

(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2005-521138

(P2005-521138A)

(43) 公表日 平成17年7月14日(2005.7.14)

(51) Int.Cl. <sup>7</sup>	F I	テーマコード (参考)
GO6N 3/00	GO6N 3/00 540	4B063
GO1N 33/53	GO1N 33/53 D	
GO1N 33/574	GO1N 33/574 Z	
// C12N 15/00	C12N 15/00 Z	
C12Q 1/68	C12Q 1/68 A	

審査請求 未請求 予備審査請求 有 (全 65 頁)

(21) 出願番号 特願2003-577211 (P2003-577211)  
 (86) (22) 出願日 平成15年3月17日 (2003.3.17)  
 (85) 翻訳文提出日 平成16年11月9日 (2004.11.9)  
 (86) 国際出願番号 PCT/NZ2003/000045  
 (87) 国際公開番号 W02003/079286  
 (87) 国際公開日 平成15年9月25日 (2003.9.25)  
 (31) 優先権主張番号 517817  
 (32) 優先日 平成14年3月15日 (2002.3.15)  
 (33) 優先権主張国 ニュージーランド (NZ)

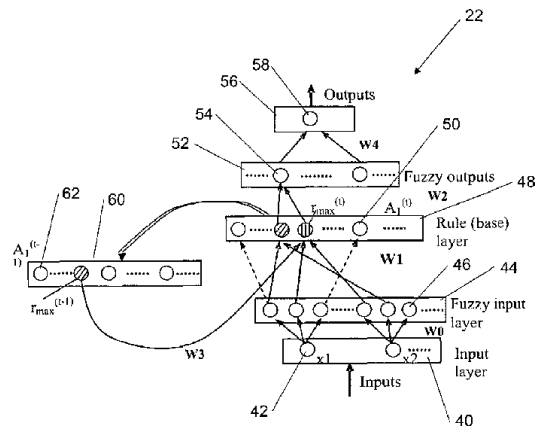
(71) 出願人 504350120  
 パシフィック エッジ バイオテクノロジー  
 リミテッド  
 ニュージーランド国, ダニーディン, セン  
 ト デイビッド ストリート 87, ザ  
 センター フォー イノベーション  
 (74) 代理人 100099759  
 弁理士 青木 篤  
 (74) 代理人 100077517  
 弁理士 石田 敬  
 (74) 代理人 100087413  
 弁理士 古賀 哲次  
 (74) 代理人 100108903  
 弁理士 中村 和広

最終頁に続く

(54) 【発明の名称】 遺伝子発現データを使用する適応学習システムの医療適用

(57) 【要約】

ニューラルネットワークモジュールが提供される。これは、遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む入力層を備える。これは、1個又は複数のルール節点を含むルールベース層、並びに1種又は複数の状態を出力するために構成された1個又は複数の出力節点を含む、出力層も備える。これは更に、遺伝子発現データと1種又は複数の状態の関係を表わしているルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備える。このモジュールを使用する方法及びシステムに加え、このシステムを利用する特異的のプロファイルが、明らかにされている。



## 【特許請求の範囲】

## 【請求項1】

遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の状態を出力するために構成された1個又は複数の出力節点を含む、出力層；及び、遺伝子発現データと1種又は複数の状態の間の関係を表わしているルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備える、ニューラルネットワークモジュール。

## 【請求項2】

遺伝子発現データと1種又は複数の状態の間の関係を決定する汎用方法であり、少なくとも：

- a) 1種又は複数の予め決定された状態のクラスにカテゴリ化された遺伝子発現データのセットを提供する工程；
- b) 該遺伝子発現データ及び該1種又は複数の予め決定された状態のクラスについて、ニューラルネットワークモジュールを訓練する工程であり、ここで、ニューラルネットワークモジュールが、遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の状態のクラスを出力するために構成された1個又は複数の出力節点を含む、出力層；及び、遺伝子発現データと1種又は複数の状態のクラスの間関係を表わすルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備える工程；並びに
- c) ルールベース層からルールを抽出する工程であり、該ルールが、遺伝子発現データと1種又は複数の状態のクラスの間関係を表わす工程を含む、方法。

10

20

## 【請求項3】

遺伝子発現データと1種又は複数の状態の間の関係を決定するシステムであり：

- a) 1種又は複数の予め決定された状態のクラスにカテゴリ化された遺伝子発現データのセットを受け取ることが可能な、入力；
- b) 該遺伝子発現データ及び該1種又は複数の予め決定された状態のクラスについて訓練可能なニューラルネットワークモジュールであり、ここでこのニューラルネットワークモジュールが、遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の状態のクラスを出力するために構成された1個又は複数の出力節点を含む、出力層；及び、遺伝子発現データと1種又は複数の状態のクラスの間関係を表わすルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備える、モジュール；並びに
- c) ルールベース層からルールを抽出する手段であり、該ルールが、遺伝子発現データと1種又は複数の状態のクラスの間関係を表わしている手段を具備する、システム。

30

## 【請求項4】

患者の状態を診断する汎用方法であり、該患者の生物学的試料から抽出された遺伝子発現データが、その遺伝子発現データと1種又は複数の状態のクラスの間関係を表わすルールを満足するかどうかを決定する工程を含み、該ルールが、請求項2記載の方法により決定される、方法。

## 【請求項5】

患者の状態を診断する汎用システムであり、該患者の生物学的試料から抽出された遺伝子発現データが、その遺伝子発現データと1種又は複数の状態のクラスの間関係を表わすルールを満足するかどうかを決定する手段を備え、該ルールが、請求項2記載の方法により決定される、システム。

40

## 【請求項6】

1種又は複数の状態に連結された過剰発現された又は過小発現された遺伝子を識別するセットを選択する方法であり：

- a) 1種又は複数の予め決定された状態にカテゴリ化された遺伝子発現データの1種又は複数のセットを提供する工程；
- b) 該遺伝子発現データ及び1種又は複数の予め決定された状態について、ニューラルネッ

50

トワークモジュールを訓練する工程であり、このニューラルネットワークが、遺伝子発現データを受け取るために構成された、1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の状態を出力するために構成された1個又は複数の出力節点を含む、出力層；及び、遺伝子発現データと1種又は複数の状態の間の関係を表わすルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備え、該適応構成要素が、入力データを基にしたルールベース層中の選択された2個又はそれよりも多いルール節点を集合するように配置されている、工程；

c) 適応構成要素に、ルールベース層内の選択された2個又はそれよりも多いルール節点を集合させる工程；

d) ルールベース層からルールを抽出する工程であり、該ルールが、遺伝子発現データと1種又は複数の状態の間の関係を表わしている、工程；

e) 抽出されたルールから、過剰発現された又は過小発現された遺伝子を同定する工程であり、この遺伝子は、1種又は複数の状態に連結された過剰発現された又は過小発現された遺伝子を識別するセットを表わしている工程を含む、方法。

