

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5963198号  
(P5963198)

(45) 発行日 平成28年8月3日(2016.8.3)

(24) 登録日 平成28年7月8日(2016.7.8)

(51) Int.Cl.	F 1		
<b>G 0 6 F</b> 19/24 (2011.01)	G 0 6 F	19/24	
<b>C 1 2 Q</b> 1/68 (2006.01)	C 1 2 Q	1/68	A
<b>G O 1 N</b> 33/50 (2006.01)	G O 1 N	33/50	P
<b>G O 1 N</b> 33/68 (2006.01)	G O 1 N	33/68	
<b>C 1 2 N</b> 15/09 (2006.01)	C 1 2 N	15/00	A

請求項の数 11 (全 26 頁)

(21) 出願番号	特願2012-211921 (P2012-211921)	(73) 特許権者	503360115
(22) 出願日	平成24年9月26日 (2012.9.26)		国立研究開発法人科学技術振興機構
(65) 公開番号	特開2014-64515 (P2014-64515A)		埼玉県川口市本町四丁目1番8号
(43) 公開日	平成26年4月17日 (2014.4.17)	(74) 代理人	110000947
審査請求日	平成27年9月18日 (2015.9.18)		特許業務法人あーく特許事務所
特許法第30条第2項適用	SCIENTIFIC REPORTS   2 : 342   (平成24年3月29日) に掲載	(72) 発明者	陳 洛南
			兵庫県神戸市東灘区向洋町中7丁目2-1 5番館808号室
		(72) 発明者	合原 一幸
			千葉県習志野市谷津4-8-8-208
		(72) 発明者	劉 銳
			東京都世田谷区北沢1-13-17-102

最終頁に続く

(54) 【発明の名称】 動的ネットワークバイオマーカーの検出装置、検出方法及び検出プログラム

(57) 【特許請求の範囲】

【請求項1】

生体に関する測定により得られた複数の因子項目についての測定データに基づいて、測定対象である生体の症状の指標となるバイオマーカーの候補を検出する検出装置であって、

前記各因子項目のそれぞれの測定データの時系列変化の相関関係に基づいて複数の因子項目を複数のクラスターに分類する分類手段と、

分類した各クラスターから、前記各因子項目のそれぞれの測定データの時系列変化及び各因子項目間の測定データの時系列変化の相関関係に基づいて予め設定された選出条件に該当するクラスターを選出する選出手段と、

選出したクラスターに含まれる因子項目をバイオマーカーの候補として検出する検出手段と

を備えることを特徴とする検出装置。

【請求項2】

請求項1に記載の検出装置であって、

前記選出手段は、

クラスター内の各因子項目間の測定データの相関を示す値の平均値を第1指数として算出する手段と、

クラスター内の因子項目の測定データと当該クラスター外の因子項目の測定データとの間の相関を示す値の平均値を第2指数として算出する手段と、

クラスター内の各因子項目について測定データの標準偏差の平均値を第3指数として算出する手段と

を含み、

複数のクラスターのうちから、第1指数、第2指数及び第3指数に基づいて、バイオマーカーとすべき因子項目を含むクラスターを選出する

ことを特徴とする検出装置。

【請求項3】

請求項2に記載の検出装置であって、

前記選出手段は、前記第1指数と、前記第2指数と、前記第3指数の逆数との積に基づく総合指数が最大であるクラスターを選択する

ことを特徴とする検出装置。

10

【請求項4】

請求項2又は請求項3に記載の検出装置であって、

前記各因子項目のそれぞれの測定データが、有意性をもって経時的に変化しているか否かを検定する差次検定手段を更に備え、

前記分類手段は、経時的变化に有意性があると検定された因子項目について分類する

ことを特徴とする検出装置。

【請求項5】

請求項4に記載の検出装置であって、

前記差次検定手段は、各因子項目の測定データ、並びに因子項目及び時系列毎に予め設定されている参照データとの比較結果に基づいて、有意性に係る検定を行う

ことを特徴とする検出装置。

20

【請求項6】

請求項5に記載の検出装置であって、

各因子項目について、対応する参照データの標準偏差の平均値を示す参照標準偏差、及び因子項目間の相関を示す値の平均値を示す参照相関値を算出する手段を更に備え、

前記検出手段は、前記第1指数が前記参照標準偏差に比べて有意性をもって増大し、前記第2指数が前記参照相関値に比べて有意性をもって減少し、かつ、前記第3指数が前記参照標準偏差に比べて有意性をもって増大した場合に、当該クラスターに含まれる項目をバイオマーカーの候補として検出する

ことを特徴とする検出装置。

30

【請求項7】

請求項1乃至請求項6のいずれかに記載の検出装置であって、

前記検出手段は、

クラスターに含まれる複数の因子項目の有意性を、測定データの統計値に基づいて検定する手段を含み、

有意性がある場合に、クラスターに含まれる項目をバイオマーカーの候補として検出する

ことを特徴とする検出装置。

【請求項8】

請求項1乃至請求項7のいずれかに記載の検出装置であって、

前記複数の因子項目のいずれかは、遺伝子に関する測定項目、タンパク質に関する測定項目、代謝物に関する測定項目、又は生体から得られる画像に関する測定項目である

ことを特徴とする検出装置。

40

【請求項9】

生体に関する測定により得られた複数の因子項目についての測定データに基づいて、測定対象である生体の症状の指標となるバイオマーカーの候補を検出する検出装置を用いた検出方法であって、

前記検出装置は、

前記各因子項目のそれぞれの測定データの時系列変化の相関関係に基づいて複数の因子

50

項目を複数のクラスターに分類する分類ステップと、

分類した各クラスターから、前記各因子項目のそれぞれの測定データの時系列変化及び各因子項目間の測定データの時系列変化の相関関係に基づいて予め設定された選出条件に該当するクラスターを選出する選出ステップと、

選出したクラスターに含まれる因子項目をバイオマーカーの候補として検出する検出ステップと

を実行することを特徴とする検出方法。

【請求項 10】

生体に関する測定により得られた複数の因子項目についての測定データに基づいて、測定対象である生体の症状の指標となるバイオマーカーの候補を検出する検出方法であって

10

、異なる時間点で採取した複数の生体サンプルのそれぞれから測定したハイスループットデータの中から、差次的生体分子を算出する分子スクリーニングステップと、

相関の高い生体分子同士を1つのクラスターとするように、前記分子スクリーニングステップで選出した前記差次的生体分子を複数のクラスターに分類するクラスター化ステップと、

前記クラスター化ステップで得られた複数のクラスターの中から、クラスター内の生体分子の間の相関の増大、クラスター内の生体分子の標準偏差の増大、及びクラスター内の生体分子と他の生体分子との間の相関の低減が最も著しいクラスターを、前記バイオマーカーの候補として先取る候補選択ステップと、

20

前記候補選択ステップで選出したバイオマーカーの候補が前記バイオマーカーであるか否かを、有意性検定によって判定する判定ステップと

を実行することを特徴とする検出方法。

【請求項 11】

コンピュータに、生体に関する測定により得られた複数の因子項目についての測定データに基づいて、測定対象である生体の症状の指標となるバイオマーカーの候補を検出する処理を実行させる検出プログラムであって、

コンピュータに、

前記各因子項目のそれぞれの測定データの時系列変化の相関関係に基づいて複数の因子項目を複数のクラスターに分類する分類ステップと、

30

分類した各クラスターから、前記各因子項目のそれぞれの測定データの時系列変化及び各因子項目間の測定データの時系列変化の相関関係に基づいて予め設定された選出条件に該当するクラスターを選出する選出ステップと、

選出したクラスターに含まれる因子項目をバイオマーカーの候補として検出する検出ステップと

を実行させることを特徴とする検出プログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、生体に関する測定により得られた複数の因子項目の測定データに基づいて、前記生体の症状の指標となるバイオマーカーの候補を検出する検出装置、検出方法及び検出プログラムに関する。

40

【背景技術】

【0002】

様々な研究結果によると、多くの疾病、特に複雑な疾病の悪化進行のプロセスは、気候システム、生態システム、経済システム等のシステムと同様に、ある臨界閾値を超えた時点、いわゆる分岐点に達すると、状態遷移が突然発生し、健康状態から急激に疾病状態に変化する（例えば、非特許文献1～5参照）。このような複雑疾病の動的メカニズムに関する研究において、疾病悪化（例えば、喘息発作、癌の発病）の進行プロセスを、時間に依存した非線形動力学システムとしてモデル化し、モデル化したシステムを観測すること

50

によって、分岐点での相転移で病気が急激に悪化することが既に判明している（非特許文献 1、6 参照）。

【0003】

図 1 は、疾病の進行プロセスを概念的に示す説明図である。図 1 の a は、疾病の進行プロセスを模式的に示している。図 1 の b、c 及び d は、進行プロセスの過程において、前述のモデル化したシステムの安定性をポテンシャル関数として示し、横軸に経過を示す時刻をとり、縦軸にポテンシャル関数の値をとって概念的に示したグラフである。図 1 の a に示すように、疾病悪化の進行プロセスは、正常状態（健康状態）、疾病前状態、疾病状態として表すことができる。正常状態において、システムは安定して、図 1 の b に黒丸の位置として示すように、ポテンシャル関数の値が最小値になっている。疾病前状態において、システムは、図 1 の c に黒丸の位置として示すように、ポテンシャル関数の値が高くなっている。したがって、外乱の影響を受けやすい状態であり、小さな外乱を受けるだけで相転移してしまう分岐点の付近、即ち、正常状態の限界に位置している。但し、当該疾病前状態は、適切な処置によって、容易に正常状態に回復することができる。一方、疾病状態において、システムは安定して、図 1 の d に黒丸の位置として示すように、ポテンシャル関数の値が大局的的最小値になっている。そのため、正常状態から相転移で生じたこの疾病状態が、正常状態へ回復することは困難である。（図中では便宜上、回復不能と表記）

10

【0004】

したがって、もし疾病前状態を検知して、疾病状態に遷移する前に、疾病状態に遷移しつつあることを患者に予告することができれば、適切な措置を取ることによって、患者を疾病前状態から正常状態へ回復させることができる可能性が高い。

