以上であるか、又は F l a X が 2 8 . 8 7 5 E U / m l 以上である、のいずれか2つ以上が当てはまる場合、このサンプルも「分類不能」であり、患者はC D 及び U C に関して「不確定」である。本明細書に記載の分類及びルールにおける参照値は訓練セットの第三四分位数に由来する。

【0386】

H. U C 対 C D 診断を予測するための第2のランダムフォレストモデリング

[0384] UC 対 CD を予測するための第 2 のランダムフォレストモデルを元の訓練データセットのサブセットから作出した。既知の非 IBD サンプル、分類不能サンプル、及びルール 1（グループ 1 とも呼ばれる。）又はルール 2（グループ 2 とも呼ばれる。）違反基準を満たすサンプルであると決定されたサンプルを訓練データセットから除外した。ルール 1 及びルール 2 違反の詳細な説明は表 17 に記載されている。

【0387】

表 17：UC 対 CD を予測するための第 2 のランダムフォレスト分類子の構築において使用されるサンプルを除外するのに用いられる違反基準

【表 17】

ルール 1 違反	ルール 2 違反
以前の診断が UC である	以前の診断が CD である
ANCA <16.8	ANCA $\geq$ 16.8 EU/ml
pANCA2 陰性	下記条件のすべてが満たされる: ASCAA <10.675 EU/ml ASCAG <14.175 EU/ml OmpC <9.4 EU/ml CBir1 <29.475 EU/ml Fla2 <34.5 EU/ml FlaX <28.875 EU/ml
下記条件の 2 つ以上が満たされる: ASCAA $\geq$ 10.675 EU/ml ASCAG $\geq$ 14.175 EU/ml OmpC $\geq$ 9.4 EU/ml CBir1 $\geq$ 29.475 EU/ml Fla2 $\geq$ 34.5 EU/ml FlaX $\geq$ 28.875 EU/ml	

10

20

【0388】

[0385] これらのサンプル除外基準のため、540 個のサンプルを用いて、UC 対 CD 診断を予測するための 2 クラスランダムフォレスト分類子を決定した。IBD 対非 IBD ランダムフォレストモデルにおいて用いられたマーカー選択の同じ方法を用いた。選択された血清－遺伝子－炎症マーカーは、ASCA-IgA、ASCA-IgG、pANCA、ANCA、OmpC、CBir1、Fla2 及び FlaX、VEGF、ECM1 及び STAT3 であった。第 2 のランダムフォレスト分類子のスコアカットオフを 0.35 に設定した。

30

【0389】

[0386] いくつかの実施形態では、sgi 診断アルゴリズムは、新しいサンプルの第 2 のモデルスコアを計算するために 11 種の血清－遺伝子－炎症生物学的マーカーに由来するスコアを適用する。スコアが 0.35 未満である場合、サンプルは CD を有する患者に由来するものであると予測される。そうでなければ、サンプルは UC を有する患者に由来するものであると予測される。

40

【0390】

I. 新しいサンプルにおける診断を予測するためのアルゴリズムの例

[0387] いくつかの実施形態では、本発明の IBD sgi 診断アルゴリズムは、17 種の血清－遺伝子－炎症（sgi）生物学的マーカー（例えば、ANCA、ASCA-A、ASCA-G、抗 FlaX、抗 A4-Fla2、pANCA、抗 OmpC、抗 CBir1、ATG16L1、CRP、SAA、ICAM、VCAM、ECM1、STAT3、VEGF、NKX2-3、及びそれらの組み合わせ）に由来する測定値を使用し、IBD 対非 IBD を予測するための第 1 のランダムフォレストモデルに基づくスコアを計算する（図 1、110 参照）。第 1 のランダムフォレストモデルは、患者が IBD を有するかどうか

50

を決定する。スコアが I B D 対非 I B D カットオフ未満（例えば 0.64 未満）である場合、サンプルは I B D を有する患者に由来するものであると予測される（125）。そうでなければ、サンプルは非 I B D を有する患者に由来するものであると予測される（120）。I B D を有すると予測されるサンプルは、アルゴリズムの次のステップに進み、これはサンプルを不確定と分類するのを除外するように設計された決定木又はルールのセットである（130）。サンプルが「分類不能」ルールのいずれかのパターンに一致する場合、アルゴリズムは、サンプルが I B D を有するが、U C 及び C D に関しては不確定と予測する（135）。そうでなければ、サンプルはアルゴリズムの次のステップに進み（140）、これは、U C 対 C D を予測するための第 2 のランダムフォレストモデルである（150）。本発明の診断アルゴリズムは、11 種の血清－遺伝子－炎症生物学的マーカー（例えば、A N C A、A S C A－A、A S C A－G、抗 F l a X、抗 A 4－F l a 2、p A N C A、抗 O m p C、抗 C B i r 1、E C M 1、S T A T 3、V E G F、及びそれらの組み合わせ）に由来する測定値を用いて、U C 対 C D を予測するための第 2 のランダムフォレストモデルに基づくモデルスコアを計算する。スコアが U C 対 C D カットオフ（例えば 0.35）未満である場合、アルゴリズムは、サンプルが C D を有すると予測する（153）。そうでなければ、アルゴリズムは、サンプルが U C を有すると予測する（155）。

10

【0391】

J. 性能評価

[0388] s g i アルゴリズムの性能を、訓練の間に用いられたのと同じ除外基準を満たす別の検証セットに対して測定した。第 1 のランダムフォレストモデルの訓練に用いられたのと同じセットの除外を適用した後、検証セットは 437 個のサンプルからなっていた。第 2 のランダムフォレストモデルの訓練の間に適用された除外の同じルール（例えば、分類不能患者、及びまた「ルール 1 / ルール 2」違反）を用いた後、U C 対 C D コールの性能を決定した。これにより、既知の U C 又は既知の C D として 199 個のサンプルが残された。その結果を表 18 に示す。データは、s g i アルゴリズムが高い感度及び特異度で I B D、C D 及び U C を予測することを示している。図 5 は、訓練及び検証コホート特性を提供し、R O C 曲線下面積が個々のマーカー単独よりも s g i アルゴリズムの方が大きいことを示す。すなわち、本発明の s g i アルゴリズムは、有利なことに、（i）I B D、非 I B D、C D 及び U C を鑑別するのに有用である、（i i）A S C A / A N C A カットオフ試験よりも良好な性能を有する、及び（i i i）I B D マーカーは低レベルであるが、依然として疾患陽性である患者を同定する。

20

30

【0392】

表 18：別の検証セットに関する s g i アルゴリズムの性能結果

【表 18】

性能	sgi
IBD／非 IBD ROC	0.871
UC／CD ROC	0.929
IBD 感度	0.736
IBD 特異度	0.896
CD 感度	0.889
CD 特異度	0.810
UC 感度	0.977
UC 特異度	0.835

40

【0393】

K. マーカー重要性

[0389] ランダムフォレストモデリングは、変数（例えばマーカー）の重要性を決定する

50

ための方法を提供する。これは、全サンプル間で1つのマーカーに関する測定値にスクランブルをかけた後、フォレスト中の木の正確度の平均下落を決定することにより行われる。このように、マーカーと予測された結果との間の任意の有意な相関は破壊されるが、同時に、マーカーの統計的分布のすべての特徴を保存する可能性がある。特定の木の訓練において用いられないサンプル（“out-of-bag”）のみを正確度決定のために用いる。次いで、他のマーカーの各々について手順を繰り返す。報告されるスコアは、正確度の平均低下をその分散で除算したものである。2つのランダムフォレストモデルがsgialgorithm中にあるため、2セットの重要性が必要である。これらのものを表19及び表20に示す。各モデルについて保持されたマーカーはすべて、少なくともいくつかの重要性を有することに留意されたい。アッセイ安定性及びコストなどの実務的関心と共に、測定された重要性を使用して、マーカーがアッセイにおいて保持されるかを決定した。

10

【0394】

表19：IBD対非IBDモデルにおいて使用されるマーカーの重要度（訓練セットにおいて重要性の高い順）

【表19】

マーカー	重要性
ANCA	4.57
ASCAA	4.32
ASCAG	3.84
pANCA	3.75
FlaX	3.67
SAA	2.96
Fla2	2.88
ICAM	2.23
OmpC	2.02
CBir1	1.84
VCAM	1.84
CRP	1.35
NKX2	0.65
ATG16L1	0.41
STAT3	0.23
ECM1	0.20
VEGF	0.19

20

30

【0395】

表20：UC対CDモデルにおいて使用されるマーカーの重要度（訓練セットにおいて重要性の高い順）

【表 2 0】

マーカー	重要性
ANCA	6.37
ASCAA	5.62
ASCAG	4.75
FlaX	4.73
Fla2	4.45
pANCA	3.62
OmpC	3.59
CBir1	3.40
ECM1	0.92
STAT3	0.38
VEGF	0.17

10

【0396】

[0390]当業者であれば、重要性の値によって、IBD対非IBDモデルでは、ANCAが最も重要であり、VEGFがあまり重要でないことを知るであろう。同様に、UC対CDモデルでは、ANCAが最も高い重要性の値を有し、VEGFが最も低い重要性の値を有する。また、当業者であれば、IBD対非IBDの診断を決定又は予測するのに使用されるアルゴリズムが、表19に列挙された1つ又は複数のマーカーを含んでもよいことを知るであろう。同様に、当業者であれば、UC対CDの診断を決定又は予測するのに使用されるアルゴリズムが、表20に列挙された1つ又は複数のマーカーを含んでもよいことを知るであろう。

20

【0397】

実施例4（血清－遺伝子－炎症マーカーを用いるIBD診断のロジスティック回帰モデリング）：

[0391]本実施例は、ロジスティック回帰モデリングに基づくIBD診断のためのアルゴリズムを作出する方法を説明する。本実施例は、一方のモデルがIBD対非IBD診断を予測することができ、他方のモデルがCD対UC診断を予測することができる、2つのロジスティック回帰モデルを含むアルゴリズムを構築する方法をさらに説明する。

30

【0398】

[0392]データをロジット関数ロジスティック曲線に適合させることにより、ロジスティック回帰モデル（ロジスティックモデル又はロジットモデルとも呼ばれる）を事象の発生確率の予測のために用いる。一部の例では、それは二項回帰に用いられる一般化された線形モデルである。

【0399】

[0393]一部の実施形態では、ロジスティック回帰モデルは、実施例3に記載のランダムフォレストモデルを構築するのに用いられたものと同じ血清－遺伝子－炎症マーカー及びコホート組成を用いる。17種の血清－遺伝子－炎症マーカーとしては、限定されるものではないが、抗サッカロミセス・セレビジア抗体（例えば、ASCA-IgA及びASCA-IgG）、抗好中球抗体（例えば、pANCA及びANCA）、抗微生物及び抗フラジェリン抗体（例えば、OmpC、CBir1、Fla2及びFlaXに対する抗体）、増殖因子（例えばVEGF）、急性期タンパク質（CRP）、アポリポタンパク質（例えばSAA）、細胞接着分子（例えば、ICAM及びVCAM）、及びSNP（例えば、ATG16L1、ECM1、NKX2-3、STAT3）が挙げられる。

40

【0400】

[0394]実施例3に記載の類似のルールに従って、訓練及び検証セットからの除外のためにサンプルを選択した。ロジスティック回帰モデルを、訓練セット中の1083個のサンプルから構築した。次いで、既知の診断の437個のサンプルの検証セットを用いて予測

50

性能についてモデルを試験した。

【0401】

[0395] ロジットモデルを、実施例3に記載のランダムフォレストモデルを構築するのに用いられたものと同じ血清－遺伝子－炎症マーカー及びコホート組成を用いて作出した。アルゴリズムを開発する方法を図6に例示する。IBD対非IBD診断を予測するためのロジスティック回帰モデルを訓練するために(610)、抗サッカロミセス・セレビジア抗体(例えば、ASCA-IgA及びASCA-IgG)、抗好中球抗体(例えば、pANCA及びANCA)、抗微生物及び抗フラジェリン抗体(例えば、OmpC、Cbir1、Fla2及びFlaXに対する抗体)、増殖因子(例えばVEGF)、急性期タンパク質(CRP)、アポリポタンパク質(例えばSAA)、細胞接着分子(例えば、ICAM及びVCAM)、及びSNP(例えば、ATG16L1、ECM1、NKX2-3、STAT3)からなる17種の血清－遺伝子－炎症マーカーを使用した。このようにして、サンプルがIBD又は非IBDを有すると予測される患者に由来するかどうかを予測することができるコンピュータ計算統計アルゴリズムが作出される(620)。

10

【0402】

[0396] IBD分類不能サンプルであると決定されたサンプルを除去した後、IBDを有すると予測されるサンプルを使用して、第2のロジットモデルを作出する(630)。実施例3に記載の決定木分析と同様、分類不能サンプルは、次の2つの基準のうちの1つを満たすものと定義される。1) ANCAスコアが16.8 EU/ml以上であり、pANCA2スコアが陽性であり、次の条件：Cbir1が29.475 EU/ml以上、Fla2が34.5 EU/ml以上、若しくはFlaXが28.875 EU/ml以上、のうちの1つ若しくは複数が当てはまる；あるいは2) pANCA2が陽性であり、次の条件：Cbir1が29.475 EU/ml以上、Fla2が34.5 EU/ml以上、若しくはFlaXが28.875 EU/ml以上、のうちのいずれか2つ以上が当てはまる。サンプルがIBD対非IBDのためのロジスティック回帰モデルに従うIBDを有すると予測され、分類不能サンプルであるか、又は不確定と決定された場合、サンプルは分類不能大腸炎(IC)を有すると予測されるか、又は不確定である。11種の血清－遺伝子－炎症マーカー(例えば、ANCA、ASCA-A、ASCA-G、抗FlaX、抗Fla2、pANCA、抗OmpC、抗Cbir1、ECM1、STAT3、VEGF、及びそれらの組み合わせ)を用いて、第2のロジットモデルを訓練した(630)。訓練及び検証されたアルゴリズムは、患者がCD対UCを有するかどうかを予測することができる(640)。