10

【請求項7】

1種又は複数の状態に連結された過剰発現された又は過小発現された遺伝子を識別するセットを選択するシステムであり；

a) 1種又は複数の予め決定された状態へカテゴリ化された遺伝子発現データの1種又は複数のセットを受け取ることが可能な、入力；

b) 該遺伝子発現データ及び1種又は複数の予め決定された状態を受け取るように適合された、ニューラルネットワークモジュールであり、このニューラルネットワークは、遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の状態を出力するために構成された1個又は複数の出力節点を含む、出力層；及び、遺伝子発現データと1種又は複数の状態の間の関係を表わしているルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備え、該適応構成要素が、入力データを基にしたルールベース層中の選択された2個又はそれよりも多いルール節点を集合するように配置されている、モジュール；

20

c) ルールベース層からルールを抽出するための、抽出構成要素であり、該ルールが、遺伝子発現データと1種又は複数の状態の間の関係を表わしている、要素；及び

30

d) 抽出されたルールから過剰発現された又は過小発現された遺伝子を同定するための、識別子であり、これらの遺伝子が、1種又は複数の状態に連結された過剰発現された又は過小発現された遺伝子を識別するセットを表わしている、識別子を備える、システム。

【請求項8】

請求項7記載のように1種又は複数の状態に連結された過剰発現された又は過小発現された遺伝子を識別するセットを選択するための、システムであり、ここで抽出されたルールが、特異的遺伝子発現を、特定の状態に連結させる、システム。

【請求項9】

遺伝子発現セットを減少する方法であり；

a) 1種又は複数の状態によりカテゴリ化された遺伝子発現データの1種又は複数のセットを提供する工程；

40

b) 該遺伝子発現データ及び1種又は複数の状態のクラスについてニューラルネットワークモジュールを訓練する工程であり、このニューラルネットワークが、遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の状態を出力するために構成された1個又は複数の出力節点を含む、出力層；及び、遺伝子発現データと1種又は複数の状態の間の関係を表わしているルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備え、該適応構成要素が、入力データを基にしたルールベース層中の選択された2個又はそれよりも多いルール節点を集合するように配置されている、工程；

c) 適応構成要素に、ルールベース層内の選択された2個又はそれよりも多いルール節点を

50

集合させる工程；

d) ルールベース層からルールを抽出する工程であり、該ルールが、遺伝子発現データと1種又は複数の状態の間の関係を表わしている工程；

e) 抽出されたルールから、遺伝子を同定する工程であり、この遺伝子が、1種又は複数の状態に連結された減少された遺伝子発現セットを表わしている工程を含む、方法。

【請求項10】

遺伝子発現セットを減少するシステムであり；

a) 1種又は複数の状態によりカテゴリ化された遺伝子発現データの1種又は複数のセットを受け取る、入力手段；

b) 該遺伝子発現データ及び1種又は複数の状態のクラスについて訓練可能なニューラルネットワークモジュールであり、このニューラルネットワークが、遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の状態を出力するために構成された1個又は複数の出力節点を含む、出力層；及び、遺伝子発現データと1種又は複数の状態の間の関係を表わしているルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備え、該適応構成要素が、入力データを基にしたルールベース層中の選択された2個又はそれよりも多いルール節点を集合するように配置されている、モジュール；

d) ルールベース層からルールを抽出するように適合された、ルール抽出手段であり、該ルールが、遺伝子発現データと1種又は複数の状態の間の関係を表わしている、手段；

e) 抽出されたルールから遺伝子を同定するように適合された識別子であり、この遺伝子は、1種又は複数の状態に連結された減少された遺伝子発現セットを表わす識別子を備える、システム。

【請求項11】

遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の状態を出力するために構成された1個又は複数の出力節点を含む、出力層；及び、遺伝子発現データと1種又は複数の状態の間の関係を表わしているルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備え、該適応構成要素が、遺伝子発現データを基にしたルールベース層中の選択された2個又はそれよりも多いルール節点を集合するように配置されているものを備える、ニューラルネットワークモジュール。

【請求項12】

請求項2記載の方法に従いニューラルネットワークから抽出された場合の、遺伝子発現データと1種又は複数の状態の間の関係を表わしている、ルール。

【請求項13】

状態を診断するためにニューラルネットワークを訓練する方法であり；

a) 1種又は複数の状態によりカテゴリ化された遺伝子発現データを提供する工程；

b) 遺伝子発現データ及び1種又は複数の状態についてニューラルネットワークモジュールを訓練する工程であり、ここでこのニューラルネットワークモジュールが、遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の状態を出力するために構成された1個又は複数の出力節点を含む、出力層；及び、遺伝子発現データと1種又は複数の状態の間の関係を表わしているルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備える、工程；

c) リーブワンアウト法を使用し、データを検定する工程；

d) 遺伝子発現データの入力を減少し、最良の検定精度を提供する工程；

e) その入力層のように、減少された遺伝子発現データを受入れるために、ニューラルネットワークモジュールを修飾する工程；

f) 修飾されたニューラルネットワークモジュールを訓練する工程；並びに

g) 適応構成要素からルールを抽出する工程を少なくとも含む、方法。

【請求項14】

更に、減少工程からの方法を反復する工程を含む、請求項13記載の状態を診断するニューラルネットワークを訓練する方法。

【請求項15】

状態を診断するために、ニューラルネットワークを訓練するシステムであり：

- a) 1種又は複数の状態によりカテゴリ化された遺伝子発現データを提供する、受け取り可能な入力；
- b) 遺伝子発現データ及び1種又は複数の状態について訓練可能なニューラルネットワークモジュールであり、ここでこのニューラルネットワークモジュールが、遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の状態を出力するために構成された1個又は複数の出力節点を含む、出力層；及び、遺伝子発現データと1種又は複数の状態の間の関係を表わしているルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備える、モジュール；
- c) 最良の検定精度を提供するために、遺伝子発現データを減少するように操作可能な、入力遺伝子発現データ減少子；
- d) ニューラルネットワークモジュールを、その入力層のように減少された遺伝子発現データを受入れるように修飾するための、ニューラルネットワーク修飾子；並びに
- e) ルールを適応構成要素から抽出するための、ルール抽出手段を備える、システム。

10

【請求項16】

抽出されたルールは、人間が読取り可能な形で表わされる、請求項15記載の状態を診断するためにニューラルネットワークを訓練するシステム。

20

【請求項17】

ニューラルネットワークモジュールが更に、1種又は複数の状態への十分な連結を示さない、ルールベース層内の節点をプルーニングするために配置されているプルーニングアルゴリズムを含む、請求項15又は16記載の状態を診断するためのニューラルネットワークを訓練するシステム。

【請求項18】

遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の予後の転帰を出力するために構成された1個又は複数の出力節点を含む、出力層；及び、遺伝子発現データと1種又は複数の予後の転帰の間の関係を表わしているルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備える、ニューラルネットワークモジュール。

30

【請求項19】

遺伝子発現データと予後の転帰の間の関係を決定する、汎用方法であり：

- a) 予め決定された予後の転帰により分類された遺伝子発現データのセットを提供する工程；
- b) 該遺伝子発現データ及び予後の転帰について、ニューラルネットワークモジュールを訓練する工程であり、ここでこのニューラルネットワークモジュールが、遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の予後の転帰を出力するために構成された1個又は複数の出力節点を含む、出力層；及び、遺伝子発現データと1種又は複数の予後の転帰の間の関係を表わすルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備える工程；
- c) ルールベース層からルールを抽出する工程であり、該ルールが、遺伝子発現データと1種又は複数の予後の転帰の間の関係を表わす工程を少なくとも含む、方法。