20

【0005】

即ち、分岐点（臨界閾値）を検出することができれば、臨界遷移の予測が可能であり、病気の早期診断を実現することができる。

【0006】

なお、従来から疾病状態を診断する方法として、バイオマーカーが用いられている。従来から用いられている通常のバイオマーカーは、生体から採取された血清、尿等の体液、更には、組織に含まれる分子レベルの DNA、RNA、蛋白質、代謝物等であり、生体内の生物学的変化を定量的に把握することができる指標である。従来のバイオマーカーによる疾病の診断方法は、ノーマルサンプル（健康状態で採取したサンプル）から抽出したバイオマーカーと異常サンプル（疾病状態で採取したサンプル）から抽出したバイオマーカーとを比較することにより、病気を診断する方法である。

30

【先行技術文献】

【非特許文献】

【0007】

【非特許文献 1】ヴェネガス・ジェイ・ジー（Venegas, J.G.）等著，「破壊変化に対する前奏曲のような喘息の自己組織（Self-organized patchiness in asthma as a prelude to catastrophic shifts.）」，（英国），ネイチャー（Nature），ネイチャーパブリッシンググループ（Nature Publishing Group），2005年，434巻，p. 777 - 782

40

【非特許文献 2】マックシャーリー・ピー・イー（McSharry, P.E.）、スミス・エル・エー（Smith, L.A.）、タラッセンコ・エル（Tarassenko, L.）著，「癇癇発作の予測：非線形方法が適切であるか。（Prediction of epileptic seizures: are nonlinear methods relevant.）」，（英国），ネイチャー・メディシン（Nature Medicine），ネイチャーパブリッシンググループ（Nature Publishing Group），2003年，9巻，p. 241 - 242

【非特許文献 3】ロベルト・ピー・ピー（Roberto, P.B.）、エリセオ・ジー（Eliseo, G.）、ジョセフ・エフ・シー（Josef, C.），「物流回帰の変化点推定のための転移モデル（Transition models for change-point estimation in logistic regression.）」，（米国），医学における統計（Statistics in Medicine），ワイリーブラックウェル（Wile

50

y-Blackwell ) , 2003年, 22巻, p. 1141 - 1162

【非特許文献4】ペック・エス (Paek, S.) 等著, 「音響神経腫に対するガンマナイフ手術後、聴覚保存 (Hearing preservation after gamma knife stereotactic radiosurgery of vestibular schwannoma.)」, (米国), キャンサー (Cancer), ワiley-ブラックウェル (Wiley-Blackwell) , 2005年, 1040巻, p. 580 - 590

【非特許文献5】リウ・ジェイ・ケイ (Liu, J.K.)、ロビット・アール・エル (Rovit, R. L.)、クッドウェル・ダブリュー・ティー (Couldwell, W.T.) , 「脳下垂体卒中 (Pituitary Apoplexy)」, (米国), セミナーズ・イン・ニューロサージェリー (Seminars in neurosurgery) , ティーム (Thieme) 2001年, 12巻, p. 315 - 320

【非特許文献6】タナカ・ジー (Tanaka, G.) , ツモト・ケイ (Tsumoto, K.) , ツジ・エス (Tsuji, S.)、アイハラ・ケイ (Aihara, K.) 著, 「前立腺癌に対する間歇的なホルモン療法のハイブリッドシステムモデルの分岐点分析。 (Bifurcation analysis on a hybrid systems model of intermittent hormonal therapy for prostate cancer.)」, (米国), フィジカル・レビュー (Physical Review) , アメリカ物理学会, 2008年, 237巻, p. 2616 - 2627

【発明の概要】

【発明が解決しようとする課題】

【0008】

しかしながら、複雑疾病の場合、臨界遷移の予測は極めて困難である。その理由は以下の通りである。

【0009】

第一に、疾病前状態は、正常状態の境界であり、分岐点に達する前に、著しい変化が生じ難い。そのため、従来のバイオマーカー、スナップショット測定等の手法による診断では、正常状態と疾病前状態とを区別することが難しい。

【0010】

第二に、様々な研究がなされている中で、分岐点を予測するための早期診断用の警告信号を精度高く検出することができる信頼性の高い疾病モデルがまだ開発されていない。特に、同じ疾病でも、個人によって、疾病悪化の進行プロセスが異なるため、モデルベースの診断方法は、成功する確率が低い。

【0011】

第三に、疾病前状態の検出対象は患者であり、通常、一人の患者から得られるサンプルの数が限られているため、長期間にわたって、予測に必要な十分なサンプルを採取することが困難である。

【0012】

しかも、従来のバイオマーカーによる疾病の診断方法は、正常状態と疾病状態とを比較することによって診断を行う方法であり、診断された時点では、患者はすでに疾病状態に陥り、以前の正常状態に戻ることは難しい。

【0013】

それに対して、本発明は、疾病状態へ遷移する前の疾病前状態を検出することができる検出装置、検出方法及び検出プログラムを提供し、さらに、疾病モデルが不要で、数少ない生体サンプルだけで診断を補助することが可能な検出装置等を提供することを目的としている。

【課題を解決するための手段】

【0014】

上記目的を達成するために、本発明に係る検出装置は、生体に関する測定により得られた複数の因子項目についての測定データに基づいて、測定対象である生体の症状の指標となるバイオマーカーの候補を検出する検出装置であって、前記各因子項目のそれぞれの測定データの時系列変化の相関関係に基づいて複数の因子項目を複数のクラスターに分類する分類手段と、分類した各クラスターから、前記各因子項目のそれぞれの測定データの時系列変化及び各因子項目間の測定データの時系列変化の相関関係に基づいて予め設定され

10

20

30

40

50

た選出条件に該当するクラスターを選出する選出手段と、選出したクラスターに含まれる因子項目をバイオマーカーの候補として検出する検出手段とを備えることを特徴とする。

【0015】

上記特徴を有する検出装置は、正常状態から疾病状態へ遷移する前の疾病前状態を示す警告信号となるバイオマーカーの候補を検出することができる。また、バイオマーカーを特定することができれば、検出対象から数少ないサンプルを採取するだけで疾病前状態を判別することができる。

【0016】

また、本発明に係る検出装置において、前記選出手段は、クラスター内の各因子項目間の測定データの相関を示す値の平均値を第1指数として算出する手段と、クラスター内の因子項目の測定データと当該クラスター外の因子項目の測定データとの間の相関を示す値の平均値を第2指数として算出する手段と、クラスター内の各因子項目について測定データの標準偏差の平均値を第3指数として算出する手段とを含み、複数のクラスターのうちから、第1指数、第2指数及び第3指数に基づいて、バイオマーカーとすべき因子項目を含むクラスターを選出することを特徴とする。

【0017】

したがって、前記第1指数、第2指数、第3指数を用いることにより、各クラスターの特性を定量的評価することができ、バイオマーカーを容易に選出することができる。

【0018】

また、本発明に係る検出装置において、前記選出手段は、前記第1指数と、前記第2指数と、前記第3指数の逆数との積に基づく総合指数が最大であるクラスターを選択することを特徴とする。

【0019】

したがって、総合指数に基づいてクラスターを選択することにより、バイオマーカーの候補となる因子項目の信頼性を高めることができる。

【0020】

また、本発明に係る検出装置は、前記各因子項目のそれぞれの測定データが、有意性をもって経時的に変化しているか否かを検定する差次検定手段を更に備え、前記分類手段は、経時変化に有意性があると検定された因子項目について分類することを特徴とする。

【0021】

したがって、経時的に著しい変化を示した因子項目を選出することにより、膨大な測定データから、バイオマーカーの候補を効率的に検出することができる。

【0022】

また、本発明に係る検出装置は、前記差次検定手段は、各因子項目の測定データ、並びに因子項目及び時系列毎に予め設定されている参照データとの比較結果に基づいて、有意性に係る検定を行うことを特徴とする。

【0023】

したがって、検出対象となる前記複数の因子項目に係る測定データの他に、参照基準となる生体サンプルを参照データとして取得することによって、測定データと参照データとを比較し、外乱を排除して検出を行うことができる。

【0024】

また、本発明に係る検出装置は、各因子項目について、対応する参照データの標準偏差の平均値を示す参照標準偏差、及び因子項目間の相関を示す値の平均値を示す参照相関値を算出する手段を更に備え、前記検出手段は、前記第1指数が前記参照標準偏差に比べて有意性をもって増大し、前記第2指数が前記参照相関値に比べて有意性をもって減少し、かつ、前記第3指数が前記参照標準偏差に比べて有意性をもって増大した場合に、当該クラスターに含まれる項目をバイオマーカーの候補として検出することを特徴とする。

【0025】

したがって、選出された因子項目がバイオマーカーとして適切か否かを判定することができる。

10

20

30

40

50

## 【0026】

また、本発明に係る検出装置において、前記検出手段は、クラスターに含まれる複数の因子項目の有意性を、測定データの統計値に基づいて検定する手段を含み、有意性がある場合に、クラスターに含まれる項目をバイオマーカーの候補として検出することを特徴とする。

## 【0027】

したがって、検定により誤検出の発生を極力防止することができる。

## 【0028】

また、本発明に係る検出装置は、前記複数の因子項目のいずれかは、遺伝子に関する測定項目、タンパク質に関する測定項目、代謝物に関する測定項目、又は生体から得られる画像に関する測定項目であることを特徴とする。

10

## 【0029】

したがって、遺伝子、タンパク質、代謝物に関して測定項目を因子項目とすることにより、生体内の生物変化を定量的に把握して、検出結果の信頼性を向上させることができる。

## 【0030】

さらに、本発明に係る検出方法は、生体に関する測定により得られた複数の因子項目についての測定データに基づいて、測定対象である生体の症状の指標となるバイオマーカーの候補を検出する検出装置を用いた検出方法であって、前記検出装置は、前記各因子項目のそれぞれの測定データの時系列変化の相関関係に基づいて複数の因子項目を複数のクラスタに分類する分類ステップと、分類した各クラスターから、前記各因子項目のそれぞれの測定データの時系列変化及び各因子項目間の測定データの時系列変化の相関関係に基づいて予め設定された選出条件に該当するクラスターを選出する選出ステップと、選出したクラスターに含まれる因子項目をバイオマーカーの候補として検出する検出ステップとを実行することを特徴とする。