20

30

【0403】

[0397] ロジットモデルアルゴリズムの性能を、既知の診断のサンプルの検証セット上でアルゴリズムを試験することにより分析することができる。実施例3に記載のランダムフォレストアルゴリズムの性能と、上述のロジスティック回帰アルゴリズムとの比較により、ランダムフォレストがより高いIBD特異度、CD感度及び特異度、及びUC特異度を示すことが示される(表21)。

【0404】

表21：アルゴリズム性能の比較

40

【表 2 1】

	ランダムフォレスト	ロジスティック回帰
IBD ROC AUC	0.871	0.862
UC ROC AUC	0.929	0.948
IBD 感度	0.736	0.744
IBD 特異度	0.896	0.858
CD 感度	0.889	0.878
CD 特異度	0.810	0.803
UC 感度	0.977	0.978
UC 特異度	0.835	0.800

10

【0 4 0 5】

[0398] 表 2 2 は、IBD、CD 及び UC を予測するためのアルゴリズムの性能を示す。感度は、式：真陽性 / (真陽性 + 偽陰性) を用いて算出される。特異度は、式：真陰性 / (真陰性 + 偽陽性) を用いて算出される。正確度は、式：(真陽性 + 真陰性) / (真陽性 + 真陰性 + 偽陽性 + 偽陰性) を用いて算出される。精度は、式：真陽性 / (真陽性 + 偽陽性) から算出される。アルゴリズムは、UC を予測するための高い感度及び正確度を有する。

【0 4 0 6】

20

表 2 2：ロジスティック回帰アルゴリズムの性能

【表 2 2】

	IBD	CD	UC
真陽性	189	79	45
真陰性	157	57	92
偽陽性	26	14	23
偽陰性	65	11	1
感度	0.744	0.878	0.978
特異度	0.858	0.803	0.800
正確度	0.792	0.845	0.851
精度	0.879	0.849	0.662

30

【0 4 0 7】

[0399] アルゴリズムの感度を分析し、一方が IBD ロジットモデルであり、他方が UC ロジットモデルである 2 つの受信者動作特性 (ROC) 曲線としてプロットすることができる。図 7 A は、IBD モデルは IBD を予測するための良好な分類システムであることを示す。図 7 B は、UC モデルは UC を予測するための良好な分類システムであることを示す。

【0 4 0 8】

[0400] IBD ロジットモデル又は IBD ランダムフォレストモデルのいずれかに基づくアルゴリズムの類似性を、予測された p 値の散布プロット (図 8) 又は平均予測 p 値の密度プロット (図 9) において比較することができる。図 9 は、ロジスティック回帰モデルと、いずれかのランダムフォレストモデルとの間にいくらかの相違が存在することを示す。

40

【0 4 0 9】

[0401] さらに、UC ロジットモデル又は UC 対 CD ランダムフォレストモデルのいずれかに基づくアルゴリズムの類似性を、予測された p 値の散布プロット (図 10) 又は平均予測 p 値の密度プロット (図 11) において比較することができる。図 11 は、ロジスティック回帰モデルと、いずれかのランダムフォレストモデルとの間にいくらかの相違が存在することを示す。

50

【0410】

実施例5（2つのランダムフォレストモデルに基づくs g i アルゴリズムを用いる改善されたIBD診断）：

[0402]本実施例は、患者のIBD状態（例えば、非IBD、UC、CD、IC、あるいはIBDであるが、UC及びCDに関して不確定）を予測するための血清－遺伝子－炎症（s g i）マーカーに基づくアルゴリズムを用いてIBD診断を改善するための方法を例示する。一部の実施形態では、s g i アルゴリズムは、コンピュータに基づく分類子を構築するために非IBD、IC、UC又はCDと分類された患者からのサンプルに由来する既知の血清－遺伝子－炎症マーカー測定値を用いる1つ又は複数のランダムフォレストモデルを含む。一部の実施形態では、s g i アルゴリズムは、IBD対非IBDを予測するランダムフォレストモデルを有してもよい。他の実施形態では、s g i アルゴリズムは、UC対CDを予測するランダムフォレストモデルを有してもよい。

10

【0411】

[0403]本発明のIBD s g i アルゴリズムの一実施形態のステップを、図1に例示する。第1に、s g i アルゴリズムは、17種の生物学的マーカー（pANCAは2回測定されるため、18個の値）の測定データに基づくランダムフォレスト分類子を使用し、患者が非IBDであるか、又はIBDを有するかどうかを予測する（110）。次に、決定木又はルールセットを使用して、IBDを有すると予測された患者が不確定診断（例えば、CD対UCに関して不確定との診断）に一致するバイオマーカーパターンを有するかどうかを決定する（130）。次いで、11種のマーカー測定値を用いる別のランダムフォレストに基づくアルゴリズムは、IBDを有すると予測される患者がUC又はCDのいずれかを有するかどうかを予測する（150）。このようにして、s g i アルゴリズムは、非IBD（120）、不確定（135）、CD（153）又はUC（155）患者からサンプルを試験分類する。

20

【0412】

[0404]s g i アルゴリズムは、非IBD、IC、UC及び／又はCD患者からのサンプルに由来する既知のマーカー測定値を含む訓練データセット（例えば訓練セット）を用いる一連の訓練、試験及びリサンプリングステップにより構築される。訓練セットは、2／3がモデル訓練セットを含み、1／3が試験セットを含むように分割される。

【0413】

[0405]いくつかの例では、訓練データセットのサンプルが、次の基準のうちのいずれかを満たす場合、それをさらなる分析から除去する。（1）対象の最後の症状の経験が医師への訪問の6か月より前である；（2）対象の最後の症状が「N/A」と記述され、対象が試験に登録した；（3）対象が試験対象患者基準／試験除外基準を満たさなかった；及び（4）対象の履歴が疾患罹患の長さに関する情報を欠く。

30

【0414】

[0406]一部の例では、UC対CDランダムフォレストモデルに関する訓練データセットからサンプルを除去する。一部の例では、次の基準を満たすサンプルを消去する。1）既知のCD診断；2）ANCA < 16.8 EU/ml；pANCA 2陰性；及び3）以下の条件： $ASCA - A \geq 10.675 \text{ EU/ml}$ 、 $ASCA - G \geq 14.175 \text{ EU/ml}$ 、 $OmpC \geq 9.4 \text{ EU/ml}$ 、 $CBir \geq 29.475 \text{ EU/ml}$ 、 $Fla2 \geq 34.5 \text{ EU/ml}$ 、又は $FlaX \geq 28.875 \text{ EU/ml}$ の2つ以上を満たす。他の例では、1）対象がUC診断を有する場合；2）ANCA $\geq 16.8 \text{ EU/ml}$ であり、pANCA 2陽性である場合；及び3）次の条件： $ASCA - A < 20.675 \text{ EU/ml}$ 、 $ASCA - G < 14.175 \text{ EU/ml}$ 、 $OmpC < 9.4 \text{ EU/ml}$ 、 $CBir < 29.475 \text{ EU/ml}$ 、 $Fla2 < 34.5 \text{ EU/ml}$ 及び $FlaX < 28.875 \text{ EU/ml}$ の全部を満たす場合、対象のサンプルを、UC対CDランダムフォレストモデルを構築するための訓練データセットから除去する。さらに別の例では、サンプルが既知の非IBDサンプルである、例えば、非IBDを有すると以前に診断されたサンプルである場合、そのサンプルを訓練セットから除外する。一部の例では、以下の条件のいずれかを満た

40

50

す、分類不能であると決定されたサンプルも訓練セットから除外する。1つの条件は、 $ANCA > 16.8 \text{ EU/ml}$ 、 $pANCA$ 陽性、及び $CBir1 \geq 29.475 \text{ EU/ml}$ 、 $Fla2 \geq 34.5 \text{ EU/ml}$ 又は $Flax \geq 28.875 \text{ EU/ml}$ である基準のうちの1つ又は複数を含む。他の条件は、 $pANCA$ 陽性、及び $CBir1 \geq 29.475 \text{ EU/ml}$ 、 $Fla2 \geq 34.5 \text{ EU/ml}$ 又は $Flax \geq 28.875 \text{ EU/ml}$ である基準のうちのいずれか2つ以上を含む。さらに別の例では、サンプルが次の基準のうちの1つを満たす場合、そのサンプルを除外する。(1) $pANCA = 1$ 及び $ANCA < 11.9 \text{ EU/ml}$; (2) $pANCA = 0$ 及び $ANCA \geq 11.9 \text{ EU/ml}$; (3) 明らかに誤差の影響を受ける; 又は (4) 生成モードにおいて再試験されるのが通常である。

10

【0415】

[0407]本実施例は、モデルに対する訓練セットを用いて IBD sg i アルゴリズムを構築するための方法を例示する。本実施例はまた、 IBD sg i アルゴリズムを検証する方法も例示する。一部の実施形態では、初期カットオフは、アルゴリズムのリサンプリングフェーズの平均カットオフと設定される。人工マーカーデータに関する確率及び予測を、最終的なモデルを平均リサンプリングモデルと比較することができるように作出することができる。一部の例では、人工マーカーデータは、類似する分布を有するマーカーの臨床的に意義のある空間に効率的に広がるデータセットである。いくつかの実施形態では、アルゴリズムは連続マーカーの統計値（例えば、八分位数又は8つの領域を含む変位値）を作出し、八分位数を無作為に拾うか、又はその八分位数内の均一に分布した無作為な値を拾うことにより、無作為マーカー値を生成する。バイナリー値マーカーの場合、 $pANCA = 1$ である場合、 $pANCA$ が1に等しいと設定されるか、又は他のすべてのマーカーは独立に生成される場合を除いて、八分位数を無作為に拾うことができる。

20

【0416】

[0408]一部の例では、分類不能大腸炎 (IC) を有する患者からサンプル进行分类する。特定の例では、 IBD を有し、 $ANCA$ レベル $\geq Q3$ であり、 $pANCA$ 陽性であり、抗 $CBir1$ 若しくは抗 $Fla2$ 若しくは抗 $Flax \geq Q3$ であると決定されるサンプルは、 IC を有する患者に由来するものであると予測される。他の例では、 IC 患者は、 $pANCA$ 陽性でもあり、 $Q3$ 以上である3つのマーカー（例えば、抗 $CBir1$ 、抗 $Fla2$ 、抗 $Flax$ ）のうちの2つを発現する、 IBD を有すると決定された患者であってもよい。表23は、“Serology 7”アッセイ (Prometheus Laboratories Inc., San Diego, CA) と比較した sg i アルゴリズムの性能を示す。注目すべきことに、表23は、 sg i アルゴリズムの性能が“Serology 7”アッセイよりも改善されることを示す。

30

【0417】

表23: sg i アルゴリズムの性能

【表 2 3】

性能	sgi	Serology7
IBD／非 ROC	0.871	0.831
UC／CD ROC	0.929	0.902
IBD 感度	0.736	0.717
IBD 特異度	0.896	0.825
CD 感度	0.889	0.871
CD 特異度	0.810	0.686
UC 感度	0.977	0.815
UC 特異度	0.835	0.778

10

【0 4 1 8】

[0409]表 2 4 は、アルゴリズムを訓練又は試験するのに用いられない 2 つのコホートに関する性能と比較した、検証セット（例えば試験セット）に関する S G I アルゴリズムの性能を示す。小児コホートは、サンプル除外ルールを行った後に 6 1 個のサンプルについて比較的小さいことに留意すべきである。検証セットに関する S G I アルゴリズムの性能は、訓練において用いられないコホートに由来するサンプルセットよりも良好である。

20

【0 4 1 9】

表 2 4：検証セットを含む s g i アルゴリズムの性能

【表 2 4】

性能	sgi 検証	アフリカ系 アメリカ人	幼児
IBD／非 ROC	0.871	0.797	0.925
UC／CD ROC	0.929	0.908	0.905
IBD 感度	0.736	0.691	0.487
IBD 特異度	0.896	0.791	1.0
CD 感度	0.889	0.893	1.0
CD 特異度	0.810	0.542	1.0
UC 感度	0.977	1.0	1.0
UC 特異度	0.835	0.857	1.0

30

40

【0 4 2 0】

[0410]データは、s g i アルゴリズムのランダムフォレストモデルと、ロジスティック回帰モデルとの比較を示す。図 1 2 は、y 軸に沿った I B D ロジスティック回帰モデルの p 値（実施例 4 参照）及び x 軸に沿った R において作出された I B D ランダムフォレストモデルの p 値の散布プロットを示す。図 1 3 は、y 軸に沿った U C 対 C D ロジスティック回帰モデルの p 値（実施例 4 参照）及び x 軸に沿った U C 対 C D ランダムフォレストモデルの p 値の散布プロットを示す。図 1 4 及び 1 5 は、2 つのアルゴリズム間の性能差を示している。

【0 4 2 1】

50

[0411] データは、2つの異なる統計解析プログラム（例えば、R及びMatlab）を用いたs g i アルゴリズムの比較を示す。図14は、IBD対非IBDモデルの比較を示す。図15は、UC対CDモデル間の比較を示す。データは、ソフトウェアプログラムが類似する結果を生成したことを示している。

【0422】

実施例6（IBD診断を予測するためのIBD s g i 診断アルゴリズム及びデータセット）：

[0412] コホートに由来するサンプルのデータの訓練セットを用いてコンピュータ計算モデルを訓練するプロセスにより、ランダムフォレストモデリングアルゴリズムを作出する。訓練されたアルゴリズムを、コホートに由来するサンプルのデータの検証セットを用いて検証する。本実施例は、ランダムフォレストモデリングにおける訓練セット及び検証セットの使用を説明する。