40

【請求項20】

該患者の生物学的試料から抽出された遺伝子発現データを、遺伝子発現データと予後の転帰の間の関係を表わすルールと相関するために適合された相関エンジンを備え、該ルールが、請求項19記載の方法により決定される、患者の状態の予後判定をする汎用システム

50

。

## 【請求項 2 1】

遺伝子発現データと予後の転帰の間の関係を決定する汎用システムであり：

- a) 予め決定された予後の転帰により分類された遺伝子発現データのセットを受け取るための、入力；
- b) 該遺伝子発現データ及び予後の転帰について訓練可能なニューラルネットワークモジュールであり、ここでこのニューラルネットワークモジュールが、遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の予後の転帰を出力するために構成された1個又は複数の出力節点を含む、出力層；及び、遺伝子発現データと1種又は複数の予後の転帰の間の関係を表わすルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備えるモジュール；
- c) ルールベース層からルールを抽出するために適合された、ルール抽出子であり、該ルールが、遺伝子発現データと1種又は複数の予後の転帰の間の関係を表わしている抽出子を備える、システム。

10

## 【請求項 2 2】

患者の状態の予後判定のための汎用方法であり、該患者の生物学的試料から抽出された遺伝子発現データを、遺伝子発現データと予後の転帰の間の関係を表わしているルールと相関することを含み、該ルールが、請求項19記載の方法により決定されている、方法。

## 【請求項 2 3】

患者の状態の予後判定のための汎用システムであり、該患者の生物学的試料から抽出された遺伝子発現データを、遺伝子発現データと予後の転帰の間の関係を表わしているルールと相関するために適合された相関手段を備え、該ルールが、請求項19記載の方法により決定される、システム。

20

## 【請求項 2 4】

1種又は複数の予後の転帰に連結された過剰発現又は過小発現された遺伝子を識別するセットを選択する方法であり：

- a) 1種又は複数の予め決定された予後の転帰へカテゴリ化された遺伝子発現データの1種又は複数のセットを提供する工程；
- b) 該遺伝子発現データ及び1種又は複数の予め決定された予後の転帰についてニューラルネットワークモジュールを訓練する工程であり、このニューラルネットワークが、遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の予後の転帰を出力するために構成された1個又は複数の出力節点を含む、出力層；及び、遺伝子発現データと1種又は複数の予後の転帰の間の関係を表わすルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備え、該適応構成要素が、入力データを基にしたルールベース層内の選択された2個又はそれよりも多いルール節点を集合するために配置されている、工程；
- c) 適応構成要素に、ルールベース層内の選択された2個又はそれよりも多いルール節点を集合させる工程；
- d) ルールベース層からルールを抽出する工程であり、該ルールが、遺伝子発現データと1種又は複数の予後の転帰の間の関係を表わしている、工程；
- e) 抽出されたルールから、過剰発現された又は過小発現された遺伝子を同定する工程であり、この遺伝子が、1種又は複数の予後の転帰に連結された過剰発現された又は過小発現された遺伝子を識別するセットを表わしている工程を含む、方法。

30

40

## 【請求項 2 5】

1種又は複数の予後の転帰に連結された過剰発現された又は過小発現された遺伝子を識別するセットを選択するシステムであり：

- a) 1種又は複数の予め決定された予後の転帰へカテゴリ化された遺伝子発現データの1種又は複数のセットを受け取ることが可能な、入力；

50

b) 該遺伝子発現データ及び1種又は複数の予め決定された予後の転帰について訓練可能な、ニューラルネットワークモジュールであり、このニューラルネットワークは、遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の予後の転帰を出力するために構成された1個又は複数の出力節点を含む、出力層；及び、遺伝子発現データと1種又は複数の予後の転帰の間の関係を表わしているルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備え、該適応構成要素が、入力データを基にしたルールベース層中の選択された2個又はそれよりも多いルール節点を集合するように配置されている、モジュール；

c) ルールベース層からルールを抽出するために適合された、ルール抽出子であり、該ルールが、遺伝子発現データと1種又は複数の予後の転帰の間の関係を表わしている、抽出子；並びに

d) 抽出されたルールから過剰発現された又は過小発現された遺伝子を同定するための、識別子であり、この遺伝子が、1種又は複数の予後の転帰に連結された過剰発現された又は過小発現された遺伝子を識別するセットを表わしている、識別子を備える、システム。

10

【請求項 26】

遺伝子発現セットを減少する方法であり；

a) 1種又は複数の予後の転帰によりカテゴリ化された遺伝子発現データの1種又は複数のセットを提供する工程；

b) 該遺伝子発現データ及び1種又は複数の状態のクラスについてニューラルネットワークモジュールを訓練する工程であり、このニューラルネットワークが、遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の予後の転帰を出力するために構成された1個又は複数の出力節点を含む、出力層；及び、遺伝子発現データと1種又は複数の予後の転帰の間の関係を表わしているルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備え、該適応構成要素が、入力データを基にしたルールベース層中の選択された2個又はそれよりも多いルール節点を集合するように配置されている、工程；

20

c) 適応構成要素に、ルールベース層内の選択された2個又はそれよりも多いルール節点を集合させる工程；

30

d) ルールベース層からルールを抽出する工程であり、該ルールが、遺伝子発現データと1種又は複数の予後の転帰の間の関係を表わしている、工程；

e) 抽出されたルールから、遺伝子を同定する工程であり、この遺伝子が、1種又は複数の予後の転帰に連結された減少された遺伝子発現セットを表わしている、工程を含む、方法。

【請求項 27】

遺伝子発現セットを減少するシステムであり；

a) 1種又は複数の予後の転帰によりカテゴリ化された遺伝子発現データの1種又は複数のセットを受け取る、入力；

b) 該遺伝子発現データ及び1種又は複数の状態のクラスについて訓練可能なニューラルネットワークモジュールであり、このニューラルネットワークが、遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の予後の転帰を出力するために構成された1個又は複数の出力節点を含む、出力層；及び、遺伝子発現データと1種又は複数の予後の転帰の間の関係を表わしているルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備え、該適応構成要素が、入力データを基にしたルールベース層中の選択された2個又はそれよりも多いルール節点を集合するように配置されている、工程；

40

d) ルールベース層からルールを抽出するように適合された、ルール抽出子であり、該ルールが、遺伝子発現データと1種又は複数の予後の転帰の間の関係を表わしている、抽出子

50

;

e)抽出されたルールから遺伝子を同定するように適合された同定手段であり、この遺伝子は、1種又は複数の予後の転帰に連結された減少された遺伝子発現セットを表わす手段を備える、システム。

【請求項28】

遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の予後の転帰を出力するために構成された1個又は複数の出力節点を含む、出力層；及び、遺伝子発現データと1種又は複数の予後の転帰の間の関係を表わしているルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備え、該適応構成要素が、遺伝子発現データを基にしたルールベース層中の選択された2個又はそれよりも多いルール節点を集合するように配置されている、ニューラルネットワークモジュール。