20

## 【0031】

また、本発明に係る検出方法は、生体に関する測定により得られた複数の因子項目についての測定データに基づいて、測定対象である生体の症状の指標となるバイオマーカーの候補を検出する検出方法であって、異なる時間点で採取した複数の生体サンプルのそれぞれから測定したハイスループットデータの中から、差次的生体分子を算出する分子スクリーニングステップと、相関の高い生体分子同士を1つのクラスターとするように、前記分子スクリーニングステップで選出した前記差次的生体分子を複数のクラスターに分類するクラスター化ステップと、前記クラスター化ステップで得られた複数のクラスターの中から、クラスター内の生体分子の間の相関の増大、クラスター内の生体分子の標準偏差の増大、及びクラスター内の生体分子と他の生体分子との間の相関の低減が最も著しいクラスターを、前記バイオマーカーの候補として先取する候補選択ステップと、前記候補選択ステップで選出したバイオマーカーの候補が前記バイオマーカーであるか否かを、有意性検定によって判定する判定ステップとを実行することを特徴とする。

30

## 【0032】

上記特徴を有する検出方法は、正常状態から疾病状態へ遷移する前の疾病前状態を示す警告信号となるバイオマーカーの候補を検出することができる。また、バイオマーカーを特定することができれば、検出対象から数少ないサンプルを採取するだけで疾病前状態を判別することができる。

40

## 【0033】

さらに、本発明に係る検出プログラムは、コンピュータに、生体に関する測定により得られた複数の因子項目についての測定データに基づいて、測定対象である生体の症状の指標となるバイオマーカーの候補を検出する処理を実行させる検出プログラムであって、コンピュータに、前記各因子項目のそれぞれの測定データの時系列変化の相関関係に基づいて複数の因子項目を複数のクラスターに分類する分類ステップと、分類した各クラスターから、前記各因子項目のそれぞれの測定データの時系列変化及び各因子項目間の測定デー

50

タの時系列変化の相関関係に基づいて予め設定された選出条件に該当するクラスターを選出する選出ステップと、選出したクラスターに含まれる因子項目をバイオマーカーの候補として検出する検出ステップとを実行させることを特徴とする。

【0034】

上記特徴を有する検出プログラムは、コンピュータにて実行することにより、コンピュータが、本発明に係る検出装置として動作する。したがって、正常状態から疾病状態へ遷移する前の疾病前状態を示す警告信号となるバイオマーカーの候補を検出することができる。また、バイオマーカーを特定することができれば、検出対象から数少ないサンプルを採取するだけで疾病前状態を判別することができる。

【発明の効果】

10

【0035】

本発明によれば、生体サンプルを診断対象から採取し、採取した生体サンプルにおいて、疾病状態になる直前の疾病前状態を示す警告信号となるバイオマーカーが存在するか否かを調べることにより、診断対象が疾病前状態にあるか否かを診断することができる。そのため、疾病悪化のモデルを構築する必要がなく、疾病悪化の駆動要素を特定する必要もなく、疾病になる前の段階で、疾病の早期診断を実現することができる。

【図面の簡単な説明】

【0036】

【図1】図1は、疾病の進行プロセスを説明する模式図である。

【図2】図2は、本発明の検出方法に係るDNBの動的特性を例示する模式図である。

20

【図3】図3は、実施の形態におけるDNBの検出方法の一例を示すフローチャートである。

【図4】図4は、実施の形態における差次的生体分子の選出処理の一例を示すフローチャートである。

【図5】図5は、実施の形態におけるDNBの候補の選出処理の一例を示すフローチャートである。

【図6】図6は、実施の形態におけるDNBの判定処理の一例を示すフローチャートである。

【図7】図7は、実施の形態におけるDNBによる疾病の早期診断方法による診断スケジュールの一例を示す図である。

30

【図8】図8は、実施の形態におけるDNBによる疾病の早期診断方法の一例を示すフローチャートである。

【図9】図9は、総合指数Iに比例した疾病リスクを表示する図形の一例である。

【図10】図10は、総合指数Iに比例した疾病リスクを表示する図形の一例である。

【図11】図11は、本発明に係る検出装置の構成例を示すブロック図である。

【図12】図12は、本発明に係る検出装置によるDNBの検出処理の一例を示すフローチャートである。

【図13】図13は、第1の検証例における診断用のデータを示すテーブルである。

【図14A】図14Aは、第1の検証例において、検出されたDNBの候補の標準偏差の平均値の時系列変化の例を示すグラフである。

40

【図14B】図14Bは、第1の検証例において、検出されたDNBの候補のクラスターのメンバー間のピアソン相関係数の絶対値の平均値の時系列変化の例を示すグラフである。

【図14C】図14Cは、第1の検証例において、検出されたDNBの候補のクラスターのメンバーと他の遺伝子との間のピアソン相関係数の絶対値の平均値の時系列変化の例を示すグラフである。

【図14D】図14Dは、第1の検証例において、検出されたDNBの候補の総合指数の平均値の時系列変化の例を示すグラフである。

【図15】図15は、第1の検証例において、ケースグループの遺伝子によって構成されたネットワークにおけるDNBの動的特性の一例を経時的に示すマップである。

50

【図16】図16は、第2の検証例における診断データのテーブルである。

【図17A】図17Aは、第2の検証例において、検出されたDNBの候補の標準偏差の平均値の時系列変化の例を示すグラフである。

【図17B】図17Bは、第2の検証例において、検出されたDNBの候補のクラスターのメンバー間のピアソン相関係数の絶対値の平均値の時系列変化の例を示すグラフである。

【図17C】図17Cは、第2の検証例において、検出されたDNBの候補のクラスターのメンバーと他の遺伝子との間のピアソン相関係数の絶対値の平均値の時系列変化の例を示すグラフである。

【図17D】図17Dは、第2の検証例において、検出されたDNBの候補の総合指数の平均値の時系列変化の例を示すグラフである。

10

【発明を実施するための形態】

【0037】

本発明の発明者らは、1つのサンプルから数千個の情報、即ち高次元データを得ることが可能なゲノムハイスループットテクノロジーを利用して、分岐プロセス理論に基づいて、複雑疾病の時間発展の数理モデルを構築し、分子ネットワークレベルにおける疾病悪化の進行メカニズムの研究を行った。その結果、疾病前状態において、臨界遷移発生前の直前分岐（突然悪化）状態を検出することができる動的ネットワークバイオマーカー（DNB: Dynamical Network Biomarker）の存在を解明した。当該動的ネットワークバイオマーカーを疾病前状態の警告信号として利用すれば、疾病モデルが不要で、数少ないサンプルだけで複雑疾病の早期診断を実現することができる。以下に、動的ネットワークバイオマーカーに基づく本発明を実施するための形態について説明する。

20

【0038】

<理論根拠>

先ず、本発明の基となる理論根拠について説明する。通常、疾病の進行プロセスに係るシステム（以下、システム（1）という）を次の式（1）で表すことができる。

【0039】

$$Z(k+1) = f(Z(k); P) \quad \dots \text{式(1)}$$

【0040】

ここで、 $Z(k) = (z_1(k), \dots, z_n(k))$  は、時刻  $k$  ( $k = 0, 1, \dots$ ) で観測されたシステム（1）の動的状態を表す変数であり、遺伝子発現量、タンパク質発現量、代謝物発現量等の情報とすることができる。詳細には、遺伝子、タンパク質等の分子に係る濃度、個数等の情報である。Pは、システム（1）の状態遷移を駆動する緩やかに変化するパラメータであり、例えば、SBP, CNV等の遺伝的要因、メチル化、アセチル化等の非遺伝的要因等の情報とすることができる。f = (f<sub>1</sub>, ..., f<sub>n</sub>) はZ(k)の非線形関数である。

30

【0041】

正常状態、疾病状態はそれぞれ、状態方程式  $Z(k+1) = f(Z(k); P)$  のアトラクタで表すことができる。複雑疾病の進行プロセスは、非常に複雑な動的特性を有するため、関数fは、数千個の変数を有する非線形関数である。しかも、システム（1）を駆動する要素Pは特定しにくいものである。そのため、正常状態及び疾病状態のシステムモデルを構築し解析することは大変困難である。

40

【0042】

それに対して、本発明の発明者らが着目したのは、正常状態から疾病状態へ遷移する直前のシステムの臨界遷移状態、即ち疾病前状態である。一般的に、システム（1）には、次の特性を有する平衡点が存在する。

【0043】

1.  $Z^*$  は、システム（1）の固定点とすれば、 $Z^* = f(Z^*; P)$ 、  
 2.  $P_c$  をシステムが分岐する閾値とすれば、 $P = P_c$  の時、ヤコビ行列  $f'(Z; P_c) / Z | Z = Z^*$  の1つの実数固有値又は1対の複素共役の固有値の絶対値が1にな

50

る、

3.  $P = P_c$  の時、一般的に、システム (1) の固有値の絶対値が 1 ではない。

【0044】

発明者らは、上記特性から、システム (1) が臨界遷移状態になると、次のような特異な特性が現れることを理論的に解明した。即ち、システム (1) が臨界遷移状態になると、システム (1) の各変数  $z_1, \dots, z_n$  のそれぞれをノードとして構成したネットワーク (1) において、一部のノードからなる支配グループ (サブネットワーク) が現れる。臨界遷移状態に現れる支配グループは、理想的には、次のような得意な特性を有する。

【0045】

(I)  $z_i$  と  $z_j$  は支配グループに属するノードである場合、

$PCC(z_i, z_j) \neq 1$  ;

$SD(z_i)$  ;

$SD(z_j)$  .

(II)  $z_i$  は支配グループに属するノードであるが、 $z_j$  は支配グループに属するノードではない場合、

$PCC(z_i, z_j) = 0$  ;

$SD(z_i)$  ;

$SD(z_j)$  境界値 .

(III)  $z_i$  と  $z_j$  は支配グループに属するノードではない場合

$PCC(z_i, z_j) \in (-1, 1) \setminus \{0\}$  ;

$SD(z_i)$  境界値 ;

$SD(z_j)$  境界値 .