10

【0423】

[0413] 本実施例は、ランダムフォレスト機械学習モデリングの診断アルゴリズムを訓練及び検証するための生物学的マーカーの測定値を有するサンプルを用いる方法を例示する。本実施例は、本明細書に記載の方法（例えば、実施例1～5参照）を用いて、非IBD、IBD、IC、CD、UC、及び／又はUC及びCDに関して不確定であるIBDを予測又は診断することができることを示す。本実施例はまた、サンプルの除外のためのルールをアルゴリズムにおいて適用することができることも例示する。

【0424】

20

[0414] 一部の実施形態では、訓練及び検証セットは、非IBD、CD及びUCなどの既知の診断を有するサンプルに由来する血清－遺伝子－炎症（s g i）マーカー（例えば、ASCA-A、ASCA-G、pANCA、ANCA、抗OmpC、抗Fla2、抗Flax、VEGF、CRP、SAA、ICAM、VCAM、ATG16L1、ECM1、NKX2-3、及びSTAT3）の測定値を含む。訓練セットにおけるサンプルの既知の診断を表29の列「Dx3」中に列挙する。ここで、「0」は非IBD診断、「1」はCD診断、「2」はUC診断を指す。用語「既知の診断」は、ランダムフォレストモデリングを含まない方法を用いる患者の診断を指す。一部の例では、既知の診断が決定される。異なるコホートに由来するサンプル中での血清－遺伝子－炎症生物学的マーカーの検出を、本明細書に記載のように実施した。表25は、訓練セットのサンプル番号1～14の測定値を示す（データセット1）。14個のサンプルを有するデータセット1は、約1490個のサンプルのn値を有するはるかに大きいデータセットの単なる代表的なサブセットにすぎない。テキストベースの生物学的マーカーデータを個別の数値に転換した（例えば、実施例3及び表16参照）。表26は、pANCA、ATG16L1、ECM1、NKX2-3及びSTAT3などのマーカーの転換された測定値を示す。

30

【0425】

表25：訓練セットにおける生物学的マーカー

【表 2 5】

サンプル 番号	識別子	データ セット	ASCA A	ASCA G	ANCA	OmpC	CBir1	Fla2	Flax	VEGF	CRP	SAA	ICAM	VCAM
1	10-IBD-07.059-081	1	3	9.1	3	3	18.1	14.7	11.6	45	13706	3935	507	544
2	10-IBD-07.079-002	1	6.2	3	8.2	3	8.1	29.9	6.5	184	751	2088	240	487
3	10-IBD-01.009-024	1	23	27.1	28.7	7.9	26.3	34.1	35.7	171	5470	4341	513	971
4	10-IBD-01.009-026	1	78.3	46	23	11.4	100	100	100	543	29610	85380	441	677
5	10-IBD-07.040-004	1	3	3	5.3	3	7.3	5.3	3.3	116	1369	2727	511	452
6	10-IBD-07.040-006	1	3	3	4.5	6.6	11	17.2	15	30	15826	7843	409	511
7	10-IBD-01.005-031	1	100.1	49.8	84.4	6.2	83.5	95.5	70	71	3430	11039	488	545
8	10-IBD-01.013-028	1	60	100.1	12	7.4	82.5	100	100	106	149199	204889	950	1273
9	10-IBD-07.059-085	1	3	9.6	3	5.1	30.6	47.9	24.6	56	386	1684	340	554
10	10-IBD-07.036-021	1	3	16.2	3	3	31.6	18.3	16.3	30	4831	139982	536	736
11	10-IBD-01.002-008	1	3	3	19.7	3	8.3	15.7	11.7	76	2618	5638	329	458
12	10-IBD-01.011-007	1	4	3	28	8.5	17.4	22.2	16	218	1385	2918	690	878
13	10-IBD-01.001-026	1	42.8	8.8	85.5	53.7	10.9	7.8	8.8	298	18723	46041	672	1118
14	10-IBD-01.001-028	1	10.6	15.2	65.5	18	19.2	21.2	23.1	479	7782	3768	417	794

10

20

30

【0 4 2 6】

表 2 6 : 訓練セットにおけるテキストベースの生物学的マーカーのバイナリー変数

【表 2 6】

サンプル番号	対象識別子	データセット	pANCA	ATG16L1	ECM1	NKX2-3	STAT3
1	10-IBD-07.059-081	1	0	0	0	0	0
2	10-IBD-07.079-002	1	0	1	0	1	0
3	10-IBD-01.009-024	1	1	1	0	1	0
4	10-IBD-01.009-026	1	0	0	0	0	1
5	10-IBD-07.040-004	1	1	0	0	0	0
6	10-IBD-07.040-006	1	0	0	0	0	0
7	10-IBD-01.005-031	1	1	1	0	0	0
8	10-IBD-01.013-028	1	1	1	0	1	1
9	10-IBD-07.059-085	1	1	0	0	0	1
10	10-IBD-07.036-021	1	0	0	0	0	0
11	10-IBD-01.002-008	1	1	0	0	0	1
12	10-IBD-01.011-007	1	1	0	0	0	0
13	10-IBD-01.001-026	1	1	0	0	0	0
14	10-IBD-01.001-028	1	1	1	0	1	1

10

【0 4 2 7】

【0415】訓練されたランダムフォレストモデルを、表 2 7 及び表 2 8 に列挙されるサンプルの検証セットを用いて検証した。再度、ここで示される検証セット（サンプル 1 5 ～ 2 8）は、約 6 7 9 個のサンプルの n 値を有するはるかにより大きいデータセットの単なる代表的なサブセットにすぎない。

20

【0 4 2 8】

表 2 7：検証セットにおける生物学的マーカー

【表 2 7】

サンプル番号	対象識別子	データセット	ASCA-A	ASCA-G	ANCA	OmpC	CBir1	Fla2	Flax	VEGF	CRP	SAA	ICAM	VCAM
15	10-IBD-07.118-074	2	11.4	45.2	11.7	3	26.7	48.4	53.2	318	270	3852	286	554
16	10-IBD-07.118-098	2	3	3	100	3.2	4	35.8	29.3	77	432	2118	264	289
17	10-IBD-07.319-164	2	3	3.2	15.3	6.3	11.4	12.1	8.8	63	978	3343	325	474
18	10-IBD-07.319-169	2	7.7	12.9	27.2	3.9	7.6	7.1	5.7	57	1417	1753	280	439
19	10-IBD-07.319-156	2	3	18.3	9.8	3	67	31.1	36.2	66	2685	1162	281	354
20	10-IBD-07.319-161	2	3	3.7	11	3	14.7	6.7	6.4	103	5831	7248	343	568
21	10-IBD-07.324-004	2	8.8	12.4	45.7	3	100	17	11.3	346	894	2118	376	706
22	10-IBD-07.320-011	2	57.3	44.5	47.4	3	100	100	100	251	16935	48907	412	596
23	10-IBD-07.136-070	2	3	7.8	7.1	3.2	16.8	41	40.8	78	1879	6519	367	371
24	10-IBD-07.136-083	2	6	8.7	6	3	30	56.7	57.6	53	4301	5288	368	502
25	10-IBD-07.118-102	2	3	5.8	20.6	3	3.2	28.2	17.6	244	838	1557	318	526
26	10-IBD-07.118-104	2	3	4	35.4	3	12.2	23.2	18.2	49	1346	1903	272	393
27	10-IBD-07.321-003	2	4	12.6	3	6.4	21.6	11	10.2	125	9459	12652	630	608
28	10-IBD-07.323-011	2	100.1	44.1	8.3	41.9	66.9	100	100	162	14971	73366	831	927

30

40

【0 4 2 9】

50

表 28：検証セットにおけるテキストベースの生物学的マーカーのバイナリー変数
【表 28】

サンプル番号	対象識別子	データセット	pANCA	ATG16L1	ECM1	NKX2	STAT3
15	10-IBD-07.118-074	2	1	0	0	0	1
16	10-IBD-07.118-098	2	1	1	0	0	0
17	10-IBD-07.319-164	2	0	0	1	1	1
18	10-IBD-07.319-169	2	0	1	1	1	1
19	10-IBD-07.319-156	2	0	0	0	0	1
20	10-IBD-07.319-161	2	0	1	0	1	1
21	10-IBD-07.324-004	2	1	0	1	1	0
22	10-IBD-07.320-011	2	1	0	0	0	0
23	10-IBD-07.136-070	2	1	0	0	1	0
24	10-IBD-07.136-083	2	0	1	1	0	0
25	10-IBD-07.118-102	2	1	0	0	1	0
26	10-IBD-07.118-104	2	1	0	1	0	0
27	10-IBD-07.321-003	2	0	1	0	0	1
28	10-IBD-07.323-011	2	1	0	0	0	1

10

【0430】

20

[0416] 表 29 は、診断アルゴリズムが訓練セット及び検証セット中のサンプルの診断を予測することができることを示す。特に、列「RF Dx 4」中の値は予測された診断を次のように記述する。「0」は非 IBD、「1」は CD、「2」は UC、「3」は IC を指す。表 29 はまた、サンプル除外基準、分類不能サンプル決定、グループ 1 違反及びグループ 2 違反に基づくアルゴリズムのデータセットにおいてサンプルが除外されたか、又は除外されなかった（例えば含まれた）ことを示す（例えば、実施例 1～4 を参照）。表 29 はまた、本発明の診断アルゴリズムを訓練及び検証する前に決定されたサンプルの既知の診断も示す（表 29、列「Dx 3」参照）。

【0431】

表 29：訓練セット及び検証セットにおけるサンプルの予測診断

30

【表 29】

サンプル 番号	対象識別子	データ セット	Dx3	IBD 除外理由	UC 除外理由	RF Dx4
1	10-IBD-07.059-081	1	1	6 か月を超える	6 か月を超える	0
2	10-IBD-07.079-002	1	2	6 か月を超える	6 か月を超える	0
3	10-IBD-01.009-024	1	1	不一致 ANCA/pANCA2	不一致 ANCA/pANCA2	1
4	10-IBD-01.009-026	1	1	不一致 ANCA/pANCA2	不一致 ANCA/pANCA2	1
5	10-IBD-07.040-004	1	0	含まれる	非 IBD	0
6	10-IBD-07.040-006	1	0	含まれる	非 IBD	0
7	10-IBD-01.005-031	1	1	含まれる	分類不能	3
8	10-IBD-01.013-028	1	1	含まれる	分類不能	3
9	10-IBD-07.059-085	1	2	含まれる	グループ 1 違反	1
10	10-IBD-07.036-021	1	2	含まれる	グループ 1 違反	1
11	10-IBD-01.002-008	1	1	含まれる	グループ 2 違反	2
12	10-IBD-01.011-007	1	1	含まれる	グループ 2 違反	2
13	10-IBD-01.001-026	1	1	含まれる	含まれる	1
14	10-IBD-01.001-028	1	1	含まれる	含まれる	1
15	10-IBD-07.118-074	2	1	臨床除外基準	臨床除外基準	1
16	10-IBD-07.118-098	2	2	臨床除外基準	臨床除外基準	3
17	10-IBD-07.319-164	2	0	不一致 ANCA/pANCA2	不一致 ANCA/pANCA2	0
18	10-IBD-07.319-169	2	0	不一致 ANCA/pANCA2	不一致 ANCA/pANCA2	0
19	10-IBD-07.319-156	2	0	含まれる	非 IBD	0
20	10-IBD-07.319-161	2	0	含まれる	非 IBD	0
21	10-IBD-07.324-004	2	2	含まれる	分類不能	3
22	10-IBD-07.320-011	2	1	含まれる	分類不能	3
23	10-IBD-07.136-070	2	2	含まれる	グループ 1 違反	1
24	10-IBD-07.136-083	2	2	含まれる	グループ 1 違反	1
25	10-IBD-07.118-102	2	1	含まれる	グループ 2 違反	2
26	10-IBD-07.118-104	2	1	含まれる	グループ 2 違反	2
27	10-IBD-07.321-003	2	1	含まれる	含まれる	0
28	10-IBD-07.323-011	2	1	含まれる	含まれる	1

【0432】

[0417] サンプル番号 1 及び 2 は除外ルールを満たしたため、s g i アルゴリズムの両方のランダムフォレストモデルを構築するために使用される訓練セットから、これらのサンプルを除外した。サンプル番号 7 及び 8 は適切な条件を満たしたため、これらのサンプルを用いて I B D 対非 I B D モデルを構築し、分類不能サンプルであると決定した。サンプル番号 9、10、11 及び 12 はグループ 1 違反（番号 9 及び 10）又はグループ 2 違反（番号 11 及び 12）に関する条件を満たしたため、U C 対 C D モデルを構築するための訓練セットから、これらのサンプルを除外した。サンプル番号 5 及び 6 は訓練セット中にあり、非 I B D を有すると予測された。サンプル番号 13 及び 14 は第 1 のランダムフォレストモデルでは I B D を有し、第 2 のモデルでは C D を有すると予測された。サンプル番号 15～28 は検証セットのものである。サンプル番号 15～18 は除外条件を満たしたため、それらを除外した。サンプル番号 19 及び 20 は s g i アルゴリズムによって非 I B D を有すると予測されたが、これは既知の診断又は以前の診断と合致する。サンプル番号 21 及び 22 は I C を有すると予測されたが、その既知の診断はそれぞれ U C 及び C D であった。サンプル番号 27 及び 28 は C D の既知の診断を有するが、s g i アルゴリズムは、番号 27 が非 I B D を有し、番号 28 が C D を有すると予測した。

【0433】

[0418] 本実施例は、s g i アルゴリズムが非 I B D、I C、U C、又は C D 診断を予測

するための妥当なモデルであることを実証する。本実施例はまた、血清学的、遺伝子、及び炎症マーカーの組み合わせた使用が I C、U C 及び C D などの I B D サブタイプを鑑別し、U C 及び C D に関して不確定である I B D を同定するための改善された診断試験を提供することも実証する。

【0434】

実施例 7 (I B D s g i 診断アルゴリズム訓練及び検証セット) :