10

【請求項29】

請求項26記載の方法でニューラルネットワークから抽出された場合に、遺伝子発現データと1種又は複数の予後の転帰の間の関係を表わしているルール。

【請求項30】

予後の転帰を提供するニューラルネットワークを訓練する方法であり：

- a)1種又は複数の予後の転帰によりカテゴリ化された遺伝子発現データを提供する工程；
- b)遺伝子発現データ及び1種又は複数の予後の転帰についてニューラルネットワークモジュールを訓練する工程であり、ここでこのニューラルネットワークモジュールが、遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の予後の転帰を出力するために構成された1個又は複数の出力節点を含む、出力層；並びに、遺伝子発現データと1種又は複数の予後の転帰の間の関係を表わしているルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備える、工程；
- c)リーブワンアウト法を使用し、データを検定する工程；
- d)遺伝子発現データの入力を減少し、最良の検定精度を提供する工程；
- e)その入力層のように、減少された遺伝子発現データを受入れるために、ニューラルネットワークモジュールを修飾する工程；
- f)修飾されたニューラルネットワークモジュールを訓練する工程；並びに
- g)適応構成要素からルールを抽出する工程を少なくとも含む、方法。

20

30

【請求項31】

更に、減少工程からの方法を反復する工程を含む、請求項30記載の予後の転帰を提供するニューラルネットワークを訓練する方法。

【請求項32】

予後の転帰を提供するために、ニューラルネットワークを訓練するシステムであり：

- a)1種又は複数の予後の転帰によりカテゴリ化された遺伝子発現データを受け取るための、入力；
- b)遺伝子発現データ及び1種又は複数の予後の転帰について訓練可能なニューラルネットワークモジュールであり、ここでこのニューラルネットワークモジュールが、遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の予後の転帰を出力するために構成された1個又は複数の出力節点を含む、出力層；及び、遺伝子発現データと1種又は複数の予後の転帰の間の関係を表わしているルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備える、モジュール；
- c)最良の検定精度を提供するために、入力遺伝子発現データを減少する、減少子；
- d)ニューラルネットワークモジュールを、その入力層のように減少された遺伝子発現データを受入れるように修飾するために適合された、修飾子；並びに
- e)ルールを適応構成要素から抽出するための、ルール抽出子を備える、システム。

40

【請求項33】

50

抽出されたルールは、人間が読取り可能な形で表わされる、請求項32記載の、予後の転帰を提供するために、ニューラルネットワークを訓練するシステム。

【請求項34】

ニューラルネットワークモジュールが更に、1種又は複数の予後の転帰への十分な連結を示さない、ルールベース層内の節点をプルーニングするために配置されているプルーニングアルゴリズムを含む、請求項32又は33記載の予後の転帰を提供するためのニューラルネットワークを訓練するシステム。

【請求項35】

プロファイリングされた疾患が、乳腺癌、前立腺腺癌、肺腺癌、結腸直腸腺癌、リンパ腫、膀胱移行細胞癌、黒色腫、子宮腺癌、白血病、腎細胞癌、膵臓腺癌、卵巣腺癌、胸膜中皮腫、及び中枢神経系癌からなる群より選択される、請求項32～34のいずれか1項記載の予後の転帰を提供するためのニューラルネットワークを訓練するシステム。

10

【請求項36】

本明細書の説明と合わせて読む図面に説明された14種の癌：乳房-、前立腺-、膀胱-、白血病-、リンパ腫-、中枢神経系-、肺-、結腸直腸-、黒色腫、子宮-、腎細胞-、膵-、卵巣-、及び胸膜癌のいずれかひとつの遺伝子プロファイル。

【請求項37】

本明細書の説明と合わせて読む図面に説明された遺伝子の少なくとも1種を含む、14種の癌：乳房-、前立腺-、膀胱-、白血病-、リンパ腫-、中枢神経系-、肺-、結腸直腸-、黒色腫、子宮-、腎細胞-、膵-、卵巣-、及び胸膜癌のいずれかひとつの遺伝子プロファイル。

20

【請求項38】

請求項36又は37記載の14種の癌のいずれかひとつの遺伝子プロファイルであり、ここでプロファイルが、そのプロファイル中の少なくとも2種の遺伝子を含む、プロファイル。

【請求項39】

請求項36～38のいずれか1項記載の14種の癌のいずれかひとつの遺伝子プロファイルであり、ここでプロファイルが、そのプロファイル中の少なくとも3種の遺伝子を含む、プロファイル。

【請求項40】

請求項36～39のいずれか1項記載の14種の癌のいずれかひとつの遺伝子プロファイルであり、ここでプロファイルが、そのプロファイル中の少なくとも4種の遺伝子を含む、プロファイル。

30

【請求項41】

本明細書の説明と合わせて読む問題のそれぞれの癌の各々に関する図面に説明された遺伝子の少なくとも1種の遺伝子発現産物に選択的に連結する少なくとも1種の核酸配列を備える、乳房-、前立腺-、膀胱-、白血病-、リンパ腫-、中枢神経系-、肺-、結腸直腸-、黒色腫、子宮-、腎細胞-、膵-、卵巣-及び胸膜癌からなる群より選択される癌を診断するための、診断キット。

【請求項42】

キットが更に、試料中に存在する遺伝子発現産物の量を定量測定することができる少なくとも1種の核酸配列の量を備える、請求項41記載の診断キット。

40

【請求項43】

本明細書の説明と合わせて読む問題のそれぞれの癌の各々に関する図面に説明された遺伝子の少なくとも1種から表された核酸配列から発現されたペプチドに選択的に結合することが可能であるリガンドを含む、乳房-、前立腺-、膀胱-、白血病-、リンパ腫-、中枢神経系-、肺-、結腸直腸-、黒色腫、子宮-、腎細胞-、膵-、卵巣-、及び胸膜癌からなる群より選択される癌を診断するための、診断キット。

【請求項44】

リガンドが、抗体又は免疫分子である、請求項43記載の診断キット。

【請求項45】

50

免疫分子が、fab断片である、請求項44記載の診断キット。

【請求項46】

リガンドが、関連するペプチドに特異的に結合可能である、請求項43～45のいずれか1項記載の診断キット。

【請求項47】

キットが更に、試料中に存在するペプチド量を定量測定することが可能なリガンドの量を備える、請求項43～45のいずれか1項記載の診断キット。

【請求項48】

患者が、乳房-、前立腺-、膀胱-、白血病-、リンパ腫-、中枢神経系-、肺-、結腸直腸-、黒色腫、子宮-、腎細胞-、膵-、卵巣-、及び胸膜癌からなる群より選択される癌に罹患しているかどうかを診断する方法であり：

a)患者から試料を単離する工程；

b)この試料が、本明細書の説明と合わせて読む問題のそれぞれの癌の各々に関する図面に説明された、少なくとも1種の遺伝子の疾患プロファイルと一致する発現レベルを含むかどうかを決定する工程を含む、方法。

【請求項49】

試料は、組織である、請求項48記載の方法。

【請求項50】

組織は、癌であることが疑われる、請求項49記載の方法。

【発明の詳細な説明】

【技術分野】

【0001】

発明の技術分野

本発明は、疾患プロファイリング及び遺伝子発現プロファイリングのための方法論を含むが、これらに限定されるものではない、遺伝子発現データを使用する適応学習システムの医療適用に関する。