【0046】

ここで、 $PCC(z_i, z_j)$  は  $z_i$  と  $z_j$  との間のピアソン相関係数であり、 $SD(z_i)$ 、 $SD(z_j)$  は  $z_i$  と  $z_j$  との標準偏差である。

【0047】

即ち、ネットワーク (1) において、上記特異な特性 (I) ~ (III) を有する支配グループの出現は、システム (1) が臨界遷移状態 (疾病前状態) に遷移していることの兆候として捉えることができる。したがって、当該支配グループを検出することにより、システム (1) の臨界遷移を検出することができる。即ち、当該支配グループを臨界遷移、即ち、疾病悪化直前の疾病前状態を示す警告信号とすることができる。そうすれば、システム (1) がいくら複雑であっても、駆動要素が不明であっても、警告信号となる支配グループだけを検出すれば、システム (1) の数理モデルを直接扱うことなく、疾病前状態であることを特定することができる。疾病前状態であることを特定することにより、疾病に対する事前対策及び早期治療を実現することが可能となる。

【0048】

本発明において、疾病前状態を示す警告信号となる当該支配グループを動的ネットワークバイオマーカー (以下、DNBと略称する) と称する。

【0049】

< DNBの特性及び判定条件 >

上述したように、DNBは、特異な特性 (I) ~ (III) を有する支配グループであって、複数のノードからなるサブネットワークとして、システム (1) が疾病前状態になるときに、ネットワーク (1) に現れるものである。ネットワーク (1) において、各ノード ( $z_1, \dots, z_n$ ) を、遺伝子、タンパク質、代謝物等の生体分子について測定の対象となる因子項目とすれば、DNBは、上記特異な特性 (I) ~ (III) を満たした一部の生体分子に係る因子項目からなるグループ (サブネットワーク) である。

【0050】

さらに、上記特異な特性 (I) ~ (III) に基づいて、DNBを判定する条件を以下の通り定めることができる。

【0051】

10

20

30

40

50

・条件 ( I ) : ネットワーク ( 1 ) に存在する一部の生体分子 ( 遺伝子、タンパク質又は代謝物 ) からなるグループであって、当該グループ内の生体分子の間のピアソン相関係数  $PC C$  の絶対値の平均値が著しく増大する。

・条件 ( I I ) : 当該グループ内の生体分子と他の生体分子との間のピアソン相関係数  $OP C C$  の絶対値の平均値が著しく低減する。

・条件 ( I I I ) : 当該グループ内の生体分子の標準偏差  $S D$  の平均値が著しく増加する。

#### 【 0 0 5 2 】

上記条件 ( I ) ~ ( I I I ) を同時に満たした生体分子からなるグループを  $D N B$  と認定する。

10

#### 【 0 0 5 3 】

$D N B$  の特性を直観的に説明するために、次に、6つのノードを有するネットワークを例として、当該ネットワークにおける  $D N B$  の動的特性を説明する。図2は、本発明の検出方法に係る  $D N B$  の動的特性を例示する模式図である。図2のaは、疾病の進行プロセスとして、正常状態、疾病前状態及び疾病状態を示している。図2のb、c及びdは、進行プロセスの過程につき、正常状態、疾病前状態及び疾病状態において、モデル化したシステム ( 1 ) の安定性を、ポテンシャル関数として示し、横軸に経過を示す時刻をとり、縦軸にポテンシャル関数の値をとって概念的に示したグラフである。また、図2のe、f及びgは、正常状態、疾病前状態及び疾病状態のそれぞれに対応したシステム ( 1 ) のネットワーク状態の一例を概念的に示している。さらに、図2のhは、疾病前状態における  $D N B$  の因子項目となる分子濃度の時間変化の一例を示している。

20

#### 【 0 0 5 4 】

ノード  $z 1 \sim z 6$  は、遺伝子、タンパク質、代謝物等の異なる種類の生体分子に係る因子項目を表すものであり、ノード  $z 1 \sim z 6$  の間の接続ラインは、ノード間の相関を示し、ラインの太さは、ピアソン相関係数  $PC C$  の大きさを示し、 $z 1 \sim z 6$  を囲む丸印の描画方法は、各ノードの標準偏差  $S D$  の大きさを示している。接続ラインの描画方法として、丸印内が空白の場合に標準偏差  $S D$  は最も小さく、一方向の斜線、二方向の斜線となるにつれて標準偏差  $S D$  が大きいことを示している。

#### 【 0 0 5 5 】

正常状態において、図2のeに示すように、各ノードは、相互の相関及びそれぞれの標準偏差がいずれも均等で、中レベルになっている。しかしながら、疾病前状態になると、他のノードに比べて、著しい特異な特性を示すグループ (  $z 1 \sim z 3$  ) が現れる。当該グループ内のノード  $z 1 \sim z 3$  は、図2のfに示すように、相互間のピアソン相関係数が著しく増大し、他のノード  $z 4 \sim z 6$  との間のピアソン相関係数が著しく低下している。また、当該グループ内のノード  $z 1 \sim z 3$  の標準偏差は大きくなっている。その理由は、当該グループ内のノード  $z 1 \sim z 3$  は、図2のhに示すように、異なる時刻 (  $t = 1$ 、 $t = 2$ 、 $t = 3$  ) における濃度の変化が激しいからである。

30

#### 【 0 0 5 6 】

しかしながら、疾病状態へ遷移した後は、図2のgに示すように、当該グループ内の各ノード  $z 1 \sim z 3$  の標準偏差はすこし大きくなっているが、各ノード間のピアソン相関係数は均等に中レベルに戻っている。即ち、当該グループ (  $z 1 \sim z 3$  ) の上記の特異な特性が消失している。

40

#### 【 0 0 5 7 】

図2に概念的に示すように、疾病前状態において、得意な特性を有する一部のノードを含む支配グループが出現する。本願において、 $D M B$  と称するこのような支配グループは、疾病前状態にあることを示し、疾病状態への遷移を予告することができる警告信号として、疾病の早期診断のためのバイオマーカーとして使用することができる。また、 $D N B$  は、通常の疾病診断に使われている静的バイオマーカーと異なり、特性が変化しているネットワークにおいて現れるサブネットワークである。そこで、本願においては、当該支配グループを  $D N B$  ( 動的ネットワークバイオマーカー ) と呼ぶこととした。

50

## 【 0 0 5 8 】

## &lt; 警告信号 &gt;

上述したように、DNBは、疾病前状態を示す警告信号として、疾病の早期診断に使用することができる。当該警告信号の強度を測るものとして、上述したDNB内のノード間のピアソン相関係数PCCの絶対値の平均値、DNB内のノードと他のノードとの間のピアソン相関係数OPCCの絶対値の平均値、及びDNBの標準偏差SDを用いることができる。さらに、DNBの特性を総合的に反映することができる総合指数Iを導入することができる。本発明において、一例として、次の式(2)で表す総合指数Iを導入する。

## 【 0 0 5 9 】

$$I = SDd \times PCCd / OPCCd \quad \dots \text{式(2)}$$

10

## 【 0 0 6 0 】

式(2)において、PCCdはDNB内のノード間のピアソン相関係数の絶対値の平均値を示し、OPCCdはDNB内のノードと他のノードとのピアソン相関係数の絶対値の平均値を示し、SDdはDNB内のノードの標準偏差の平均値を示す。当該式(2)から分かるように、SDd及びPCCdが増大し、かつ、OPCCdが減少すると、当該総合指数Iは大幅に増大する。そのため、DNBの特性を高い感度で検出することができる。さらに、当該総合指数Iの値から、疾病状態との距離をある程度把握することができる。

## 【 0 0 6 1 】

## &lt; DNBの検出方法 &gt;

図3は、実施の形態におけるDNBの検出方法の一例を示すフローチャートである。本発明の検出方法においては、まず、生体に関する測定により測定データを得ることが必要である。DNAチップなどのハイスループット技術を利用すれば、1つの生体サンプルから2万個以上の遺伝子を測定することができる。統計的に分析するために、本発明において、測定対象から複数(6個以上)の生体サンプルを異なる時間点で採取し、採取した生体サンプルを測定して得られた測定データを集計して統計データを取得する。本発明におけるDNBの検出方法は、主に、図3に示すように、ハイスループットデータの取得処理(s1)と、差次的生体分子の選出処理(s2)と、クラスター化処理(s3)と、DNBの候補の選出処理(s4)と、有意性分析によるDNBの判定処理(s5)とを含む。次に、これらの各処理について詳細に説明する。

20

## 【 0 0 6 2 】

ステップs1のハイスループットデータの取得処理は、検出対象となるサンプルをケースサンプルとし、参照用サンプルをコントロールサンプルとし、各サンプルから、それぞれのハイスループット生理学的データ、即ち、生体分子の発現量の測定データ(例えば、マイクロアレイデータ)を取得する処理である。参照用サンプルとは、検査の対象となる患者から事前に採取したサンプル、採取時に最初に採取したサンプル等のサンプルであり、測定装置の校正等の目的を有するコントロールサンプルとして用いられる。コントロールサンプルは、必ずしも必要ではないが、誤差要因を排除し、測定の信頼性を向上させるために有用である。

30

## 【 0 0 6 3 】

ステップs2の差次的生体分子の選出処理は、発現量に著しい変化を示した生体分子を選択する処理である。図4は、実施の形態における差次的生体分子の選出処理の一例を示すフローチャートである。図4は、図3に示したステップs2の差次的生体分子の生体処理を詳細に示したものである。

40

## 【 0 0 6 4 】

図4に示すように、まず、n個のケースサンプルからそれぞれ測定したハイスループットデータ(生体分子の発現量)に基づく統計データをD1cとし、コントロールサンプルから測定したデータをDrとする(s21)。次に、各ケースサンプル生体分子D1cに対して、スチューデントのt検定を行い、コントロールサンプルのハイスループットデータDrに比べて発現量に著しい変化を示した生体分子D2cを選出する(s22)。ステップs22では、発現量に著しい変化を示した生体分子D2cを選出する方法として、ス

50

チューデントの  $t$  検定を例示しているが、特に手法を限定するものではない。例えば、マン・ホイットニーの  $U$  検定等の他の検定方法を適用することも可能であり、このようなノンパラメトリック手法による検定は、母集団となる  $D1c$  が正規分布に従わない場合に、特に有効である。また、スチューデントの  $t$  検定をする場合においても、有意水準の値は、 $0.05$ 、 $0.01$  等の値を適宜設定することができる。

#### 【0065】

次に、誤発現率  $FDR$  を用いて、各ケースサンプル生体分子  $D2c$  に対して、複数の比較 (multiple comparisons) 又は複数のスチューデントの  $t$  検定を補正し、補正後の各ケースサンプル遺伝子又はタンパクデータ  $D3c$  を選出する ( $s23$ )。次に、2倍変更方法 (The two-fold change method) を用いて、補正後の各ケースサンプル遺伝子又はタンパクデータ  $D3c$  から標準偏差  $SD$  が比較的著しく変化する  $Dc$  を差次的生体分子として選出する。ここで選出された差次的生体分子  $Dc$  は、コントロールサンプルの生体分子  $Dr$  に比べて、著しい差異を示すだけでなく、自分の平均値からも大きく逸脱している。ステップ  $s23$  においても、検定方法はスチューデントの  $t$  検定に限るものではない。