[0419] 本実施例は、ランダムフォレストモデリングにおける訓練セット及び検証セットの使用を説明する。コホートからのサンプルのデータの訓練セットを使用するコンピュータ計算モデルを訓練するプロセスによってランダムフォレストモデリングアルゴリズムを作出する。訓練されたアルゴリズムを、コホートからのサンプルのデータの検証セットを用いて検証する。

10

【0435】

[0420] 図 16 は、訓練及び／又は検証コホート中に存在する血清学的マーカーの各々の % を示す。左側のベン図は、C D における C B i r 1、F l a 2 及び F l a X 抗体の重複を示す。右側のベン図は、非 I B D における C B i r 1、F l a 2 及び F l a X 抗体の重複を示す。

【0436】

[0421] 図 17 は、訓練及び／又は検証コホート中に存在する炎症マーカーの各々の % を示す。左側のベン図は、I B D における C R P 及び S A A の重複を示す。中央のベン図は、非 I B D における C R P 及び S A A の重複を示す。右側のベン図は、健常対照における C R P 及び S A A の重複を示す。

20

【0437】

[0422] 図 18 は、訓練及び／又は検証コホート中の 2 つ以上の遺伝子マーカーに関して陽性である患者数を示す。結果は、2 つ以上の遺伝子マーカーを有する I B D 患者の % が非 I B D 患者よりも高いことを示す。傾向は統計的に有意である。

【0438】

[0423] さらに、本明細書に提示される結果は、I B D の発症が、炎症及び組織損傷をもたらす遺伝的に疑われる人々における多様な抗原（例えば、酵母、細菌、フラジェリン、自己抗体）に対する脱調節された免疫学的応答を反映することを示す。

【0439】

[0424] 3 クラスのマーカー（血清、遺伝子及び炎症マーカー）間の複雑な相互作用を分析するランダムフォレストモデルは、単一のクラスのマーカー（例えば血清マーカー）を使用するよりも、I B D 対非 I B D 患者及び C D 対 U C 患者を鑑別する検出パターンにおいてより正確である。

30

【0440】

実施例 8 (血清学的、遺伝子及び炎症 (s g i) マーカーの組み合わせは非 I B D、クローン病、及び潰瘍性大腸炎患者を正確に鑑別する) :

[0425] 本実施例は、非 I B D、クローン病、及び潰瘍性大腸炎を正確に鑑別するための本発明の方法の使用を説明する。本実施例はまた、I B D の同定及び階層化を改善することができる確立された、及び新しいバイオマーカーの組み合わせを同定した試験も記載する。

40

【0441】

[0426] I B D の診断は、臨床検査、画像診断、組織病理に関する内視鏡検査、及び実験室試験の組み合わせに基づく、複雑なものである。血清学的マーカーは、重要な付加情報を提供することができる。I B D 診断を補助するための血清学の使用は、文献に広く記載されている。典型的には I B D と関連する自己-及び抗-微生物抗体に加えて、遺伝的バリエーション並びに血管新生及び炎症分子が I B D の同定をより良好に支援することができる。I B D 診断のための理想的な手段は、(1) 他の胃腸 (G I) 障害を有する患者から I B D 患者を識別する、及び (2) C D から U C を識別する。本明細書に記載の試験においては、血清学的バイオマーカーのパネル（例えば、A S C A - A、A S C A - G、A N C

50

A、pANCA、抗OmpC、抗CBir1)の診断正確度を、4つの遺伝子バリエーション(例えば、ATG16L1 rs2241880、NKX2-3 rs10883365、ECM1 rs3737240、及びSTAT3 rs744166)、炎症マーカー(例えば、CRP、SAA、ICAM、VCAM、VEGF)、及び血清学的マーカー(例えば、ASCA-A、ASCA-G、ANCA、pANCA、抗OmpC、抗CBir1、抗A4-Fla2、抗Flax)を含む17種の血清、遺伝子、及び炎症(sgi)バイオマーカーのパネルと比較した。

【0442】

[0427]よく特徴付けられた患者サンプルを、北米における8の研究機関及び42の地域医療機関から収集した。1,520人の患者からのサンプルが含まれていた:572人のCD、328人のUC、183人の健常なボランティア対照、及び過敏性腸症候群、肝炎、慢性便秘、GERD、セリアック病、憩室炎などを含む437人の非IBD疾患対照。全コホートから選択された1,083人の患者訓練セットに由来するデータを用いて、診断アルゴリズムを構築した。最終的なアルゴリズムの性能を、437人の患者の別の検証セットを用いて測定した。受信者動作特性(ROC)曲線下面積(AUC)を、連続変数としてランダムフォレストスコア(例えば、各分類についての決定木の割合)を用いて算出した。訓練中に選択されたカットオフスコアを適用して、感度及び特異度を算出した。

【0443】

[0428]ROC分析を用いて、血清学的バイオマーカーのみのパネルの診断正確度と、4つの遺伝子バリエーション(例えば、ATG16L1 rs2241880、NKX2-3 rs10883365、ECM1 rs3737240、及びSTAT3 rs744166)、5つの炎症マーカー(例えば、CRP、SAA、ICAM、VCAM、及びVEGF)、及び2つの他の抗微生物血清学的マーカー(例えば、A4-Fla2及びFlaxに対する抗体)を含む17種のマーカーパネルとを比較した。さらなるマーカーは、IBD対非IBD鑑別、AUCを0.80(95%CI=0.7682~0.8507)から0.87(95%CI=0.8442~0.9070)まで増加させた($p<0.001$)。UC対CD鑑別は、0.78(95%CI=0.7144~0.8403)から0.93(95%CI=0.8585~0.9354)まで増加した($p<0.001$)。図19及び20を参照されたい。IBD対非IBD並びにUC対CDの鑑別の両方の増加は統計的に有意であると決定された($p<0.001$)。IBDについては18%、CDについては18%、及びUCについては29%の誤差率の低下があった。疾患可能性を同定する性能における感度は、IBDについては74%、CDについては89%、及びUCについては98%であった。疾患を確認する性能における特異度は、IBDについては90%、CDについては81%、及びUCについては84%であった。

【0444】

[0429]これらの結果は、血清、遺伝子、及び炎症マーカーの組み合わせを用いて、血清学的マーカーのみよりも高い正確度で、IBD、CD及びUC患者をより良好に分類することができることを示している。

【0445】

実施例9(pANCA決定木):

[0430]本実施例は、pANCA2について陰性である患者標本の集団における診断予測を改善するために、例えば、偽陽性IBD予測率を最小化するために開発された決定木を例示する。いくつかの実施形態では、pANCA2定義は、陰性(0若しくは「検出されず」)又は弱い(+1)pANCA決定を有するすべての標本を含む。臨床的に特性評価され、生検により確認された患者標本の大きい群上でのその決定木及びその性能特性を開発するための方法が本明細書に記載される。

【0446】

開発方法:

[0431]試験対象患者基準/試験除外基準を満たした元のpANCA2(一)訓練及び検証標本からの実験室データを、新しい決定木を開発、試験及び検証する目的で新しいファ

10

20

30

40

50

イル中で解析した。元の pANCA 2 (－)、ANCA ELISA > 11.9 EU/mL 訓練のみのサブグループからの実験室データを用いて、新しい決定木を開発したが、残りは検証及び性能測定基準の算出のために取っておいた。生検（参照）結果に対する疾患予測を比較する共通の混同マトリックスを用いて、決定木内の潜在的なルールの効果を評価した。既知の UC 標本に関する CD 予測（及びその逆）は偽陽性とスコア化されたが、既知の CD 又は UC 標本のいずれかに関する不確定 IBD（IC）予測は評価のために真陽性とスコア化されたことに留意することが重要である。このルールは、IC 予測を全標本の 5% 未満に限定するために開発された。

【0447】

[0432] この群の標本を規定する主な特徴は、好中球抗核抗体（ANCA）に関する低い又は検出不可能な血清学的証拠である。IFA 又は ELISA アッセイにおける ANCA 反応性は、潰瘍性大腸炎（UC）の予測のための主要な決定因子であり、したがって、新しい決定木を開発するために取り組んだのは、UC 予測を行うための規定のルールを適用する前に、任意の特定の標本に関して、その標本における非 IBD 及び／又はクローン病（CD）のパターンを同定することによって UC 予測をまず除外することであった。

【0448】

直接クローン病予測：

[0433] 第 1 に、選択された標本からのデータを再検討し、個々の CD マーカー（例えば、ASCA-A、ASCA-G、OmpC、Cbir1、Fla2、FlaX）の各々に関する潜在的な値を、正確な予測を誘導するその能力について個別に評価した。反復評価により、ASCA 又は OmpC のいずれかに関する特定のアッセイの値が、過剰である場合、これを用いて CD 予測を行うことができると決定された。ASCA-IgA ≥ 69 EU/mL、ASCA-IgG ≥ 40 EU/mL 又は OmpC ≥ 60 EU/mL のいずれかである場合、標本は CD であると予測され、その標本はさらなる考慮から除去された。

【0449】

「計数」の使用による CD 推定：

[0434] 次に、複数の CD マーカーの組み合わせを、CD 予測を誘導するその能力について評価した。CD「計数」の概念を用いて、CD 予測を独立にもたらしことができる CD マーカーの複数の潜在的な組み合わせを効率的に管理した。再度、各アッセイについて記述された参照範囲に基づくものであるスコアリングの公正で単純なスキームをもたらし、高度に反復的なプロセスを用いた。最終的な CD 計数手順は ASCA-A、ASCA-G 及び OmpC マーカーを等しく計量するが、その対応する参照範囲を超える少なくとも 2 つのフラジェリンマーカーが CD 計数に対して正の影響を有することが必要である。こうして、その対応する参照上限に等しいか、又はそれを超える任意の CD マーカー（又はフラジェリンの対）は、CD カウンターにおける +1 の増加をもたらす。さらに、任意のフラジェリンが 100 EU/mL 以上のアッセイ値を有することが分かった場合、さらなる +1 を CD カウンターに加える。この ≥ 100 EU/mL ルールを満たすフラジェリンの数にもかかわらず、総計数は +1 だけ増加する。最終的な CD 計数が 2 以上である任意の標本を CD と指定する。

【0450】

pANCA 陰性標本：

[0435] CD 予測が上記手順の使用により行われなかった任意の残りの訓練標本を、さらなる分析のために pANCA (－)（検出されず）及び pANCA +1（弱く検出される）サブグループに分割した。pANCA (－) 群において、ANCA ELISA 参照範囲の使用は、非 IBD 標本の指定において有用であることが分かった。この指定の限界は、この予測における性能を平衡化する記述された上側の参照間隔のわずかに上に設定された。また、UC 予測を行うために ANCA ELISA 値限界を設定することができるが、この値は直接的な非 IBD 及び UC 予測の間にギャップを残したことも決定された。クリーンな非 IBD 対 UC 予測を行うことができなかった ANCA ELISA 値ギャップは、既知の非 IBD、CD 及び UC 標本の複雑な混合物を含むことが分かった。そこで、

不確定 I B D 指定を C D 計数 = 1 を有する任意の標本に適用したが、このギャップ中の他のすべてのものを非 I B D と指定した。

【0451】

p A N C A + 1 標本：

[0436]最後に、残りの p A N C A + 1 標本からのデータを再検討した。記述された上側の参照間隔よりも低い A N C A E L I S A 値を用いて、この群における U C 予測から非 I B D を効率的に区別することができると決定された。

【0452】

決定木：

[0437]ルールセットは次の通りである。p A N C A 2 陰性（すなわち、p A N C A = 0 又は + 1）であると分かった I B D s g i （血清、遺伝子及び炎症）プロセッシングに提出された任意の患者標本を、s g i アルゴリズムキューから除去し、臨床予測を行うために以下の決定行列にかけた。

1) 直接クローン病 (C D) 予測

a. $A S C A - I g A \geq 69 \text{ EU/mL}$ である場合、結果を「I B D : クローン病に一致するパターン」と指定する。あるいは

b. $A S C A - I g A \geq 40 \text{ EU/mL}$ である場合、結果を「I B D : クローン病に一致するパターン」と指定する。あるいは

c. $O m p C - I g A \geq 60 \text{ EU/mL}$ である場合、結果を「I B D : クローン病に一致するパターン」と指定する。あるいは

2) 「計数」の使用による C D 推定

a. 総「計数」を決定する。

i. $A S C A - I g A \geq 8.5 \text{ EU/mL}$ (参照範囲より上) である場合、C D 計数に + 1 を加える。

i i. $A S C A - I g A \geq 17.8 \text{ EU/mL}$ (参照範囲より上) である場合、C D 計数に + 1 を加える。

i i i. $O m p C - I g A \geq 10.9 \text{ EU/mL}$ (参照範囲より上) である場合、C D 計数に + 1 を加える。

i v. 対応する参照範囲 ($C b i r 1 - I g G \geq 78.4 \text{ EU/mL}$ 、 $A 4 - F l a 2 - I g G \geq 44.8 \text{ EU/mL}$ 、 $F l a X - I g G \geq 33.4 \text{ EU/mL}$) より上のフラジェリンマーカーが 2 以下である場合、C D 計数に + 1 を加える。

v. 任意のフラジェリン $\geq 100 \text{ EU/mL}$ である場合、C D 計数に + 1 を加える。

b. 総 C D 計数 ≥ 2 である場合、結果を「I B D : クローン病に一致するパターン」と指定する。あるいは

3) p A N C A 陰性標本

a. p A N C A = 0 (検出されず) 及び A N C A E L I S A $< 20 \text{ EU/mL}$ である場合、結果を「I B D に一致しないパターン」と指定する。あるいは

b. p A N C A = 0 (検出されず) 及び A N C A E L I S A $> 27.4 \text{ EU/mL}$ である場合、結果を「I B D : 潰瘍性大腸炎に一致するパターン」と指定する。あるいは

c. p A N C A = 0 (検出されず) 及び $20 \leq A N C A E L I S A \leq 27.4 \text{ EU/mL}$ であり、上記 2 b に由来する C D 計数 = 0 である場合、結果を「I B D に一致しないパターン」と指定する。あるいは

d. p A N C A = 0 (検出されず) 及び $20 \leq A N C A E L I S A \leq 27.4 \text{ EU/mL}$ であり、上記 2 b に由来する C D 計数 = 1 である場合、結果を「I B D : クローン病対潰瘍性大腸炎に関して不確定に一致するパターン」と指定する。あるいは