【背景技術】

【0002】

発明の背景

癌のような、多くの疾患及び障害は、非常に複雑な遺伝子及び表現型の異常並びに推定不可能な生物学的挙動を有している。癌細胞は、クローン細胞進化、複数の遺伝子突然変異、ゲノム不安定性、及び誤った遺伝子発現の成功した世代の終点を表わしている。

癌の生物学的挙動は、複数の要因、最も重要なことには、個々の癌の生物学的特徴であるが、年齢、性別、人種、遺伝的体質などのような、患者の生物学、並びに癌の位置により決定される。この癌の生物学的及び遺伝的複雑さは、個々の癌は、予想できない臨床経過を辿り、患者の確実でない転帰を伴うことを意味している。

【0003】

特定の癌について複数の治療選択肢が利用可能である場合、患者に関する正確な予後判定が必要であり、その結果治療は、その患者の個々の疾患についてテーラーメイドすることができる。

現在臨床医が利用可能である、癌の分類及び予後評価のための臨床及び情報のツールは、特に個々の患者に適用した場合に、重大な限界がある。個々の患者の特定の癌において、所定の癌に関する臨床及び生物学的情報を統合することは、望ましいであろう。

【0004】

遺伝子発現データは、標準のマイクロアレイデータを用い入手可能である。Affymetrix (商標)により入手可能であるような遺伝子発現マイクロアレイは、大量のデータを提供し、及び罹患した又は異常な組織を、健全な正常組織と比較することにより、患者の特定の疾患又は状態を特徴付けるために使用することができる。しかし得られたデータは、特定の状態又は疾患に関する意味のある情報を得るために処理することが困難である。

この問題点は、患者治療に関する医療適用において特に緊急である。臨床環境において

10

20

30

40

50

患者管理に影響を及ぼすために、臨床の決定支援システムは、高いレベルの信頼性がなければならない。Shippらは、患者の予後の階層化に関する機械学習技術の可能性を精巧に説明したが、それらの方法は、患者の30%を、それらの治療転帰の予測に関して誤分類した。これらは、B-細胞リンパ腫の治癒症例の70%正確な予後判定を達成し、及びこの症例の12%は、実際の致命的転帰とは対照的に治癒と誤って推定された。この精度は、このモデルの臨床適用にとって適切ではない。Alizadehらにより示された同じデータに関するモデルも、臨床には適用することができない。

#### 【0005】

先行技術の機械学習を使用する方法の別の難点は、これらは、新規データが利用可能になるべきモデルをアップデートする容易な手段を提供しないことが多いことである。代わりに、このモデルの完全な再訓練が必要となる。これは集中的コンピュータリソースに  
10  
与するので、これは時間をかかりかつ潜在的に経費がかかる。使用されるシステムは、システムの完全な再訓練を伴わずに、新規データの追加に適合できることが望ましい。

#### 【0006】

本発明者らは、一般的な見解として、進化コネクショニストシステム(Evolving Connectionsist Systems)(ECOS)技術を利用する技術は、従来の統計及びニューラルネットワーク技術と比較した場合に、下記の利点があることがわかった：(i)これらは、それらの訓練に使用されるデータの複雑性を反映している柔軟性のある構造を有する；(ii)これらは、クラスタ化及び分類/予測の両方を行う；(iii)これらのモデルは、古いデータに関して再訓練される必要がなく、新規データに適合することができる；(iv)これらは、試料の異なるサブクラスのルール(プロファイル)を抽出するために使用することができる。これ  
20  
らのルール(プロファイル)は、付加されたいくつかの統計係数によりファジイである。

従って本発明の対象は、遺伝子発現データと1種又は複数の状態又は予後の転帰の間の関係を決定する方法を提供すること、又は少なくとも有用な選択肢を伴う公開を提供することである。

#### 【発明の開示】

#### 【0007】

##### 発明の概要

広範な表現において、本発明のひとつの局面は、遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む入力層；1個又は複数のルール節点を含むルールベース層；及び、遺伝子発現データと1種又は複数の状態の間の関係を表わしているルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備える、ニューラルネットワークモジュールを含む。  
30

#### 【0008】

別の局面において、本発明は、遺伝子発現データと1種又は複数の状態の間の関係を決定する汎用方法を提供し、少なくとも：

- a) 1種又は複数の予め決定された状態のクラスにカテゴリ化された遺伝子発現データのセットを提供する工程；
- b) 該遺伝子発現データ及び該1種又は複数の予め決定された状態のクラスについて、ニューラルネットワークモジュールを訓練する工程であり、ここで、ニューラルネットワークモジュールが、遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の状態のクラスを出力するために構成された1個又は複数の出力節点を含む、出力層；及び、遺伝子発現データと1種又は複数の状態のクラスの間  
40  
の関係を表わすルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備える工程；並びに
- c) ルールベース層からルールを抽出する工程であり、該ルールが、遺伝子発現データと1種又は複数の状態のクラスの間  
50  
の関係を表わす工程を含む。

#### 【0009】

別の局面において、本発明は、遺伝子発現データと1種又は複数の状態の間の関係を決

定するシステムを提供し、これは：

- a) 1種又は複数の予め決定された状態のクラスにカテゴリ化された遺伝子発現データのセットを受け取ることが可能な、入力；
- b) 該遺伝子発現データ及び該1種又は複数の予め決定された状態のクラスについて訓練可能なニューラルネットワークモジュールであり、ここでこのニューラルネットワークモジュールが、遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の状態のクラスを出力するために構成された1個又は複数の出力節点を含む、出力層；及び、遺伝子発現データと1種又は複数の状態のクラスの間の変換関係を表わすルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備える、モジュール；並びに
- c) ルールベース層からルールを抽出する手段であり、該ルールが、遺伝子発現データと1種又は複数の状態のクラスの間の変換関係を表わしている手段を備える。

10

【0010】

別の局面において、本発明は、患者の生物学的試料から抽出された遺伝子発現データが、その遺伝子発現データと1種又は複数の状態のクラスの間の変換関係を表わすルールを満足するかどうかを決定する工程を含み、該ルールが本発明の方法により決定される、患者の状態を診断するための汎用方法を提供する。

別の局面において、本発明は、患者の生物学的試料から抽出された遺伝子発現データが、その遺伝子発現データと1種又は複数の状態のクラスの間の変換関係を表わすルールを満足するかどうかを決定する手段を備え、該ルールが、本発明の方法により決定される、患者の状態を診断する汎用システムを提供する。

20

【0011】

別の局面において、本発明は、1種又は複数の状態に連結された過剰発現された又は過小発現された遺伝子を識別するセットを選択する方法を提供し、この方法は：

- a) 1種又は複数の予め決定された状態にカテゴリ化された遺伝子発現データの1種又は複数のセットを提供する工程；
- b) 該遺伝子発現データ及び1種又は複数の予め決定された状態について、ニューラルネットワークモジュールを訓練する工程であり、このニューラルネットワークが、遺伝子発現データを受け取るために構成された、1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の状態を出力するために構成された1個又は複数の出力節点を含む、出力層；及び、遺伝子発現データと1種又は複数の状態の間の変換関係を表わすルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備え、該適応構成要素が、入力データを基にしたルールベース層中の選択された2個又はそれよりも多いルール節点を集合するように配置されている、工程；
- c) 適応構成要素に、ルールベース層内の選択された2個又はそれよりも多いルール節点を集合させる工程；
- d) ルールベース層からルールを抽出する工程であり、該ルールが、遺伝子発現データと1種又は複数の状態の間の変換関係を表わしている、工程；
- e) 抽出されたルールから、過剰発現された又は過小発現された遺伝子を同定する工程であり、この遺伝子は、1種又は複数の状態に連結された過剰発現された又は過小発現された遺伝子を識別するセットを表わしている工程を含む。