#### 【0066】

次に、クラスター化処理 (図3の  $s3$ ) を行う。ここでいうクラスター化処理とは、複数の生体分子を、互いに相関の高いグループにそれぞれ分類する処理であり、生体分子が分類されたそれぞれのグループをクラスターと称する。即ち、互いに相関の高い生体分子を1つのクラスターにするように、図4に示すステップ  $s24$  で選択した差次的生体分子  $Dc$  を、 $n$  個のクラスターに分類する。得られた全てのクラスターは、潜在的な支配グループ、即ち、検出すべき  $DNB$  の候補となる。

#### 【0067】

次に、図3に示す  $DNB$  の候補の選出処理 ( $s4$ ) を行う。図5は、実施の形態における  $DNB$  の候補の選出処理の一例を示すフローチャートである。図5は、図3に示したステップ  $s4$  の  $DNB$  の候補の選出処理を詳細に示したものである。即ち、 $DNB$  の候補の選出処理は、図5に示す  $DNB$  の候補の選出処理のフローチャートに基づいて処理を行う。図5に示す循環ループで、すべてのクラスター ( $k$ ) ( $k = 1, \dots, n$ ) について、クラスター内のノードの間のピアソン相関係数の絶対値の平均値  $PCCd(k)$ 、クラスター内のノードと他のノードとの間のピアソン相関係数の絶対値の平均値  $OPCCd(k)$ 、クラスター内のノードの標準偏差の平均値  $SDd(k)$ 、及び総合指数  $I(k)$  を計算する ( $s41 \sim s46$ )。そして、すべてのクラスターから、総合指数  $I$  の値が一番大きいクラスターを  $DNB$  の候補として選出する ( $s47$ )。

#### 【0068】

次に、図3に示す有意性分析による  $DNB$  の判定処理 (図3の  $s5$ ) を行う。図6は、実施の形態における  $DNB$  の判定処理の一例を示すフローチャートである。図6は、図3に示したステップ  $s5$  の優位性分析による  $DNB$  の判定処理を詳細に示したものである。即ち、上述した  $DNB$  の判定条件 ( $I$ )  $\sim$  ( $III$ ) に基づいて、ステップ  $s47$  で  $DNB$  の候補として選出されたクラスター ( $m$ ) は  $DNB$  であるか否かを判定する。様々な有意性分析によって判定することができるが、一例として、図6に示す  $DNB$  の判定処理のフローチャートに基づいて処理を行う。

#### 【0069】

図6に示すように、まず、コントロールサンプルから取得したデータの各ノード間のピアソン相関係数の絶対値の平均値  $PCCDr$ 、及び各ノードの標準偏差の平均値  $SDdr$  をそれぞれ計算する ( $s51$ 、 $s52$ )。そして、コントロールサンプルのピアソン相関係数の平均値  $PCCDr$  に比べて、ステップ  $s47$  で選出したクラスター ( $m$ ) 内の各ノード間のピアソン相関係数の絶対値の平均値  $PCCd(m)$  が有意に増大したか否かを判断する ( $s53$ )。有意に増大していないと判断した場合 ( $NO$ )、 $DNB$  が存在しないとの結果を出力し、処理を終了する ( $s57$ )。他方、有意に増大したと判断した場合 ( $YES$ )、次のステップ  $s54$  に進む。ステップ  $s54$  において、コントロールサンプルのピアソン相関係数の平均値  $PCCDr$  に比べて、クラスター ( $m$ ) 内の各ノードと他の

10

20

30

40

50

ノードとのピアソン相関係数の平均値  $OPCCd(m)$  が有意に低減したか否かを判断する (s 5 4)。有意に低減していないと判断した場合 (NO)、DNB が存在しないとの結果を出力し (s 5 7)、処理を終了する。他方、有意に低減したと判断した場合 (YES)、次のステップ s 5 5 に進む。ステップ s 5 5 において、コントロールサンプルの標準偏差平均値  $SDr$  に比べて、クラスター (m) 内のノードの標準偏差平均値  $SDd(m)$  が有意に増大したか否かを判断する。有意に増大していないと判断した場合 (NO)、DNB が存在しないと判断して (s 5 7)、処理を終了させる。他方、有意に増大したと判断した場合、クラスター (m) を DNB と認定して (s 5 6)、処理を終了させる。

#### 【 0 0 7 0 】

< DNB による疾病の早期診断方法 >

診断スケジュールとしては、一定な間隔を開けて、複数回診断を行い、毎回の診断で数個のサンプルを取ることが望ましい。図 7 は、実施の形態における DNB を用いた疾病の早期診断の診断スケジュールの一例を示す図である。図 7 に示すように、複数の段階 (段階 1、段階 2、...、段階 T) で、サンプルを採取する。各段階で採取するサンプルの数は、データの精度を確保するために、通常、1つの段階において、6 個以上のサンプルを採取することが望ましい。2つの連続した段階の間隔は、疾病の状況によって、数日、数週間、数カ月、又は数年と長く設定することができるが、各段階では、短い期間において異なる時間点でサンプルを採取することが望ましい。例えば、6 個のサンプルを 1 日のうちの 6 つの時間点で採取する。各時間点の間隔は、状況に応じて、例えば、数分間、数時間とすることができる。

#### 【 0 0 7 1 】

図 8 は、実施の形態における DNB による疾病の早期診断方法の一例を示すフローチャートである。図 8 に示すように、DNB による疾病の早期診断方法は、主に、サンプル採取処理 (s 1 0 0) と、差次的生体分子の選出 (s 2 0 0) と、DNB の候補の選出処理 (s 3 0 0) と、有意性分析による DNB の判定処理 (s 4 0 0) と、診断結果の出力処理 (s 5 0 0) とを含む。次に、これらの処理内容について詳細に説明する。

#### 【 0 0 7 2 】

サンプル採取処理 (s 1 0 0) : 一般の疾病診断方法と同じように、診断する疾病に応じて、必要な生理データを取るためのサンプルを採取する。例えば、肝臓障害の場合、血液、肝臓の組織などのサンプルを採取する。

#### 【 0 0 7 3 】

また、診断する際に診断対象から採取したサンプルをケースサンプルとする他に、参照用のサンプルとして、診断対象以外の健康な人から採取したサンプル、又は診断対象から最初に採取したサンプルをコントロールサンプルとすることができる。

#### 【 0 0 7 4 】

差次的生体分子の選出 (s 2 0 0) : 図 4 に示した差次的生体分子の選出処理フローチャートにしたがって、上記ステップ s 1 0 0 で採取したサンプルから、差次的生体分子を選出する。

#### 【 0 0 7 5 】

DNB の候補の選出処理 (s 3 0 0) : 図 5 に示した DNB の候補の選出フローチャートにしたがって、上記ステップ s 2 0 0 で選出した差次的生体分子から、DNB の候補となる支配グループを選出する。

#### 【 0 0 7 6 】

有意性分析による DNB の判定処理 (s 4 0 0) : 図 6 に示した有意性分析による DNB の判定方法を示すフローチャートにしたがって、上記ステップ s 3 0 0 で選出した DNB の候補が DNB であるか否かを判定する。

#### 【 0 0 7 7 】

診断結果の出力 (s 5 0 0) : 上記ステップ s 4 0 0 において、DNB が存在しないと判断した場合、上記ステップ s 3 0 0 で選出した DNB の候補のデータを次回の診断の参照データとして記憶装置に保存し、異常が認めない旨の診断結果を出力する。一方、上記

10

20

30

40

50

ステップs 400において、DNBと認定されたクラスターが存在すると判断した場合、認定されたクラスターの生体分子データをDNBのメンバーとして記憶し、疾病前状態にある旨の診断結果を出力する。また、検出したDNBに関連づけた診断結果を出力することもできる。なお、ここでは、診断結果として示しているが、詳細には、例えば、医師による診断の参考となる結果であり、ステップs 500にて出力される診断結果は、医師による診断そのものではなく、医師の診断の補助となり、診断を支援すべく診断の参考となる出力データである。

【0078】

例えば、DNBの特性を総合的に反映する総合指数Iを出力することができる。当該総合指数Iが大きければ大きいほど、分岐点に近いので、総合指数Iに比例した疾病リスクを直観的に見ることができ、図形又は画像の形で出力すると、ワーニング効果が高くなる。

10

【0079】

図9及び図10は、総合指数Iに比例した疾病リスクを表示する図形の一例である。図9において、矢印全体は疾病前状態（発病前期）を示し、矢印により示される方向の流れが疾病状態（発病）の経時変化を指しており、矢印内の左側に位置するひし形マークは診断で得られた総合指数Iの値に応じて位置が変化する発病リスクポインターである。総合指数Iの値が大きければ大きいほどひし形マークが矢印の右側に近づくように設定する。

【0080】

また、以前にもDNBを用いて疾病の早期診断を受けたことがある場合、前回診断で得られた総合指数と一緒に、図10のように、総合指数Iに比例した疾病リスクを表示することができる。図10において、点線のひし形マークが2011年7月1日の診断で得られた総合指数を示し、実線のひし形マークが2011年9月1日の診断で得られた総合指数を示している。ひし形マークの位置変動から、疾病状態に近づいていくことを直観的に判断することができる。

20

【0081】

また、DNBに関連づけた情報として、検出したDNBを含んだ全ネットワーク、又は当該DNBを含んだ一部のネットワークのマップ（例えば、後述の図15）を出力することができる。

【0082】

さらに、DNBのメンバーである生体分子のリストを出力することもできる。上述したように、DNBは、正常状態から疾病状態へ遷移する疾病前状態に現れるものであるが、DNBとして検出された生体分子、即ち遺伝子、タンパク質又は代謝物自身が、必ずしも疾病を悪化させる要因となる病的遺伝子、タンパク質又は代謝物とは限らない。ただし、DNBメンバーの一部が疾病と関連していることが判明している。

30

【0083】

そのため、検出したDNBのメンバーに含まれた、特定の疾病と関連性のあるもの（遺伝子、タンパク質又は代謝物）を抽出すると、例えば、医師の診断により、検診者等の診断対象が発症する可能性のある疾病をある程度把握することができる。

【0084】

したがって、上述した「診断結果の出力」（図8のs 500）に続いて、遺伝子、タンパク質又は代謝物と疾病との対応関係を記憶したデータベースを利用して、検出したDNBから疾病と関連性のある遺伝子、タンパク質又は代謝物を抽出して、診断の参考となる診断結果として出力することができる。