4) p A N C A + 1 標本

a. p A N C A = + 1 及び A N C A E L I S A $< 13.7 \text{ EU/mL}$ である場合、結果を「I B D に一致しないパターン」と指定する。あるいは

b. p A N C A = + 1 及び A N C A E L I S A $\geq 13.7 \text{ EU/mL}$ である場合、

結果を「IBD：潰瘍性大腸炎に一致するパターン」と指定する。

【0453】

ルールの効果：

[0438] pANCA 2（－）標本からの無傷のデータセットに対する上記ルールの適用の累積的效果を、表30及び表32に記載する。表30は、ANCA ELISA値 ≥ 11.9 EU/mLである検証標本に対するこれらのルールの効果を実証するが、表32は、ANCA ELISA値 < 11.9 EU/mLであるこれらの検証標本からのデータを含む。この相違の理由は、ANCA ELISA値 < 11.9 EU/mLである標本からの予測は、アルゴリズムの性能がこれらの標本に対して十分であると考えられるため、IBD sgi アルゴリズムの性能測定基準の算出に含まれていたが、性能の算出からANCA ELISA値 ≥ 11.9 EU/mLである標本を除去し、その群における予測を行うための代替的な方法を開発することに同意されたということである。この変化は、IBD sgi アルゴリズムが、いくつかの理由から、ANCA ELISA値 ≥ 11.9 EU/mLであるpANCA 2（－）標本に関する予測を行うことができなかったことを意味すると解釈されるべきではない。ランダムフォレスト法は、本質的には任意の単一のマーカーに依拠しないパターン照合手段であるため、この方法から誘導されたアルゴリズムをさらに用いて、単一成分マーカーの不正確な（又は欠けている）値を有する標本に関する予測を行うことができる。さらに、IBD sgi アルゴリズムを17種のマーカーを用いて訓練し、予測的カットオフをAUC曲線よりも高く設定して、存在する場合、欠けているか、又は不正確な個々のマーカーの効果を改善させた。このように、検証標本から誘導される性能特性はpANCA 2（－）検証標本のサブセットを含まないが、そのANCA ELISA値により規定される、2つのpANCA（－）群におけるsgi アルゴリズムの実際の全体的な正確度は非常に類似している（表31及び33中のランダムフォレストアルゴリズムの性能を参照）。事実、これらの2つの群の間で算出されるIBD sgi 感度及び偽陽性率は事実上同一である。pANCA 2（－）、ANCA ELISA、 11.9 EU/mLの群はランダムフォレストアルゴリズムの訓練及び性能検証の両方に含まれていたため、ランダムフォレスト性能が性能検証に含まれていない他方の群に関して本質的に同一であったという事実は、後者の群について行われた臨床予測が前者の群について行われたものと同様に信頼できることを意味する。

【0454】

[0439]以下の表は、ANCA ELISAアッセイ値 ≥ 11.9 EU/mLであるpANCA 2（－）標本に対する、決定木予測（表30）対ランダムフォレストアルゴリズムにより行われた予測（表31）の効果を実証する混同マトリックスを含む。

【0455】

表30：決定木性能

【表30】

		臨床診断(生検)	
		陽性	陰性
決定木	陽性	138	49
	陰性	37	65

感度=(真陽性/(真陽性+偽陰性)) $\times 100=78.9\%$

特異度=(真陰性/(真陰性+偽陽性)) $\times 100=57\%$

正確度=((真陽性+真陰性)/合計) $\times 100=70.2\%$

偽陽性率=(偽陽性/偽陽性+真陰性) $\times 100=43\%$

【0456】

表31：ランダムフォレスト性能

10

20

30

40

【表 3 1】

		臨床診断(生検)	
		陽性	陰性
決定木	陽性	130	86
	陰性	27	46

感度 = (真陽性 / (真陽性 + 偽陰性)) × 100 = 82.8%

特異度 = (真陰性 / (真陰性 + 偽陽性)) × 100 = 34.9%

正確度 = ((真陽性 + 真陰性) / 合計) × 100 = 60.9%

偽陽性率 = (偽陽性 / (偽陽性 + 真陰性)) × 100 = 65.2%

10

【0 4 5 7】

[0440] 決定木の感度はランダムフォレストアルゴリズムよりもわずかに低い、決定木の全体的な正確度は実質的により高く、それぞれ 70.2% 対 60.9% であり、偽陽性率もランダムフォレストアルゴリズムと比較して決定木法において実質的に低く、それぞれ 43% 対 65.2% であることに留意されたい。

【0 4 5 8】

[0441] 表 3 2 は、ANCA ELISA アッセイ値 < 11.9 EU/mL である pANCA 2 (一) 標本に対する、決定木予測対ランダムフォレストアルゴリズムにより行われた予測 (表 3 3) の効果を実証する混同マトリックスを含む。

20

【0 4 5 9】

表 3 2 : 決定木性能

【表 3 2】

		臨床診断(生検)	
		陽性	陰性
決定木	陽性	271	54
	陰性	289	589

30

感度 = (真陽性 / (真陽性 + 偽陰性)) × 100 = 48.4%

特異度 = (真陰性 / (真陰性 + 偽陽性)) × 100 = 91.6%

正確度 = ((真陽性 + 真陰性) / 合計) × 100 = 71.5%

偽陽性率 = (偽陽性 / (偽陽性 + 真陰性)) × 100 = 8.4%

【0 4 6 0】

表 3 3 : ランダムフォレスト性能

【表 3 3】

		臨床診断(生検)	
		陽性	陰性
決定木	陽性	213	311
	陰性	287	392

40

感度 = (真陽性 / (真陽性 + 偽陰性)) × 100 = 42.6%

特異度 = (真陰性 / (真陰性 + 偽陽性)) × 100 = 55.8%

正確度 = ((真陽性 + 真陰性) / 合計) × 100 = 50.3%

偽陽性率 = (偽陽性 / (偽陽性 + 真陰性)) × 100 = 44.2%

【0 4 6 1】

[0442] この群の標本において、決定木の感度はランダムフォレストアルゴリズムよりもわずかに高いが、決定木法の特異度及び全体的な正確度はランダムフォレストアルゴリズム

50

ムよりも実質的に高い。決定木法は偽陽性に対して真陰性予測を効率的に平衡化し、偽陽性率を最小化しながら、特異度を劇的に改善させる。

【0462】

結論：

[0443]決定木法を含む提唱されたルールは、pANCA 2（－）の特定のサブグループ状況ではランダムフォレストアルゴリズムよりも正確なIBD sgi予測をもたらすと上記データから結論付けられる。偽陽性率は有意に低く、ここで対象の標本は既知のIBDマーカーに対する血清反応性が最小である非常に初期の疾患である可能性が高いため、この観察の重要性は無視されるべきではない。偽IBD予測がなされる率を最小化して、IBDを有さない患者に対する不必要なリスク（例えば、さらなる診断手順又は不必要な介入に由来するリスク）を回避することが重要である。

10

【0463】

[0444]いくつかの態様では、本明細書に記載の決定木法を、pANCA 2（－）アッセイ結果を有する臨床標本における予測診断を行うための代替経路として実装することができる。特定の実施形態では、改変されたアルゴリズムは、そのpANCA 2値に応じて、決定木法又はランダムフォレストアルゴリズムのいずれかに対して個々の標本を指向させる第1のステップとしての初期ノードを含む。

【0464】

実施例10（遺伝子、血清及び炎症マーカーに関するアッセイの説明）：

[0445]本実施例は、本発明のIBD診断アルゴリズムにおいて用いられる血清－遺伝子－炎症（sgi）マーカーを検出又は測定するためのアッセイの例を記載する。

20

【0465】

A. SNPジェノタイピングのためのTaqMan法の概説

[0446]一部の実施形態では、IBD sgi診断パネルにおけるSNPジェノタイピングアッセイは、TaqMan SNPジェノタイピングに基づく。定量的ジェノタイピングアッセイは、ヒト血液ゲノムDNA（gDNA）又はライセートサンプルを用いて、IBD（特にCD及びUC）の診断及び／又は鑑別診断に有用であり得る。これらの対立遺伝子識別PCRアッセイは、配列の5'側に2つの異なる蛍光染料を含む2つの特異的オリゴヌクレオチド配列（例えば、FAM染料又はVIC染料を含む蛍光性プローブ）であって、それらの各々が、副溝結合剤（融点エンハンサー）に連結された配列の3'側に非蛍光クエンチャーを有する前記配列を使用する。

30

【0466】

[0447]PCR増幅の間に、各プローブは標的DNA上のフォワード及びリバースプライマーの間でその相補的配列に特異的にアニーリングする。DNAポリメラーゼは固有の5'ヌクレアーゼ活性を有するため、ゲノム配列にハイブリダイズしたプローブの選択的切断が起こる。これは、クエンチャーからのリポーター染料の分離に起因する蛍光の増加をもたらす。従って、ある染料対他の染料（FAM対VIC）の選択的増加は、対立遺伝子が試験したゲノムDNA中に存在することを示す。

【0467】

B. 対立遺伝子識別

40

[0448]各プローブの染料の相対蛍光強度の検査によってサンプルの遺伝子型を決定することができる。ABIのSDSソフトウェアを用いて、2つの染料の強度のグラフプロットを作出することができる。ホモ接合性の野生型遺伝子型を、FAMの相対的に高い強度（y軸）及びVICの非常に低い強度（x軸）によって決定する。ホモ接合性の変異遺伝子型は、高いVIC（x軸）、低いFAM（y軸）を有する、類似するが反対のパターンを示す。データが分析されたら、ABIソフトウェアを用いて対立遺伝子「コール」を作出し、それぞれのSNPに関する遺伝子型を決定する。

【0468】

[0449]いくつかの実施形態では、SNPジェノタイピングアッセイは、読み取り前及び読み取り後のデータ収集のためのABI 7500 Fastの使用を、Applied

50

Biosystems Veriti 96-Well Fast Thermal Cycler (ABI Veriti) 上でのオフラインPCR増幅ステップと組み合わせるワークフローを用いる。この組み合わせにより、1人の実験室操作者が、複数のABI Veritiマシン上での複数のジェノタイピングの実行（プレート当たり最大で96個のサンプル）、次いで、1つのABI 7500 Fast装置上での各反応プレートの迅速な事後読み取りを同時に実施することができる。SNPアッセイについては、PCR反応混合物を自動又は手動でTecan Freedom EVO 100自動ワークステーションにより96ウェルプレート中で調製し、予備読み取りをABI 7500 FastリアルタイムPCRシステム上で実施したら、反応プレートをオフライン増幅ステップのためにABI Veritiに移す。増幅が完了した時、事後読み取り及びデータ分析のために反応プレートをABI 7500 Fast装置に移し戻す。このワークフローにより、増幅前及び増幅後測定のために単一のABI 7500 Fast装置を用いて複数のプレートを同時に増幅することができる。

10

【0469】

C. アッセイデータ報告

[0450] IBD sgi 診断試験は、免疫応答に影響する遺伝的感受性を評価するための4つのSNPジェノタイピングアッセイ（表34参照）を含む。

【0470】

表34：SNPアッセイデータ出力

【表34】

20

アッセイ名	SNP ID	ABI 読み出し(csv ファイル)	遺伝子型(コール)
ATG16L1 SNP	rs2241880	VIC_1	ホモ接合性 AA
		両方	ホモ接合性 AG
		FAM_1	ホモ接合性 GG
ECM1 SNP	rs3737240	VIC_2	ホモ接合性 CC
		両方	ホモ接合性 CT
		FAM_2	ホモ接合性 TT
NKX2-3 SNP	rs10883365	VIC_10	ホモ接合性 AA
		両方	ホモ接合性 AG
		FAM_10	ホモ接合性 GG
STAT3 SNP	rs744166	VIC_11	ホモ接合性 AA
		両方	ホモ接合性 AG
		FAM_11	ホモ接合性 GG

30

【0471】

D. ATG16L1 SNP rs2241880ジェノタイピングアッセイ

[0451] ATG16L1 遺伝子は第2染色体上に位置し、自食作用の間のオートファゴソームの形成に関与することが知られるタンパク質をコードする。ATG16L1 1155 (A>G) SNP (SNP ID: rs2241880) は、AA位置についてT300AであるcDNA配列のコード領域位置1155A>G内に位置する。ATG16L1 SNP rs2241880ジェノタイピングアッセイは、全血サンプルから調製されたゲノムDNA (gDNA) 又はライセートサンプルを用いてIBD患者におけるATG16L1 1155 (A>G) SNPを検出するための実験室で開発された試験 (LDT) である。この定性的ジェノタイピングアッセイはIBDの診断に有用である。

40

【0472】

[0452] 各遺伝子型について、以前にジェノタイピングされた血液ライセート又はgDNAサンプルを各試行のための対照として用いる。これらの対照の遺伝子型を、両方の鎖上で配列決定することにより確認する。非鋳型対照 (NTC) も各試行に含まれる。

【0473】

50

[0453] g D N A 対照を、全血サンプルから単離される D N A、又は同じ全血サンプルから調製される血液ライセートから調製する。下記の 3 つの対照を各アッセイにおいて泳動する。

A T G 1 6 L 1 1 1 5 5 (A > G) S N P ジェノタイピングアッセイ対照 G G
A T G 1 6 L 1 1 1 5 5 (A > G) S N P ジェノタイピングアッセイ対照 A G
A T G 1 6 L 1 1 1 5 5 (A > G) S N P ジェノタイピングアッセイ対照 A A