30

40

【0012】

別の局面において、本発明は、1種又は複数の状態に連結された過剰発現された又は過小発現された遺伝子を識別するシステムを提供し、このシステムは：

- a) 1種又は複数の予め決定された状態へカテゴリ化された遺伝子発現データの1種又は複数のセットを受け取ることが可能な、入力；
- b) 該遺伝子発現データ及び1種又は複数の予め決定された状態を受け取るように適合された、ニューラルネットワークモジュールであり、このニューラルネットワークは、遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の状態を出力するために構成さ

50

れた1個又は複数の出力節点を含む、出力層；及び、遺伝子発現データと1種又は複数の状態の間の関係を表わしているルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備え、該適応構成要素が、入力データを基にしたルールベース層中の選択された2個又はそれよりも多いルール節点を集合するように配置されている、モジュール；

c)ルールベース層からルールを抽出するための、抽出構成要素であり、該ルールが、遺伝子発現データと1種又は複数の状態の間の関係を表わしている、要素；並びに

d)抽出されたルールから過剰発現された又は過小発現された遺伝子を同定するための、識別子であり、これらの遺伝子が、1種又は複数の状態に連結された過剰発現された又は過小発現された遺伝子を識別するセットを表わしている識別子を備える。

10

【0013】

好ましくは、この抽出されたルールは、特異的遺伝子発現を特定の状態に連結させる。

【0014】

別の局面において、本発明は、遺伝子発現セットを減少する方法が提供され、この方法は：

a)1種又は複数の状態によりカテゴリ化された遺伝子発現データの1種又は複数のセットを提供する工程；

b)該遺伝子発現データ及び1種又は複数の状態のクラスについてニューラルネットワークモジュールを訓練する工程であり、このニューラルネットワークが、遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の状態を出力するために構成された1個又は複数の出力節点を含む、出力層；及び、遺伝子発現データと1種又は複数の状態の間の関係を表わしているルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備え、該適応構成要素が、入力データを基にしたルールベース層中の選択された2個又はそれよりも多いルール節点を集合するように配置されている、工程；

20

c)適応構成要素に、ルールベース層内の選択された2個又はそれよりも多いルール節点を集合させる工程；

d)ルールベース層からルールを抽出する工程であり、該ルールが、遺伝子発現データと1種又は複数の状態の間の関係を表わしている、工程；

e)抽出されたルールから、遺伝子を同定する工程であり、この遺伝子が、1種又は複数の状態に連結された減少された遺伝子発現セットを表わしている工程を含む。

30

【0015】

別の局面において、本発明は、遺伝子発現セットを減少するシステムを提供し、このシステムは：

a)1種又は複数の状態によりカテゴリ化された遺伝子発現データの1種又は複数のセットを受け取る、入力手段；

b)該遺伝子発現データ及び1種又は複数の状態のクラスについて訓練可能なニューラルネットワークモジュールであり、このニューラルネットワークが、遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の状態を出力するために構成された1個又は複数の出力節点を含む、出力層；及び、遺伝子発現データと1種又は複数の状態の間の関係を表わしているルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備え、該適応構成要素が、入力データを基にしたルールベース層中の選択された2個又はそれよりも多いルール節点を集合するように配置されている、モジュール；

40

d)ルールベース層からルールを抽出するように適合された、ルール抽出手段であり、該ルールが、遺伝子発現データと1種又は複数の状態の間の関係を表わしている、手段；

e)抽出されたルールから遺伝子を同定するように適合された識別子であり、この遺伝子は、1種又は複数の状態に連結された減少された遺伝子発現セットを表わす識別子を備える。

【0016】

50

別の局面において、本発明は、遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の状態を出力するために構成された1個又は複数の出力節点を含む、出力層；及び、遺伝子発現データと1種又は複数の状態の間の関係を表わしているルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備え、該適応構成要素が、遺伝子発現データを基にしたルールベース層中の選択された2個又はそれよりも多いルール節点を集合するように配置されているものを備える、ニューラルネットワークモジュールを提供する。

【0017】

別の局面において、本発明は、本発明の方法に従いニューラルネットワークから抽出された場合の、遺伝子発現データと1種又は複数の状態の間の関係を表わしている、ルールを提供する。 10

【0018】

別の局面において、本発明は、状態を診断するためにニューラルネットワークを訓練する方法を提供し、この方法は：

- a) 1種又は複数の状態によりカテゴリ化された遺伝子発現データを提供する工程；
- b) 遺伝子発現データ及び1種又は複数の状態についてニューラルネットワークモジュールを訓練する工程であり、ここでこのニューラルネットワークモジュールが、遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の状態を出力するために構成された1個又は複数の出力節点を含む、出力層；並びに、遺伝子発現データと1種又は複数の状態の間の関係を表わしているルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備える、工程； 20
- c) リーブワンアウト (leave one out) 法を使用し、データを検定する工程；
- d) 遺伝子発現データの入力を減少し、最良の検定精度を提供する工程；
- e) その入力層のように、減少された遺伝子発現データを受入れるために、ニューラルネットワークモジュールを修飾する工程；
- f) 修飾されたニューラルネットワークモジュールを訓練する工程；
- g) 適応構成要素からルールを抽出する工程；
- h) 任意に減少工程からこの方法を反復する工程を少なくとも含む。 30

【0019】

別の局面において、本発明は、状態を診断するためのニューラルネットワークを訓練するシステムを提供し、このシステムは：

- a) 1種又は複数の状態によりカテゴリ化された遺伝子発現データを提供する、受け取り可能な入力；
- b) 遺伝子発現データ及び1種又は複数の状態について訓練可能なニューラルネットワークモジュールであり、ここでこのニューラルネットワークモジュールが、遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の状態を出力するために構成された1個又は複数の出力節点を含む、出力層；及び、遺伝子発現データと1種又は複数の状態の間の関係を表わしているルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備える、モジュール； 40
- d) 最良の検定精度を提供するために、遺伝子発現データを減少するように操作可能な、入力遺伝子発現データ減少子；
- e) ニューラルネットワークモジュールを、その入力層のように減少された遺伝子発現データを受入れるように修飾するための、ニューラルネットワーク修飾子；
- g) ルールを適応構成要素から抽出するための、ルール抽出手段を備える。

【0020】

好ましくは、前記抽出されたルールは、人間が読取り可能な形で表わされる。好ましい態様において、ニューラルネットワークモジュールは更に、1種又は複数の状態への十分 50