40

【0085】

そうすれば、健康診断を受ける人から例えば血液を採取して、その血液から得た遺伝子、タンパク質又は代謝物のデータからDNBを検出した場合、当該DNBに含まれている遺伝子、タンパク質又は代謝物と関連性がある疾病をある程度特定することができ、検診者が持つ潜在的疾病を早期段階で診断することができる。

【0086】

50

< 検出装置 >

以上詳述したDNBの検出方法は、コンピュータを用いた検出装置として本発明を具現化することができる。図11は、本発明に係る検出装置の構成例を示すブロック図である。図11に示す検出装置1は、パーソナルコンピュータ、サーバコンピュータに接続されるクライアントコンピュータ、その他各種コンピュータを用いて実現される。検出装置1は、制御部10、記録部11、記憶部12、入力部13、出力部14、取得部15等の各種機構を備えている。

【0087】

制御部10は、CPU (Central Processing Unit) 等の回路を用いて構成され、検出装置1全体を制御する機構である。

10

【0088】

記録部11は、HDD (Hard Disk Drive) 等の磁気記録機構、SSD (Solid State Disk) 等の不揮発性半導体記録機構等の不揮発性の補助記録機構である。記録部11には、本発明に係る検出プログラム11a等の各種プログラム及びデータが記録されている。

【0089】

記憶部12は、SDRAM (Synchronous Dynamic Random Access Memory)、SRAM (Static Random Access Memory) 等の揮発性の主記憶機構である。

【0090】

入力部13は、キーボード、マウス等のハードウェア、及びドライバ等のソフトウェアを備える入力機構である。

20

【0091】

出力部14は、モニタ、プリンタ等のハードウェア、及びドライバ等のソフトウェアを備える出力機構である。

【0092】

取得部15は、外部から各種データを取得する機構である。具体的には、通信網を介してデータを取り込むLAN (Local Area Network) ポート、測定機器に接続可能なパラレルケーブル等の専用線に接続するポート等の各種ハードウェア、及びドライバ等のソフトウェアである。

【0093】

そして、記録部11に記録されている検出プログラム11aを、記憶部12に記憶し、制御部10の制御に基づき実行することにより、コンピュータは、検出プログラム11aに係る各種手順を実行し、本発明の検出装置1として機能する。なお、便宜上、記録部11及び記憶部12として分けているが、両者とも各種情報の記録という同様の機能を有するものであり、装置の仕様、運用形態等に応じていずれの機構に記録させるかは、適宜決定することが可能である。

30

【0094】

図12は、本発明に係る検出装置1によるDNBの検出処理の一例を示すフローチャートである。本発明に係る検出装置1の処理は、前述のDNBの検出処理を実行するものである。検出装置1の制御部10は、取得部15により、生体に関する測定により得られた複数の因子項目についての測定データを取得する (Sc1)。ステップSc1は、図3のステップs1として示したハイスループットデータの取得処理に対応する。なお、ここではコンピュータ処理の対象として表現するため、因子項目として表現しているが、ここでいう因子項目とは、前述のDNBのノードとなり得る遺伝子に関する測定項目、タンパク質に関する測定項目、代謝物に関する測定項目等の測定項目を示している。なお、CTスキャン等の測定装置にて出力される体内画像から得られる画像に関する測定項目を用いることも可能である。

40

【0095】

制御部10は、取得した因子項目のそれぞれの測定データが、有意性をもって経時的に変化しているか否かを検定し、検定した結果に基づいて差次的生体分子を選出する (Sc2)。ステップSc2は、図3のステップs2として示した差次的生体分子の選出処理に

50

対応する。

【0096】

したがって、ステップSc2の処理は、制御部10が、各因子項目の測定データ、並びに因子項目及び時系列毎に予め設定されている参照データとの比較結果に基づいて、有意性に関する検定を行い（Sc21）、経時変化に有意性があると検定された因子項目を選出する処理（Sc22）を含む。即ち、図4に示した各種処理を実行する。なお、検出装置1が参照データとして処理するデータとは、コントロールサンプルであり、例えば、最初に取得したサンプルをコントロールサンプルとする等の設定に基づき、検出装置1は、当該サンプルに対して参照データとしての取り扱いを行う。

【0097】

制御部10は、選出した各因子項目のそれぞれの測定データの時系列変化の相関関係に基づいて複数の因子項目を複数のクラスターに分類する（Sc3）。ステップSc3は、図3のステップs3として示したクラスター化処理に対応する。

【0098】

制御部10は、分類した各クラスターから、各因子項目のそれぞれの測定データの時系列変化及び各因子項目間の測定データの時系列変化の相関関係に基づいて予め設定された選出条件に該当するクラスターを選出する（Sc4）。ステップSc4は、図3のステップs4として示したDNBの候補の選出処理に対応する。

【0099】

したがって、ステップSc4の処理は、制御部10が、それぞれのクラスターに対し、クラスター内の各因子項目の測定データの相関を示す値の平均値を第1指数として算出し（Sc41）、クラスター内の因子項目の測定データと当該クラスター外の因子項目の測定データとの間の相関を示す値の平均値を第2指数として算出し（Sc42）、クラスター内の各因子項目について測定データの標準偏差の平均値を第3指数として算出する処理（Sc43）を含む。さらに、ステップSc4の処理は、制御部10が、第1指数と、第2指数と、第3指数の逆数との積に基づく総合指数を算出し（Sc44）、算出した総合指数が最大であるクラスターを選出する処理（Sc45）を含む。即ち、図5に示した各種処理を実行する。なお、第1指数、第2指数、第3指数及び総合指数としては、例えば、クラスター内のノードの間のピアソン相関係数の絶対値の平均値PCCd(k)、クラスター内のノードと他のノードとの間のピアソン相関関数の絶対値の平均値OPPCd(k)、クラスター内のノードの標準偏差の平均値SDd(k)、及び総合指数I(k)がそれぞれ用いられる。

【0100】

制御部10は、選出したクラスターに含まれる因子項目をバイオマーカーの候補として検出する（Sc5）。ステップSc5は、図3のステップs5として示したDNBの判定処理に対応する。

【0101】

したがって、ステップSc5の処理は、制御部10が、各因子項目について、対応する参照データの標準偏差の平均値を示す参照標準偏差を算出し（Sc51）、因子項目間の相関を示す値の平均値を示す参照相関値を算出する（Sc52）。そして、第1指数が参照標準偏差に比べて有意性をもって増大し、第2指数が前記参照相関値に比べて有意性をもって減少し、かつ、第3指数が参照標準偏差に比べて有意性をもって増大した場合に、当該クラスターに含まれる項目をバイオマーカーとして検出する処理（Sc53）を含む。即ち、図6に示した各種処理を実行する。なお、参照標準偏差及び参照相関値としては、例えば、各ノード間のピアソン相関係数の絶対値の平均値PCCDr及び各ノードの標準偏差の平均値SDDrがそれぞれ用いられる。

【0102】

そして、制御部10は、バイオマーカーの候補として検出した因子項目を出力部14から出力し（Sc6）、処理を終了する。

【0103】

10

20

30

40

50

## &lt; 第 1 の検証例 &gt;

本発明の D N B による疾病の早期診断方法の診断精度を検証するために、肺障害が引き起こされたマウスの実験データを用いて、本発明の診断方法で診断を行い、その診断結果を実際の病状進行状況と比較して、本発明の診断方法の有効性を検証した。次に、当該検証例を詳細に説明する。実験データは、複数の実験用 C D - 1 雄マウスをケースグループと、コントロールグループとに分けて、ケースグループを通常の空気環境に、コントロールグループを有毒ガスであるホスゲンが含まれている空気環境に置き、そして、両グループのマウスの健康状況を観察し、ホスゲンの吸入による急性肺損傷の分子レベルのメカニズムを調べるといふ実験から得られたものである。当該実験データを用いて、本発明の診断方法を用いて、ホスゲン曝露されているケースグループのマウスの健康状況を診断した。通常、マウスは一定量のホスゲンを吸入すると、ホスゲン誘発性肺障害を発症する。

10

## 【 0 1 0 4 】

図 1 3 は、第 1 の検証例における診断用のデータを示すテーブルである。図 1 3 に示すように、診断対象は、ホスゲン誘発性肺障害にかかったマウス ( C D - 1 雄マウス ) であり、サンプル採取対象は、診断対象となるケースグループ及び参照対象となるコントロールグループのマウスの肺組織であり、サンプリングポイントは、実験開始後の 0、0.5、1、4、8、12、24、48、72 時間が経過した時間点であり、D N B の検出に用いた遺伝子の数は 2 2 6 9 0 個である。

## 【 0 1 0 5 】

本発明の診断方法を用いて、具体的に、次の処理を行った。

20

## 【 0 1 0 6 】

まず、それぞれのサンプルから測定したハイスループット遺伝子データから差次的発現遺伝子を選択する。各々のサンプリングポイントにおいて、6 つのケースサンプルと 6 つコントロールサンプルが提供されている。最初のサンプリングポイント ( 0 h ) において、ケースサンプルのデータはコントロールサンプルのデータと同じである。

## 【 0 1 0 7 】

各々のサンプリングポイントにおいて、有意水準  $p < 0.05$  のスチューデントの t 検定を用いて、それぞれ、 $A = [ 0, 53, 184, 1325, 1327, 738, 980, 1263, 915 ]$  の差次的遺伝子を選出した。

## 【 0 1 0 8 】

そして、選出した差次的遺伝子のセット A に対して、誤発現率 ( F D R ) 及び 2 倍変更スクリーニングを用いることにより、各サンプリングポイントに対して、それぞれ、 $B = [ 0, 29, 72, 195, 269, 163, 173, 188, 176 ]$  の遺伝子を選出した。

30

## 【 0 1 0 9 】

選出した遺伝子セット B に対して、各サンプリングポイントにおいて、相関の高いものを 1 つのクラスターにまとめるクラスター化を行い、それぞれ、40 個のクラスターを得た。

## 【 0 1 1 0 】

さらに、各サンプリングポイントにおいて、得られた 40 個のクラスター内の全ての遺伝子に対して、データの正規化を行う。そして、各サンプリングポイントにおいて、正規化されたコントロールグループ及びケースグループにおける各クラスターのそれぞれの標準偏差の平均値 S D d ( 第 3 指数 )、クラスターのメンバー間のピアソンの相関係数の絶対値の平均値 P C C d ( 第 2 指数 )、クラスターのメンバーと他の遺伝子との間のピアソン相関係数の絶対値の平均値 O P C C d ( 第 1 指数 )、及び総合指数 I を計算する。