【 0 4 7 4 】

E . E C M 1 S N P ジェノタイピングアッセイ

[0454] E C M 1 遺伝子は第 1 染色体上に位置し、腸壁の障壁機能の維持に関与する細胞外マトリックスタンパク質 1 をコードし、N F - κ B 炎症経路を活性化する。E C M 1 5 5 8 (C > T) S N P (S N P I D : r s 3 7 3 7 2 4 0) は、A A 位置について T 1 3 0 M である c D N A 配列のコード領域位置 5 8 8 C > T 内に位置する。

【 0 4 7 5 】

[0455] E C M 1 S N P r s 3 7 3 7 2 4 0 ジェノタイピングアッセイは、全血サンプルから調製された g D N A 又はライセートサンプルを用いて I B D 患者における E C M 1 5 5 8 (C > T) S N P 遺伝子型を検出するための実験室で開発された試験 (L D T) である。この定性的ジェノタイピングアッセイは I B D の診断に有用である。

【 0 4 7 6 】

[0456] 各遺伝子型について、以前にジェノタイピングされた血液ライセート又は g D N A サンプルを各試行のための対照として用いる。これらの対照の遺伝子型を、両方の鎖上で配列決定することにより確認する。非鋳型対照 (N T C) も各試行に含まれる。g D N A 対照を、全血サンプル、又は同じ全血サンプルから調製される血液ライセートから調製する。

E C M 1 5 5 8 (C > T) S N P ジェノタイピングアッセイ対照 T T
E C M 1 5 5 8 (C > T) S N P ジェノタイピングアッセイ対照 C T
E C M 1 5 5 8 (C > T) S N P ジェノタイピングアッセイ対照 C C

【 0 4 7 7 】

F . N K X 2 - 3 S N P ジェノタイピングアッセイ

[0457] N K X 2 - 3 遺伝子は第 1 0 染色体上に位置する。N K X 2 - 3 (A > G) S N P (S N P I D : r s 1 0 8 8 3 3 6 5) はイントロン中に位置する (位置 1 0 1 2 8 7 7 6 4) 。N K X 2 - 3 r s 1 0 8 8 3 3 6 5 ジェノタイピングアッセイは、全血サンプルから調製されるゲノム D N A (g D N A) 又はライセートサンプルを用いて I B D 患者において N K X 2 - 3 (A > G) S N P 遺伝子型を検出するための実験室で開発された試験 (L D T) である。この定性的ジェノタイピングアッセイは I B D の診断に有用である。

【 0 4 7 8 】

[0458] 各遺伝子型について、以前にジェノタイピングされた血液ライセート又は g D N A サンプルを各試行のための対照として用いる。これらの対照の遺伝子型を、両方の鎖上で配列決定することにより確認する。非鋳型対照 (N T C) も各試行に含まれる。D N A 対照を、全血サンプル、又は同じ全血サンプルから調製される血液ライセートから調製する。

N K X 2 - 3 (A > G) S N P ジェノタイピングアッセイ対照 G G
N K X 2 - 3 (A > G) S N P ジェノタイピングアッセイ対照 A G
N K X 2 - 3 (A > G) S N P ジェノタイピングアッセイ対照 A A

【 0 4 7 9 】

G . S T A T 3 S N P ジェノタイピングアッセイ

[0459] S T A T 3 遺伝子は第 1 7 染色体上に位置し、炎症性腸疾患 (I B D) などの種々の自己免疫障害において重要な役割を果たしている。S T A T 3 (A > G) S N P (S N P I D : r s 7 4 4 1 6 6) は、3 ' U T R の 5 ' フランキング領域に位置する (位置 4 0 5 1 4 2 0 1) 。S T A T 3 r s 7 4 4 1 6 6 ジェノタイピングアッセイは、全

10

20

30

40

50

血サンプルから調製される g D N A 又はライセートサンプルを用いて I B D 患者において S T A T 3 (A > G) S N P 遺伝子型を検出するための実験室で開発された試験 (L D T) である。この定性的ジェノタイピングアッセイは I B D の診断に有用である。

【0480】

[0460]各遺伝子型について、以前にジェノタイピングされた血液ライセート又は g D N A サンプルを各試行のための対照として用いる。これらの対照の遺伝子型を、両方の上で配列決定することにより確認する。

S T A T 3 (A > G) S N P ジェノタイピングアッセイ対照 G G

S T A T 3 (A > G) S N P ジェノタイピングアッセイ対照 A G

S T A T 3 (A > G) S N P ジェノタイピングアッセイ対照 A A

10

【0481】

H. F l a X アッセイ

[0461]一部の実施形態では、F l a X アッセイは組換え F l a X 細菌抗原を用いて、ヒト血清中の抗 F l a X I g G 抗体を測定する。このアッセイは I B D (特に C D 及び U C) の診断及び鑑別診断に有用である。

【0482】

[0462]一部の実施形態では、前記アッセイは、標準的な 96 ウェルプレートフォーマットでの自動化間接 E L I S A の実行である。最初に、組換え F l a X 抗原をプレートのウェル表面上に被覆した後、遮断する。希釈された患者血清 (1 / 100) をウェルに添加し、インキュベートする。未結合の分子を洗浄除去した後、アルカリホスファターゼで標識された検出抗体を、ウェル中でインキュベートする。2回目の洗浄後、比色シグナルの現像のために基質を添加する。停止溶液の添加時に、光密度 (O D) についてプレートを読み取る。標準曲線から抗 F l a X 抗体濃度を算出し、血清 1 ミリリットル当たりの E L I S A 単位 (E U / m L) で報告する。

20

【0483】

I. F l a 2 アッセイ

[0463]一部の実施形態では、F l a 2 アッセイは組換え F l a 2 細菌抗原を用いて、ヒト血清中の抗 F l a 2 I g G 抗体を測定する。F l a 2 アッセイは I B D (特に C D 及び U C) の診断及び鑑別診断に有用である。

【0484】

[0464]一部の実施形態では、前記アッセイは、標準的な 96 ウェルプレートフォーマットでの自動化間接 E L I S A の実行である。最初に、組換え F l a 2 抗原をプレートのウェル表面上に被覆した後、遮断する。希釈された患者血清 (1 / 100) をウェルに添加し、インキュベートする。未結合の分子を洗浄除去した後、アルカリホスファターゼで標識された検出抗体を、ウェル中でインキュベートする。2回目の洗浄後、比色シグナルの現像のために基質を添加する。停止溶液の添加時に、光密度 (O D) についてプレートを読み取る。標準曲線から抗 F l a 2 抗体濃度を算出し、血清 1 ミリリットル当たりの E L I S A 単位 (E U / m L) で報告する。

30

【0485】

J. V E G F アッセイ

[0465]一部の実施形態では、ヒト V E G F キットは、血清中のヒト V E G F を測定するためのサンドイッチ抗原検出 E L I S A アッセイである (T h e r m o S c i e n t i f i c I n c .) 。このアッセイは I B D (特に C D 及び U C) の診断及び鑑別診断に有用である。

40

【0486】

[0466]一部の実施形態では、前記アッセイは標準的な 96 ウェルプレートフォーマットで実施される。プレートを、製造業者による抗 V E G F 抗体 (捕捉抗体) で被覆する。試験時に、希釈された患者血清 (1 / 2) を V E G F 捕捉結合インキュベーションのためにウェルに添加する。未結合の分子を洗浄除去した後、ビオチン化検出抗体を添加し、それは V E G F の第 2 の部位に結合する。2回目の洗浄後、3, 3', 5, 5' - テトラメチ

50

ルベンジジン (TMB) と反応して比色シグナルを生成するストレプトアビジン-西洋わさびペルオキシダーゼを添加し、光密度 (OD) 読み取り装置により測定する。OD の値を、標準曲線に対して算出して、VEGF 濃度に関する値を生成し、血清 1 ミリリットル当たりのピコグラム (pg/mL) で報告する。

【0487】

[0467] 製造業者の推奨プロトコールに基づいて、Tecan Freedom EVO 150 システムを用いることにより、このアッセイを自動化した。

【0488】

K. CRP、SAA、ICAM1 及び VCAM1 アッセイ

[0468] 一部の実施形態では、ヒト血管損傷 II アッセイは、MULTI-ARRAY に基づく多重化サンドイッチ抗原検出アッセイである。このアッセイは、ヒト血清中の CRP、SAA、ICAM1 及び VCAM1 を測定するために用いられる (Meso Scale Discovery Inc.)。このアッセイは IBD (特に CD 及び UC) の診断及び鑑別診断に有用である。

【0489】

[0469] 一部の実施形態では、前記アッセイは、96 ウェルプレートフォーマットでのユニークな電気化学発光検出及びパターン化されたアレイを用いる。表面電極ウェル上の CRP、SAA、ICAM1 及び VCAM1 の各分析物に対する 4 つの捕捉抗体でプレートを被覆する。試験時に、希釈された患者血清 (1/1000) を、抗原捕捉インキュベーションのためにウェルに添加する。未結合の分子を洗浄除去した後、化学発光 (SULFO-TAG) 標識された検出抗体を添加し、これは分析物の第 2 の部位に結合する。2 回目の洗浄後、エネルギーを発光放射光シグナルに変換する電子励起を印加し、相対発光単位 (RLU) により測定することにより、プレートを読み取る。RLU 値を、標準曲線に対して算出して、濃度に関する値を生成し、CRP 及び SAA については血清 1 リットル当たりのミリグラム (mg/L)、又は VCAM1 及び ICAM1 については 1 ミリリットル当たりのマイクログラム (µg/mL) で報告する。

【0490】

[0470] 製造業者の推奨プロトコールに基づいて、Tecan Freedom EVO 150 システム及び PerkinElmer Janus システムを用いることにより、これらのアッセイを自動化した。

【0491】

実施例 11 (IBD sgi 試験のための参照範囲及び参照コール) :

[0471] 本実施例は、本発明の IBD 診断アルゴリズムにおいて検出又は測定された血清-遺伝子-炎症 (sgi) マーカーの各々のための参照範囲及び参照コールを例示する。いくつかの実施形態では、参照範囲又はコールは、特定の sgi マーカーに関するカットオフ値、切点、変異の存在若しくは非存在、又は IFA パターンの存在若しくは非存在に対応する。

【0492】

[0472] IBD sgi 試験のための参照範囲及び参照コールを個々のアッセイの各々に基づいて決定した (表 35 参照)。

【0493】

表 35 : IBS sgi アッセイのための参照範囲及び参照コール

10

20

30

40

【表 3 5】

アッセイ名	参照値／コール
ASCA IgA ELISA	<8.5 EU/mL
ASCA IgG ELISA	<17.8 EU/mL
抗 OmpC ELISA	<10.9 EU/mL
抗 CBir1IgG ELISA	<78.4 EU/mL
抗 A4-Fla2 IgG ELISA	<44.8 EU/mL
抗 Flax IgG ELISA	<33.4 EU/mL
ANCA ELISA	<19.8 EU/mL
IBD 特異的 pANCA IFA 核周囲パターン	検出されず
IBD 特異的 pANCA DNase 感受性	検出されず
ATG16L1 SNP (rs2241880)	A/A, A/G
ECM1 SNP (rs3737240)	C/C, C/T
NKX2-3 SNP (rs10883365)	A/A, A/G
STAT3 SNP (rs744166)	G/G, A/G
ICAM1	<0.54 µg/mL
VCAM1	<0.68 µg/mL
CRP	<13.2 mg/L
SAA	<10.9 mg/L
VEGF	<345 pg/mL

10

20

【0494】

[0473]血清学的及び炎症マーカーアッセイに関して、参照範囲は、参照範囲上限の10000個の推定値にわたる95%下側信頼区間(LCI)であり、参照範囲上限はマーカー測定値の中央95%の上側境界値である。各範囲限界のLCIを、健常対照のデータベースから繰り返し120人の対象を無作為に選択するシミュレーションを用いて決定した。

30

【0495】

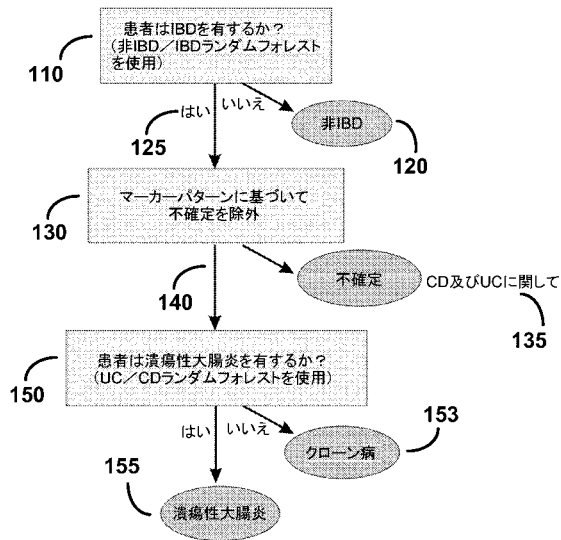
[0474]遺伝子アッセイ(ATG16L1 SNP rs2241880、ECM1 SNP rs3737240、NKX2-3 SNP rs10883365及びSTAT3 SNP rs744166)は、定性的ジェノタイピングアッセイである。ATG16L1 SNP rs2241880については、A/A又はA/Gの参照値／コールは変異が検出されないことに等しい。ECM1 SNP rs3737240については、C/C又はC/Tの参照値／コールは変異が検出されないことに等しい。NKX2-3 SNP rs10883365については、A/A又はA/Gの参照値／コールは変異が検出されないことに等しい。STAT3 SNP rs744166については、G/G又はA/Gの参照値／コールは変異が検出されることに等しい。