な連結を示さない、ルールベース層内の節点をブルーニングするために配置されているブルーニングアルゴリズムを含む。

【0021】

本発明の別の局面の広範な意味において、遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の予後の転帰を出力するために構成された1個又は複数の出力節点を含む、出力層；及び、遺伝子発現データと1種又は複数の予後の転帰の間の関係を表わしているルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備える、ニューラルネットワークモジュールを含む。

【0022】

別の局面において、本発明は、遺伝子発現データと予後の転帰の間の関係を決定する、汎用方法を提供し、この方法は少なくとも：

- a) 予め決定された予後の転帰により分類された遺伝子発現データのセットを提供する工程；
- b) 該遺伝子発現データ及び予後の転帰について、ニューラルネットワークモジュールを訓練する工程であり、ここでこのニューラルネットワークモジュールが、遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の予後の転帰を出力するために構成された1個又は複数の出力節点を含む、出力層；及び、遺伝子発現データと1種又は複数の予後の転帰の間の関係を表わすルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備える工程；
- c) ルールベース層からルールを抽出する工程であり、該ルールが、遺伝子発現データと1種又は複数の予後の転帰の間の関係を表わす工程を含む。

【0023】

別の局面において、本発明は、該患者の生物学的試料から抽出された遺伝子発現データを、遺伝子発現データと予後の転帰の間の関係を表わすルールと相関するために適合された相関エンジンを備え、該ルールは、本発明の方法により決定される、患者の状態の予後判定をする汎用システムを提供する。

【0024】

別の局面において、本発明は、遺伝子発現データと予後の転帰の間の関係を決定する汎用システムを提供し、この方法は：

- a) 予め決定された予後の転帰により分類された遺伝子発現データのセットを受け取るための、入力；
- b) 該遺伝子発現データ及び予後の転帰について訓練可能なニューラルネットワークモジュールであり、ここでこのニューラルネットワークモジュールが、遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の予後の転帰を出力するために構成された1個又は複数の出力節点を含む、出力層；及び、遺伝子発現データと1種又は複数の予後の転帰の間の関係を表わすルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備える、モジュール；
- c) ルールベース層からルールを抽出するための、ルール抽出子であり、該ルールが、遺伝子発現データと1種又は複数の予後の転帰の間の関係を表わしている抽出子を備える。

【0025】

別の局面において、本発明は、患者の状態の予後判定のための汎用方法であり、該患者の生物学的試料から抽出された遺伝子発現データを、遺伝子発現データと予後の転帰の間の関係を表わしているルールと相関することを含み、該ルールは、本発明の方法により決定される、患者の状態の予後判定する汎用方法を提供する。

【0026】

別の局面において、本発明は、患者の生物学的試料から抽出された遺伝子発現データを、遺伝子発現データと予後の転帰の間の関係を表わしているルールと相関するために適合

10

20

30

40

50

された相関手段を備え、該ルールは、本発明の方法により決定される、患者の状態を予後判定する汎用システムを提供する。

【0027】

別の局面において、本発明は、1種又は複数の予後の転帰に連結された過剰発現又は過小発現された遺伝子を識別するセットを選択する方法を提供し、これは：

a) 1種又は複数の予め決定された予後の転帰へカテゴリ化された遺伝子発現データの1種又は複数のセットを提供する工程；

b) 該遺伝子発現データ及び1種又は複数の予め決定された予後の転帰についてニューラルネットワークモジュールを訓練する工程であり、このニューラルネットワークが、遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の予後の転帰を出力するために構成された1個又は複数の出力節点を含む、出力層；及び、遺伝子発現データと1種又は複数の予後の転帰の間の関係を表わすルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備え、該適応構成要素が、入力データを基にしたルールベース層内の選択された2個又はそれよりも多いルール節点を集合するために配置されている、工程；

c) 適応構成要素に、ルールベース層内の選択された2個又はそれよりも多いルール節点を集合させる工程；

d) ルールベース層からルールを抽出する工程であり、該ルールが、遺伝子発現データと1種又は複数の予後の転帰の間の関係を表わしている工程；

e) 抽出されたルールから、過剰発現された又は過小発現された遺伝子を同定する工程であり、この遺伝子が、1種又は複数の予後の転帰に連結された過剰発現された又は過小発現された遺伝子を識別するセットを表わしている工程を含む。

【0028】

別の局面において、本発明は、1種又は複数の予後の転帰に連結された過剰発現された又は過小発現された遺伝子を識別するセットを選択するシステムを提供し、これは：

a) 1種又は複数の予め決定された予後の転帰へカテゴリ化された遺伝子発現データの1種又は複数のセットを受け取ることが可能な、入力；

b) 該遺伝子発現データ及び1種又は複数の予め決定された予後の転帰について訓練可能な、ニューラルネットワークモジュールであり、このニューラルネットワークは、遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の予後の転帰を出力するために構成された1個又は複数の出力節点を含む、出力層；及び、遺伝子発現データと1種又は複数の予後の転帰の間の関係を表わしているルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備え、該適応構成要素が、入力データを基にしたルールベース層中の選択された2個又はそれよりも多いルール節点を集合するように配置されている、モジュール；

d) ルールベース層からルールを抽出するために適合された、ルール抽出子であり、該ルールが、遺伝子発現データと1種又は複数の予後の転帰の間の関係を表わしている、抽出子；

e) 抽出されたルールから過剰発現された又は過小発現された遺伝子を同定するための、識別子であり、この遺伝子が、1種又は複数の予後の転帰に連結された過剰発現された又は過小発現された遺伝子を識別するセットを表わしている、識別子を備える。

【0029】

別の局面において、本発明は、遺伝子発現セットを減少する方法を提供し、これは：

a) 1種又は複数の予後の転帰によりカテゴリ化された遺伝子発現データの1種又は複数のセットを提供する工程；

b) 該遺伝子発現データ及び1種又は複数の状態のクラスについてニューラルネットワークモジュールを訓練する工程であり、ニューラルネットワークが、遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点

10

20

30

40

50

を含む、ルールベース層；1種又は複数の予後の転帰を出力するために構成された1個又は複数の出力節点を含む、出力層；並びに、遺伝子発現データと1種又は複数の予後の転帰の間の関係を表わしているルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備え、該適応構成要素が、入力データを基にしたルールベース層中の選択された2個又はそれよりも多いルール節点を集合するように配置されている、工程；

c) 適応構成要素に、ルールベース層内の選択された2個又はそれよりも多いルール節点を集合させる工程；

d) ルールベース層からルールを抽出する工程であり、該ルールが、遺伝子発現データと1種又は複数の予後の転帰の間の関係を表わしている、工程；

e) 抽出されたルールから、遺伝子を同定する工程であり、この遺伝子が、1種又は複数の予後の転帰に連結された減少された遺伝子発現セットを表わしている、工程を含む。

【0030】

別の局面において、本発明は、遺伝子発現セットを減少するシステムを提供し、これは：

a) 1種又は複数の予後の転帰によりカテゴリ化された遺伝子発現データの(1種)又は複数のセットを受け取る、入力；

b) 該遺伝子発現データ及び1種又は複数の状態のクラスについて訓練可能なニューラルネットワークモジュールであり、このニューラルネットワークが、遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の予後の転帰を出力するために構成された1個又は複数の出力節点を含む、出力層；及び、遺伝子発現データと1種又は複数の予後の転帰の間の関係を表わしているルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備え、該適応構成要素が、入力データを基にしたルールベース層中の選択された2個又はそれよりも多いルール節点を集合するように配置されている、工程；

d) ルールベース層からルールを抽出するように適合された、ルール抽出子であり、該ルールが、遺伝子発現データと1種又は複数の予後の転帰の間の関係を表わしている、抽出子；

e) 抽出されたルールから遺伝子を同定するように適合された同定手段であり、この遺伝子は、1種又は複数の予後の転帰に連結された減少された遺伝子発現セットを表わす手段を備える。