40

## 【 0 1 1 1 】

そして、各サンプリングポイントにおいて、計算したケースグループにおける各クラスターから、総合指数 I が一番大きいクラスターを D N B の候補として選出し、さらに、当該 D N B の候補に対して、コントロールグループの標準偏差の平均値 S D c、遺伝子間のピアソン相関係数の絶対値の平均値 P C C c を基準に、有意性分析により、D N B である

50

か否かを判断する。その結果、各サンプリングポイントにおいて、DNBとなるクラスターの数は、[ 0、0、0、0、1、0、0、0、0 ]であった。

【 0 1 1 2 】

即ち、第5のサンプリングポイント(8h)でDNBが検出された、当該DNBは220個の遺伝子を有する第111番目のクラスターである。

【 0 1 1 3 】

図14Aは、第1の検証例において、検出されたDNBの候補の標準偏差の平均値SDdの時系列変化の例を示すグラフである。図14Bは、第1の検証例において、検出されたDNBの候補のクラスターのメンバー間のピアソン相関係数の絶対値の平均値PCCdの時系列変化の例を示すグラフである。図14Cは、第1の検証例において、検出されたDNBの候補のクラスターのメンバーと他の遺伝子との間のピアソン相関係数の絶対値の平均値OPCCdの時系列変化の例を示すグラフである。図14Dは、第1の検証例において、検出されたDNBの候補の総合指数Iの時系列変化の例を示すグラフである。図14A~図14Dにおいて、横軸は時間段階tを表し、縦軸は、それぞれ、標準偏差の平均値SDd(図14A)、クラスターのメンバー間のピアソン相関係数の絶対値の平均値PCCd(図14B)、クラスターのメンバーと他の遺伝子との間のピアソン相関係数の絶対値の平均値OPCCd(図14C)、及び総合指数Iを示している(図14D)。破線は、ケースグループから検出したDNBの候補の各種指数の経時変化を示し、実線は、コントロールグループから選んだ1クラスターの各種指数の経時変化を示している。

10

【 0 1 1 4 】

図14A~図14Dから分かるように、第4の時間段階(即ち、4時間経過)から、DNBの候補の第1指数PCCd、第3指数SDd及び総合指数Iは大きく増加し始め、第5の時間段階(即ち、8時間経過)でピークに達している。一方、DNBの候補の第3指数OPCCdは第2の時間段階から低減し、同じ第5の時間段階(即ち、8時間経過)で極小値を示している。

20

【 0 1 1 5 】

また、DNBの動的特性を直観的に示すために、DNBを含む全遺伝子ネットワークの動的特性を図15に示す。図15は、第1の検証例において、ケースグループの遺伝子によって構成されたネットワークにおけるDNBの動的特性の一例を経時的に示すマップである。図15には、順に0.5、1、4、8、12、24、48、72hの各サンプリングポイントで、ケースグループの遺伝子ネットワーク(3452個の遺伝子、9238個のリンク)が表示されており、「○」で表示しているノードはDNBの候補に属する遺伝子であり、「□」で表示しているノードはDNBの候補のノード付近の他の遺伝子であり、ノードとノードとの間のラインは両ノードの相関性を示している。また、「○」の色の濃さが遺伝子の標準偏差SDの大きさを示し、両ノードを接続ラインの濃さが両ノードの相関係数PCCの絶対値の大きさを示している。全てのマップはCytoscape(<http://www.cytoscape.org/>)を用いて構成されたものである。

30

【 0 1 1 6 】

図15に示すように、DNBの候補の特性(SD、PCC)が時間の経過とともに変化し、他の遺伝子と同じ振る舞いをする通常のクラスターから、次第にDNBへと進化していく。図15のeに示す第5の段階(8hが経過した時)で、DNBとしての特徴が最も著しく、疾病前状態(8h)である警告信号をはっきり示している。しかしながら、疾病状態へ遷移した後(24h、48h、72h)、DNBのメンバーは、再び他の遺伝子と同じ振る舞いをするようになっている。

40

【 0 1 1 7 】

当該結果から、疾病前状態は第5の時間段階の近くに存在し、第5の時間段階の後、システムは疾病状態へ遷移するということが判明される。

【 0 1 1 8 】

したがって、本発明のDNBによる疾病の早期診断方法で診断する場合、第4の時間段階での診断で、疾病になる警告信号が少し見えているため、近いうち疾病が悪化するとの

50

診断結果を出すことができる。そして、第5の時間段階での診断では、疾病になる警告信号をはっきり見えているため、間もなく疾病になるとの診断結果を出すことができる。

【0119】

一方、実際のマウス実験の結果では、ケースグループのマウスは、ホスゲンを吸入してから8時間後に、肺水腫が起こし、12時間後に50%~60%が死亡し、24時間後に、さらに60%~70%が死亡した。

【0120】

したがって、本発明のDNBによる早期診断の診断結果は、マウスの実際の病気悪化状況と完全に一致しているということができる。

【0121】

<第2の検証例>

上述した第1の検証例において、動物実験のデータを用いて、本発明のDNBによる疾病の早期診断方法の有効性を検証した。本検証例では、B細胞リンパ腫の臨床データを用いて、さらに、本発明のDNBによる疾病の早期診断方法の診断精度を検証する。

【0122】

図16は、第2の検証例における診断データをリストしたテーブルである。図16に示すように、臨床像、病変及びフローサイトメトリーに基づいて、サンプルは、安静期(P1)、活性期(P2)、限界期(P3)、転移期(P4)、侵攻期(P5)の5つの段階のグループに分けられている。各段階において、サンプルの数は、それぞれ、5、3、6、5、7である。各段階において、脾腫は、それぞれ、「None」、「None」、「+/-」、「+」、「+++」である。各段階において、フローサイトメトリーは、それぞれ、「正常安静」、「正常活性」、「異常」、「ミックスト」、「B-1クローン」である。また、安静期(P1)で採取したサンプルをコントロールサンプルとし、他の段階(P2~P5)で採取したサンプルをケースサンプルとする。

【0123】

上記26個のサンプルから測定した遺伝子発現データから13712個の遺伝子を対象に、上述したDNBによる疾病の早期診断方法で診断を行った。その結果として、図17A~図17Dに、ケースグループの遺伝子から検出したDNBの候補の各指数を示す。図17Aは、第2の検証例において、検出されたDNBの候補の標準偏差の平均値SDdの時系列変化の例を示すグラフである。図17Bは、第2の検証例において、検出されたDNBの候補のクラスターのメンバー間のピアソン相関係数の絶対値の平均値PCCdの時系列変化の例を示すグラフである。図17Cは、第2の検証例において、検出されたDNBの候補のクラスターのメンバーと他の遺伝子との間のピアソン相関係数の絶対値の平均値OPCCdの時系列変化の例を示すグラフである。図17Dは、第2の検証例において、検出されたDNBの候補の総合指数Iの時系列変化の例を示すグラフである。

【0124】

図17A~図17Dにおいて、横軸は段階(P1~P4)の番号を示しており、縦軸は、それぞれ、標準偏差の平均値SDd(図17A)、クラスターのメンバー間のピアソン相関係数の絶対値の平均値PCCd(図17B)、クラスターのメンバーと他の遺伝子との間のピアソン相関係数の絶対値の平均値OPCCd(図17C)、及び総合指数Iを示している(図17D)。

【0125】

図17A~図17Dから明らかなように、第2の段階(P2)、即ち活性期において、DNBの候補の総合指数Iがピークに達しており、疾病状態を示す警告信号が最も強くなっている。この診断結果は、実際の病変と完全に一致している。実際の臨床データでも、活性期の直後の限界期から病状が悪化して、脾腫が「+/-」となり、フローサイトメトリーが「異常」になっている。したがって、本検証例におけるDNB特定の解析結果と実際の臨床データと完全に一致している。

【0126】

また、通常の診断では、図16に示すように、活性期における脾腫が「None」、フ

10

20

30

40

50

ローサイトメトリーが「正常活性」となっているため、異常が認めないとの診断結果になる。それに対して、本発明のDNBによる疾病の早期診断方法で診断する場合、活性期において、疾病前状態を示す警告信号(DNB)を検出したため、「異常の兆候が見られる」との診断結果を患者に伝えることができる。そのため、患者は早期段階で治療措置を取ることにより、病状の悪化を止める。

【0127】

したがって、本発明のDNBによる疾病の早期診断は、リンパ腫のような複雑疾病の早期診断に大変有効であることが検証されている。

【0128】

また、本検証例において検出されたDNBの中に、22個の遺伝子とTFsが含まれており、そのうち、13個の遺伝子が明らかにB細胞リンパ腫と関連しており、さらに、当該13個の遺伝子のうちの8個が増殖のマスタレギュレーターであることが判明されている。したがって、本発明におけるDNBは、疾病前状態を示す警告信号として、早期段階で異常の兆候を患者に伝えることができるだけでなく、疾病と関連している遺伝子も具体的に特定することができるため、複雑疾病の治療、製薬に非常に役に立つと思われる。

【0129】

前記実施の形態は、本発明の無数に存在する実施例の一部を開示したに過ぎず、疾病の種類、検出すべき目的等の様々な要因を加味して適宜設計変更することが可能である。特に、因子項目としては、生体に関する測定により得られた情報であれば、様々な測定データを用いることができる。例えば、前述の遺伝子、タンパク質、代謝物に関する測定データに限らず、CTスキャン等の測定装置にて出力される体内画像に基づいて、各部位の様々な状況を数値化することにより、測定データとして用いることが可能である。更には、画像以外にも、声又は体内から発せられる音を測定し、数値化した上で測定データとして用いることも可能である。

【符号の説明】

【0130】

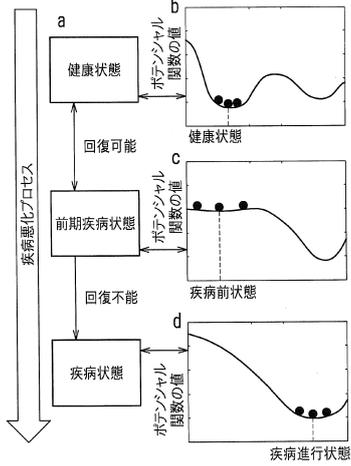
- 1 検出装置
- 10 制御部
- 11 記録部
- 12 記憶部
- 13 入力部
- 14 出力部
- 15 取得部
- 11a 検出プログラム