40

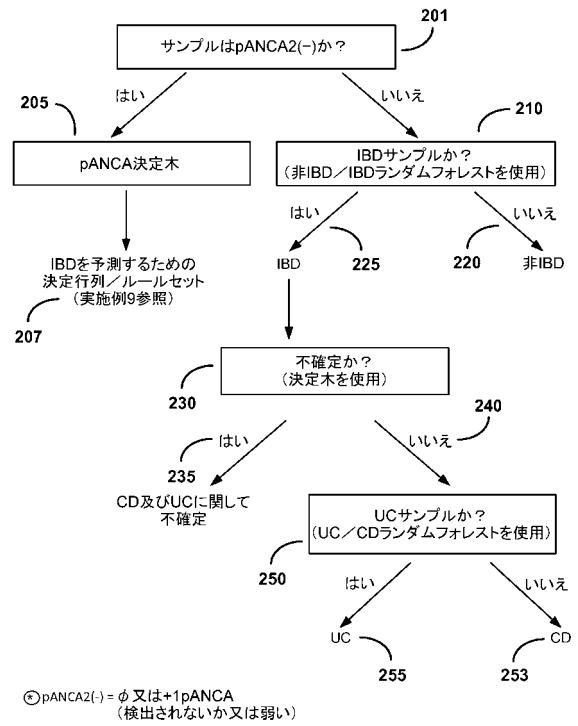
【0496】

[0475]本明細書の記載は例示的なものであり、限定的なものではないことが理解されるべきである。多くの実施形態は、本明細書の記載を読む際に当業者に明らかとなるであろう。それ故、本発明の範囲は、本明細書の記載を参照して決定されるべきではなく、添付の特許請求の範囲を、そのような特許請求の範囲の権利が及ぶ等価物の全範囲と共に参照して決定されるべきである。特許出願、特許、PCT公報、及びGenbank受託番号を含むあらゆる論文及び参考文献の開示は、参照によりすべての目的のために本明細書に組み入れられる。

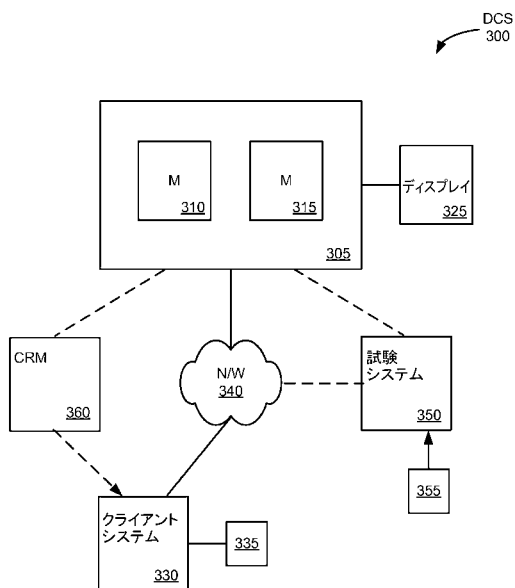
【図 1】



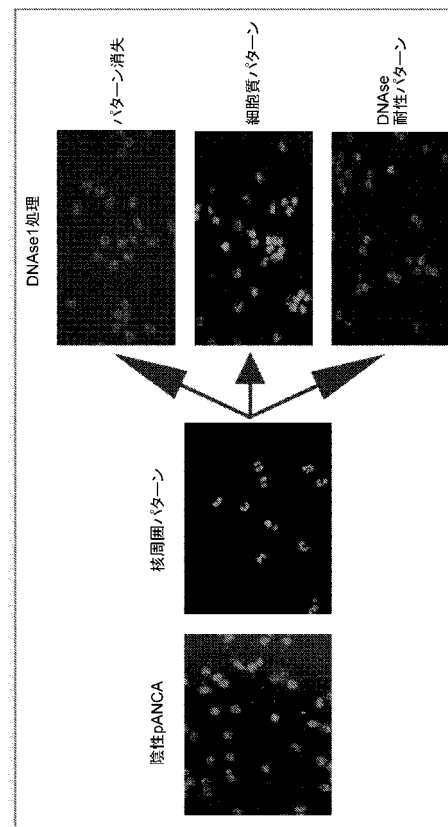
【図 2】



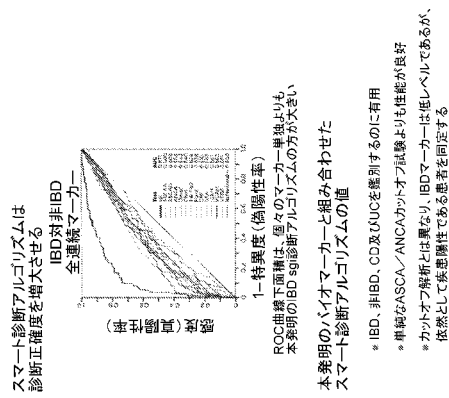
【図 3】



【図 4】



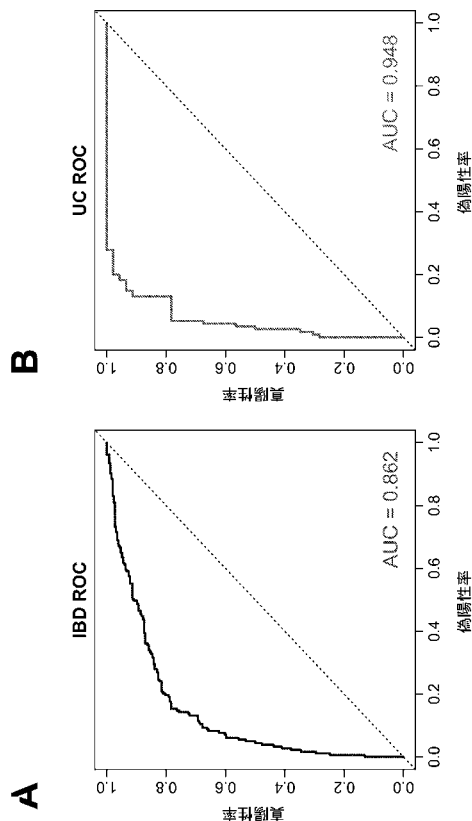
【図 5】



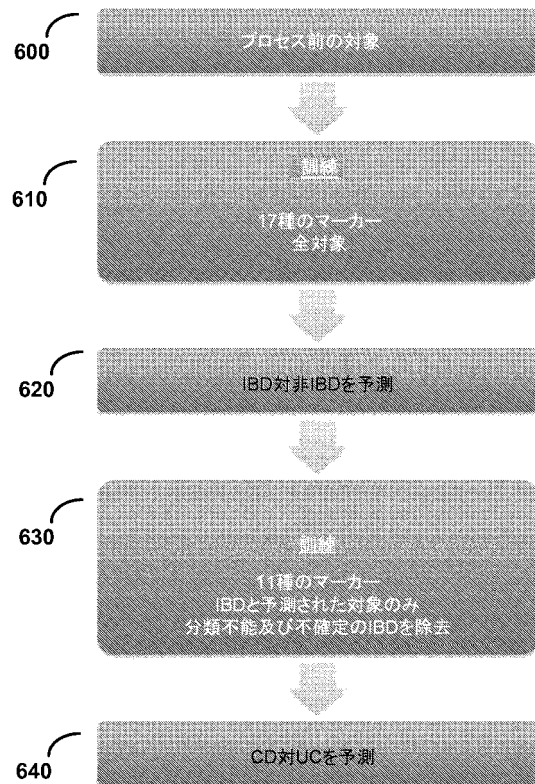
42の地理医療機関及び8の三次医療センターからの1520のよく特徴付けられたサンプル

訓練コホート		検証コホート		合計	
疾患のマーカーバターンを認識するようにアルゴリズムを訓練		アルゴリズムの性能を試験			
CD	419 (39%)	153 (35%)	572 (37%)	1083	437
UC	177 (17%)	101 (23%)	278 (18%)	1083	437
GI	514 (54%)	123 (28%)	637 (42%)	1083	437
合計	1110	377	1487	1083	437

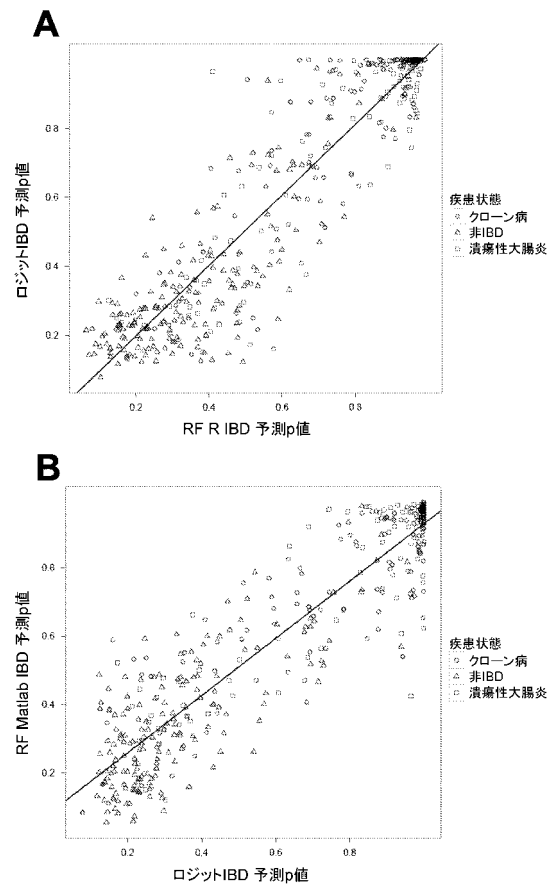
【図 7】



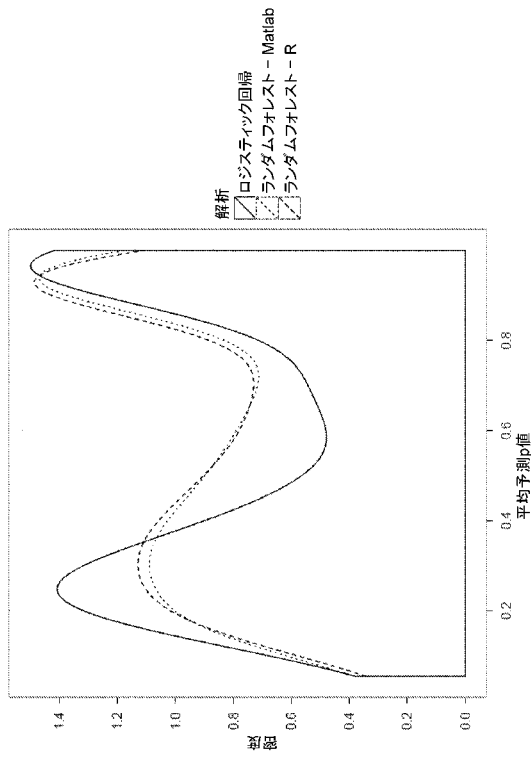
【図 6】



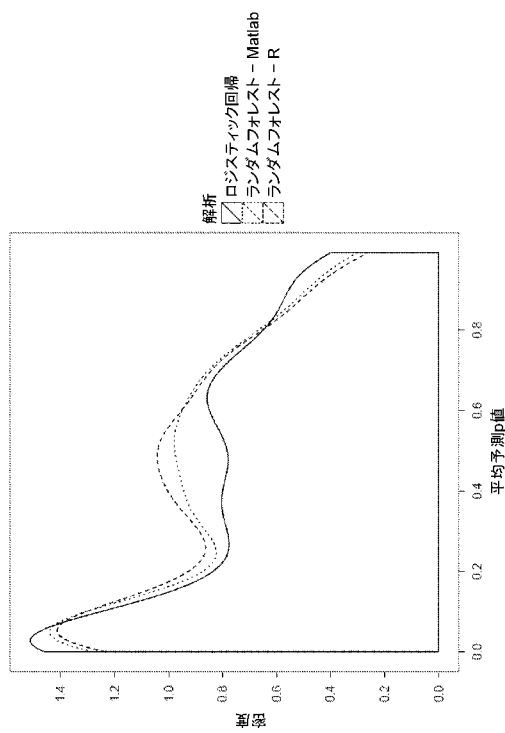
【図 8】



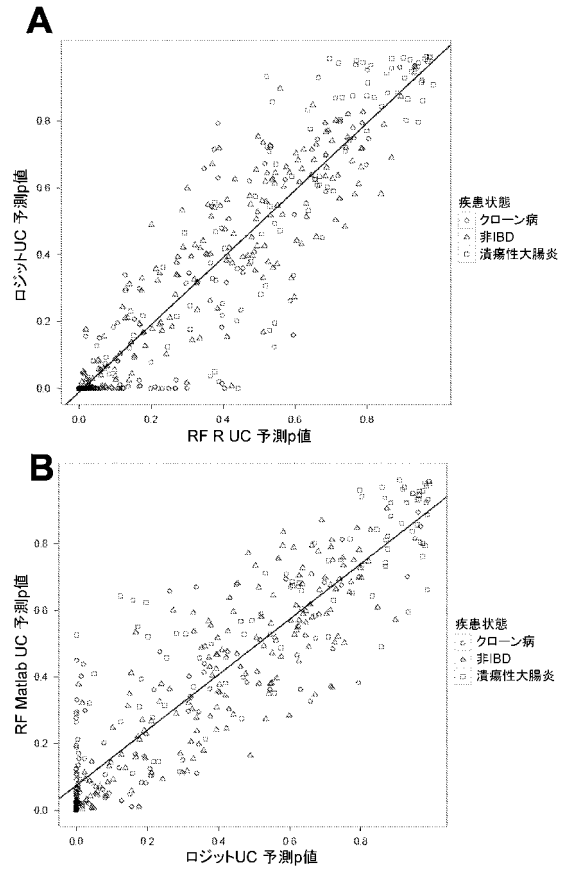
【図 9】



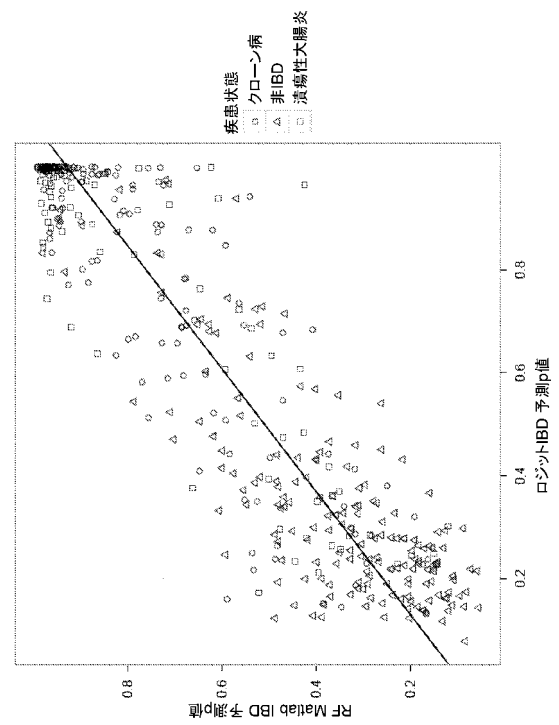
【図 1 1】



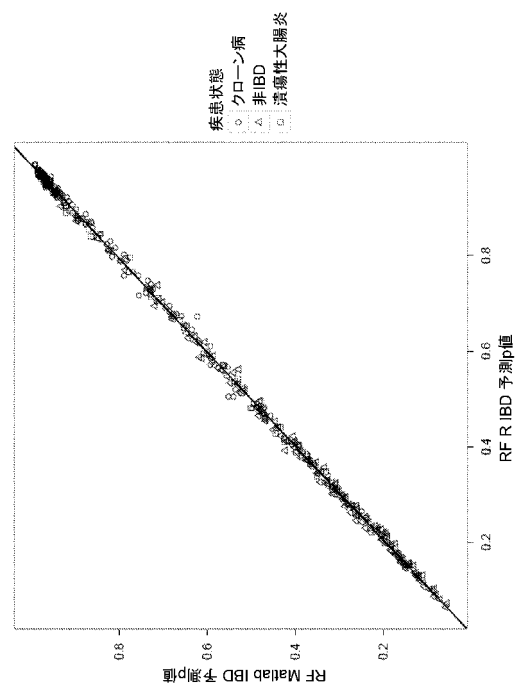
【図 1 0】



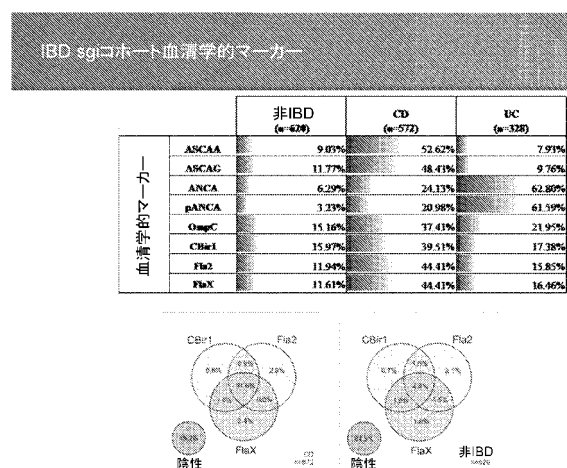
【図 1 2】



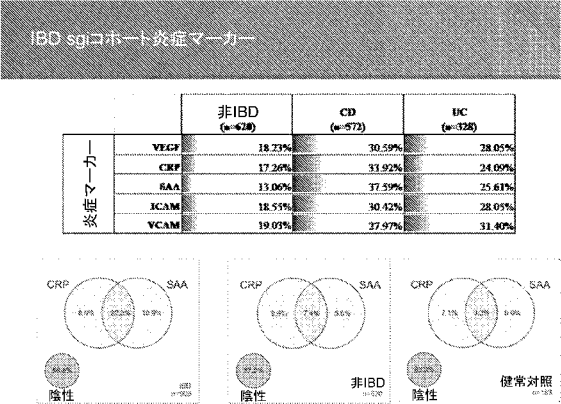
【図 14】



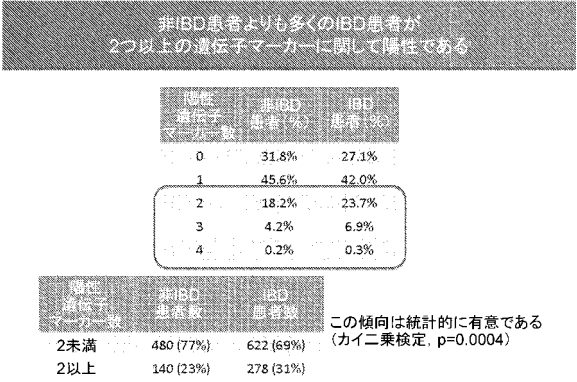
【図 16】



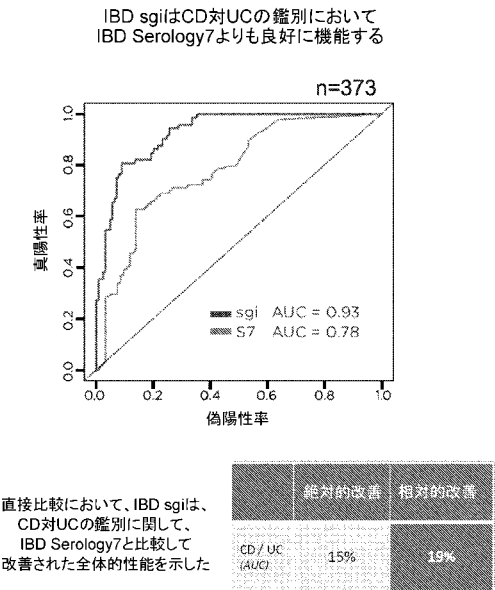
【図 17】



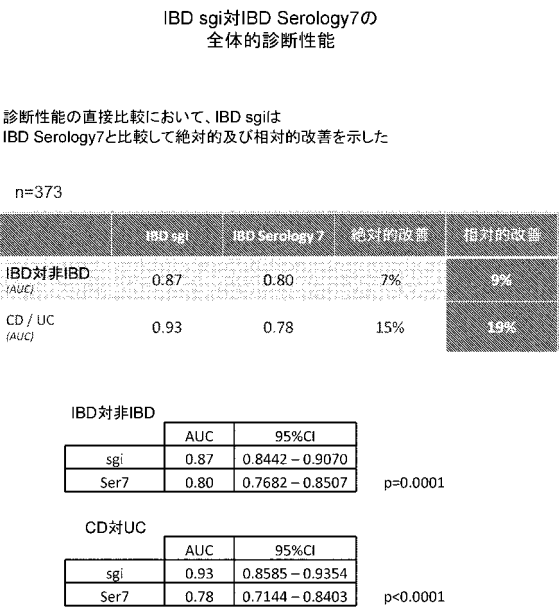
【図 18】



【図 19】



【図 20】



【配列表】

20155027400000001.app

【国際調査報告】

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 12/61202

A. CLASSIFICATION OF SUBJECT MATTER IPC(8) - G06F 19/00; G01N 33/48, 33/53 (2013.01) USPC - 702/19; 436/63; 435/4 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) IPC(8): G06F 19/00; G01N 33/48, 33/53 (2013.01) USPC: 702/19; 436/63; 435/4 Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched USPC: 436/501, 63; 435/6.1, 7.1, 4; 702/19 (keyword limited; terms below) Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) PatBase; Google; PubMed Search terms: inflammatory bowel disease, IBD, random forst, crohn's disease, ANCA, pANCA, pANCA2, ATG16L1, CRP, SAA, VEGF		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2011/0045476 A1 (BARKEN et al.) 24 February 2011 (24.02.2011) abstract; para [0034]-[0036], [0090]-[0092], [0094], [0118]-[0119], [0121]-[0125], [0177], [0195], [0233], [0277], [0307], [0359]-[0360], [0460]-[0461], [0534]-[0535], [0539], [0541]; Tables 1A, 26, 29	1-78
A	US 2008/0131439 A1 (LOIS et al.) 5 June 2008 (05.06.2008) entire document	1-78
A	US 2011/0218956 A1 (LOIS et al.) 8 September 2011 (08.09.2011) entire document	1-78
A	US 2011/0159521 A1 (GONG et al.) 30 June 2011 (30.06.2011) entire document	1-78
☐ Further documents are listed in the continuation of Box C. ☐		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search 12 March 2013 (12.03.2013)		Date of mailing of the international search report 22 MAR 2013
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-3201		Authorized officer: Lee W. Young PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 12/61202

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:
Please see extra sheet

1. ☒ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

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

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/US 12/61202

Continuation of Box III: Lack of Unity of Invention

This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees must be paid.

Groups I+: Claims 1-78, directed to a method for diagnosing inflammatory bowel disease (IBD) and/or a clinical subtype thereof in an individual, said method comprising:

(a) analyzing a sample obtained from said individual to determine the presence, level or genotype of one or more markers selected from the group consisting of a serological marker, a genetic marker, an inflammation marker, and a combination thereof in said sample to obtain a marker profile;

(b) applying a first random forest statistical analysis to said marker profile to obtain a decision whether said marker profile is an IBD sample or a non-IBD sample; wherein the first invention is limited to wherein the marker is a serological marker (claims 1, 3, 12-14 and 19-78) (applicants may opt for additional marker types to be searched by specifying the marker type and paying an additional invention search fee for each elected marker type).

The inventions listed as Groups I+ do not relate to a single general inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

The special technical feature of the claims of Groups I+ claims is the marker type used to determine a diagnosis for IBD, wherein each subgroup is directed to a different marker type.

The common technical elements shared by the above groups is that they are related to a method for diagnosing inflammatory bowel disease (IBD) and/or a clinical subtype thereof in an individual, said method comprising:

(a) analyzing a sample obtained from said individual to determine the presence, level or genotype of one or more markers selected from the group consisting of a serological marker, a genetic marker, an inflammation marker, and a combination thereof in said sample to obtain a marker profile;

(b) applying a first random forest statistical analysis to said marker profile to obtain a decision whether said marker profile is an IBD sample or a non-IBD sample; (c) applying a decision tree or set of rules to said sample designated as an IBD sample to determine if said IBD sample is an inconclusive sample; and (d) if the sample is not inconclusive, but is an IBD sample, applying a second random forest statistical analysis to said IBD sample to determine a clinical subtype. These common technical elements do not represent an improvement over the combined prior art of US 2008/0131439 A1 to Lois et al. (hereinafter "Lois") and US 2011/0251100 A1 to Li et al. (hereinafter "Li"), as follows:

Lois teaches a method for diagnosing IBD and/or a clinical subtype thereof in an individual (para [0006]-[0009]), said method comprising: (a) analyzing a sample obtained from said individual to determine the presence or level of one or more markers selected from a serological markers in said sample to obtain a marker profile (para [0007]-[0008] - anti-neutrophil antibody; para [0035] - ANCA); and (b) applying a first random forest statistical analysis to said marker profile to obtain a decision whether said marker profile is an IBD sample or a non-IBD sample (para [0009], [0064]). Although Lois does not explicitly recite applying a second random forest analysis to determine a clinical subtype for the sample, Lois teaches the use of a random forest (para [0148]) to determine whether a sample is associated with a clinical subtype of IBD (para [0146]), it would have been obvious to a person of ordinary skill in the art to use first random forest to determine a diagnosis of IBD, and a second random forest to further determine classification based on the teaching of Lois. Although Lois also does not expressly recite applying a decision tree or set of rules to said sample designated as an IBD sample to determine if said IBD sample is an inconclusive sample, in a related disclosure, Li teaches the use of machine learning to determine decision rules for classification of different IBD subtypes (para [0154]). It would have been obvious to a person of ordinary skill in the art to apply a set of decision rules, such as those developed via machine learning by Li, to indeterminate samples that did not produce an obvious classification based on the methods of Lois, in order to enable a more accurate diagnosis and determination of a course of effective treatment.

Therefore, the inventions of Groups I+ lack unity of invention under PCT Rule 13 because they do not share a same or corresponding special technical feature.

フロントページの続き

(51)Int.Cl. F I テーマコード (参考)
G 0 1 N 33/543 5 4 5 A
C 1 2 N 15/00 A

(31)優先権主張番号 61/570,271

(32)優先日 平成23年12月13日(2011.12.13)

(33)優先権主張国 米国(US)

(81)指定国 AP(BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), EA(AM, AZ, BY, KG, KZ, RU, TJ, TM), EP(AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG), AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, 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, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC

(特許庁注：以下のものは登録商標)

1. F I R E W I R E
2. J A V A
3. J A V A S C R I P T

(72)発明者 プリンセン, フレッド
アメリカ合衆国, カリフォルニア州, ラ ホヤ, カークウッド プレイス 6 1 0
(72)発明者 ロックトン, スティーブン
アメリカ合衆国, カリフォルニア州, サン ディエゴ, マウント タミ レーン 3 3 1 9
(72)発明者 クローナー, リサ, ジェイ.
アメリカ合衆国, カリフォルニア州, サン ディエゴ, ラムズデル コート 1 1 7 0 4
(72)発明者 フレッチャー, フレデリック, エー.
アメリカ合衆国, カリフォルニア州, サン ディエゴ, デール ストリート 1 4 2 9
(72)発明者 ストックフィッシュ, トーマス
アメリカ合衆国, カリフォルニア州, エスコンディード, キャニオン ロード 2 3 3 6
(72)発明者 シング, シャラット
アメリカ合衆国, カリフォルニア州, ランチョ サンタ フェ, トップ オブ ザ モーニング ウェイ 8 1 7 1

Fターム(参考) 4B024 AA11 CA01 CA09 CA20 HA08 HA11
4B063 QA01 QA13 QQ03 QQ28 QQ42 QQ52 QR32 QR35 QR55 QR62
QS25 QS34 QX02