10

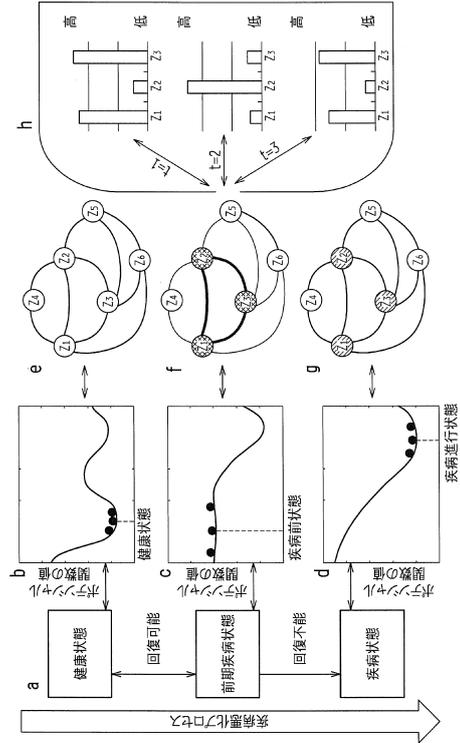
20

30

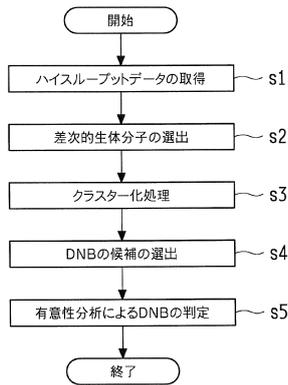
【図1】



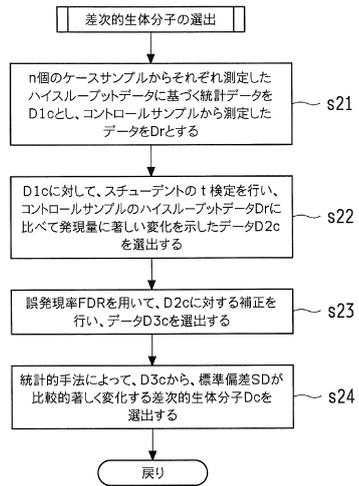
【図2】



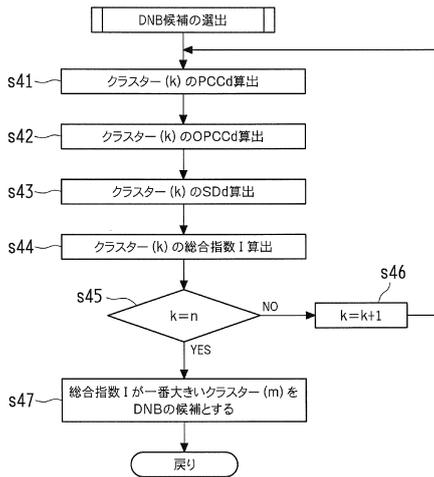
【図3】



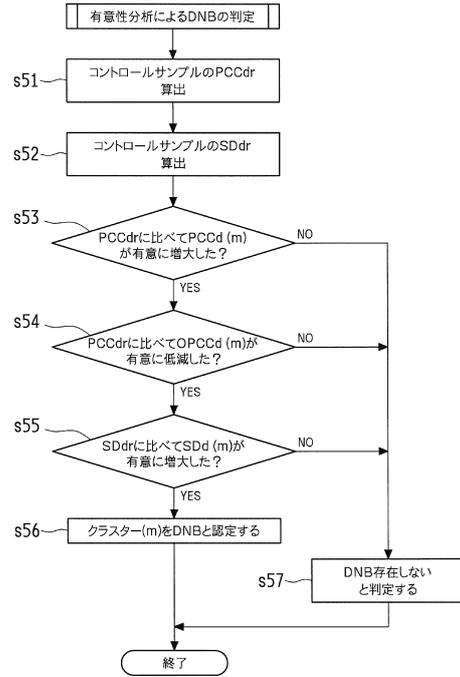
【図4】



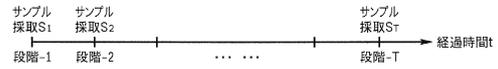
【図5】



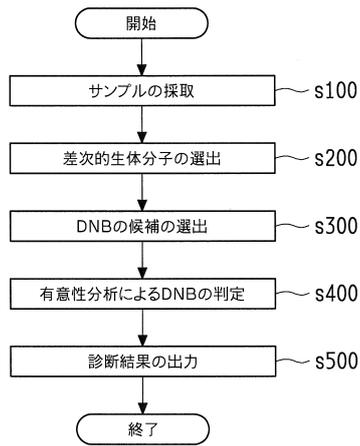
【図6】



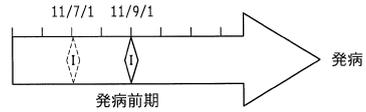
【図7】



【図8】



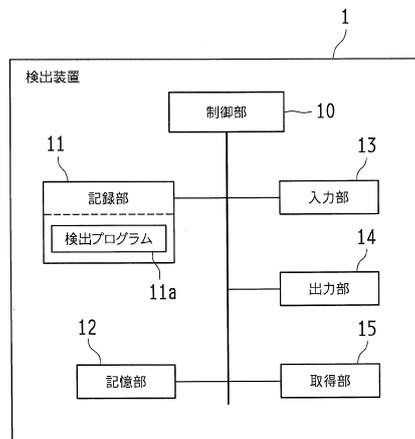
【図10】



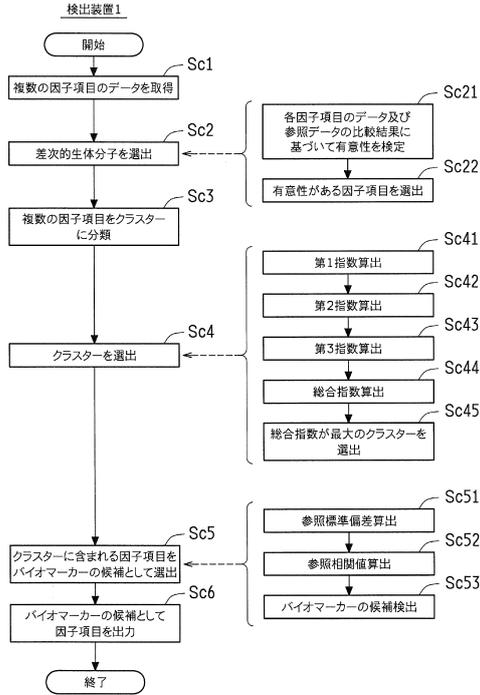
【図9】



【図11】



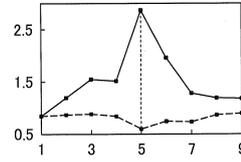
【図12】



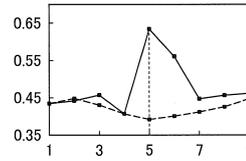
【図13】

診断対象	ホスゲン誘発性肺障害にかかったマウス(CD-1雄マウス)
サンプル採集対象	診断対象とするケースグループのマウスの肺組織 参照対象とするコントロールグループのマウスの肺組織
サンプルリングポイント	0, 0.5, 1, 4, 8, 12, 24, 48, 72(h)
観察対象	22690個の遺伝子

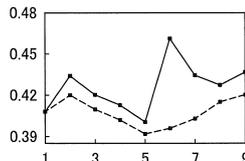
【図14A】



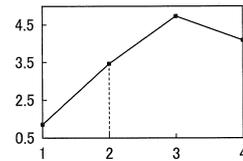
【図14B】



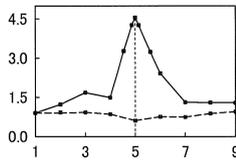
【図14C】



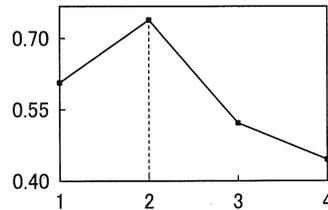
【図17A】



【図14D】



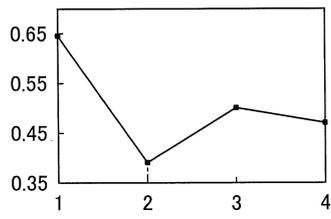
【図17B】



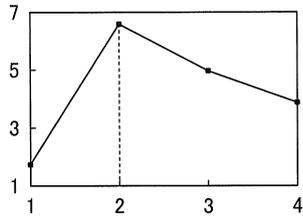
【図16】

サンプル採集対象	B細胞リンパ腫患者			
	病変	サンプル数	脾腫	フローサイトメトリー
P1	安静期	5	None	正常安静
P2	活性期	3	None	正常活性
P3	限界期	6	+/-	異常
P4	転移期	5	+	ミックスト
P5	侵襲期	7	+++	B-1クローン
ケースグループ	P2~P5で採集したサンプル			
コントロールグループ	P1で採集したサンプル			
観察対象	13712個の遺伝子			

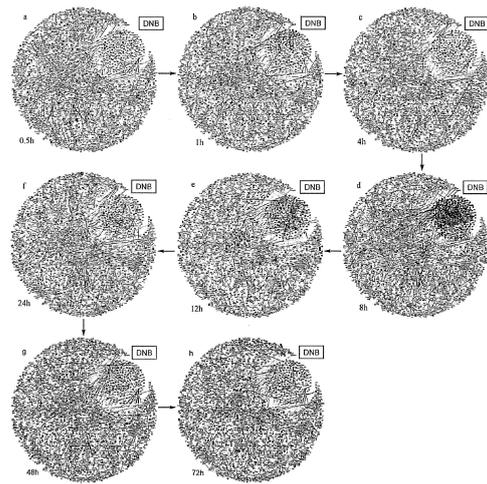
【 17 C】



【 17 D】



【 15】



---

フロントページの続き

(72)発明者 劉 治平

中華人民共和国山東省済南市 山東大学 教工1楼1-2-501室

(72)発明者 李 美儀

中華人民共和国上海市岳陽路319号中科大厦

審査官 梅岡 信幸

(56)参考文献 特開2009-057337(JP,A)

国際公開第2008/102825(WO,A1)

特表2012-514783(JP,A)

特表2005-527904(JP,A)

特表2004-524604(JP,A)

特開2012-094143(JP,A)

(58)調査した分野(Int.Cl., DB名)

G06F 19/10 - 19/28

C12Q 1/68

G01N 33/50

G01N 33/68

C12N 15/09