【0031】

別の局面において、本発明は、遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の予後の転帰を出力するために構成された1個又は複数の出力節点を含む、出力層；及び、遺伝子発現データと1種又は複数の予後の転帰の間の関係を表わしているルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備え、該適応構成要素が、遺伝子発現データを基にしたルールベース層中の選択された2個又はそれよりも多いルール節点を集合するように配置されている、ニューラルネットワークモジュールを提供する。

【0032】

別の局面において、本発明は、本発明の方法でニューラルネットワークから抽出された場合に、遺伝子発現データと1種又は複数の予後の転帰の間の関係を表わしているルールを提供する。

【0033】

別の局面において、本発明は、予後の転帰を提供するニューラルネットワークを訓練する方法を提供し、この方法は、少なくとも：

a) 1種又は複数の予後の転帰によりカテゴリ化された遺伝子発現データを提供する工程；

b) 遺伝子発現データ及び1種又は複数の予後の転帰についてニューラルネットワークモジ

10

20

30

40

50

ュールを訓練する工程であり、ここでこのニューラルネットワークモジュールが、遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の予後の転帰を出力するために構成された1個又は複数の出力節点を含む、出力層；並びに、遺伝子発現データと1種又は複数の予後の転帰の間の関係を表わしているルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備える、工程；

c)リープワンアウト法を使用し、データを検定する工程；

d)遺伝子発現データの入力を減少し、最良の検定精度を提供する工程；

e)その入力層のように、減少された遺伝子発現データを受入れるために、ニューラルネットワークモジュールを修飾する工程；

f)修飾されたニューラルネットワークモジュールを訓練する工程；

g)適応構成要素からルールを抽出する工程、

h)任意に、減少工程からの方法を反復する工程を含む。

#### 【0034】

別の局面において、本発明は、予後の転帰を提供するために、ニューラルネットワークを訓練するシステムを提供し、このシステムは：

a)1種又は複数の予後の転帰によりカテゴリ化された遺伝子発現データを受け取るための、入力；

b)遺伝子発現データ及び1種又は複数の予後の転帰について訓練可能なニューラルネットワークモジュールであり、ここでこのニューラルネットワークモジュールが、遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の予後の転帰を出力するために構成された1個又は複数の出力節点を含む、出力層；及び、遺伝子発現データと1種又は複数の予後の転帰の間の関係を表わしているルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備える、モジュール；

d)最良の検定精度を提供するために、入力遺伝子発現データを減少する、減少子；

e)ニューラルネットワークモジュールを、その入力層のように減少された遺伝子発現データを受入れるように修飾するために適合された、修飾子；

g)ルールを適応構成要素から抽出するための、ルール抽出子を備える。

#### 【0035】

好ましくは、前記抽出されたルールは、人間が読取り可能な形で表わされる。好ましい態様において、ニューラルネットワークモジュールは更に、1種又は複数の予後の転帰への十分な連結を示さない、ルールベース層内の節点をブルーニングするために配置されているブルーニングアルゴリズムを含む。

#### 【0036】

好ましい態様において、抽出されたルールは、乳腺癌、前立腺腺癌、肺腺癌、結腸直腸腺癌、リンパ腫、膀胱移行細胞癌、黒色腫、子宮腺癌、白血病、腎細胞癌、膵臓腺癌、卵巣腺癌、胸膜中皮腫、及び中枢神経系癌のプロファイリングを提供する。

#### 【0037】

別の好ましい態様において、本発明の方法及びシステムは、乳腺癌、前立腺腺癌、肺腺癌、結腸直腸腺癌、リンパ腫、膀胱移行細胞癌、黒色腫、子宮腺癌、白血病、腎細胞癌、膵臓腺癌、卵巣腺癌、胸膜中皮腫、及び中枢神経系癌のプロファイリングに適用される。

別の態様において、抽出されたルールは、DLBCL(びまん性大細胞型Bリンパ腫)のプロファイリングができる。

別の好ましい態様において、本発明の方法及びシステムはDLBCL(びまん性大細胞型Bリンパ腫)に適用される。

#### 【0038】

別の局面において、本発明は、本明細書の説明と合わせて読む図面に説明された、14種の癌：乳房-、前立腺-、膀胱-、白血病-、リンパ腫-、中枢神経系-、肺-、結腸直腸-、黒色腫、子宮-、腎細胞-、膵臓-、卵巣-、及び胸膜の癌のいずれかひとつに関する遺伝子プ

10

20

30

40

50

ロファイリングを提供する。

別の局面において、本発明は、本明細書の説明と合わせて読む図面に説明された遺伝子の少なくとも1種を含む14種の癌：乳房-、前立腺-、膀胱-、白血病-、リンパ腫-、中枢神経系-、肺-、結腸直腸-、黒色腫、子宮-、腎細胞-、膵臓-、卵巣-、及び胸膜の癌のいずれかひとつに関する遺伝子プロファイルを提供する。

好ましくはこのプロファイルは、プロファイル中の少なくとも2種の遺伝子を、より好ましくはプロファイル中の少なくとも3種の遺伝子を、最も好ましくはプロファイル中の少なくとも4種の遺伝子を含む。

#### 【0039】

別の局面において、本発明は、本明細書の説明と合わせて読む問題のそれぞれの癌の各々に関する図面に説明された遺伝子の少なくとも1種の遺伝子発現産物に選択的に連結する少なくとも1種の核酸配列を備える、乳房-、前立腺-、膀胱-、白血球-、リンパ球-、中枢神経系-、肺-、結腸直腸-、黒色腫、子宮-、腎細胞-、膵臓-、卵巣-、及び胸膜の癌からなる群より選択される癌を診断する、診断キットを提供する。

好ましくは、このキットは、試料中に存在する遺伝子発現産物を定量測定することができる少なくとも1種の核酸配列の量も備える。

#### 【0040】

別の局面において、本発明は、本明細書の説明と合わせて読む問題のそれぞれの癌の各々に関する図面に説明された少なくとも1種の遺伝子により表された核酸配列により発現されたペプチドに選択的に結合することが可能であるリガンドを備える、乳房-、前立腺-、膀胱-、白血病-、リンパ腫-、中枢神経系-、肺-、結腸直腸-、黒色腫、子宮-、腎細胞-、膵-、卵巣-、及び胸膜の癌からなる群より選択される癌を診断するための診断キットに関する。

好ましくは、このリガンドは、抗体又は免疫分子、例えばfab断片である。好ましい態様において、このリガンドは、問題のペプチドと特異的に結合することが可能である。好ましくは、このキットは、試料中に存在するペプチド量を定量測定することができるリガンドの量も備える。

#### 【0041】

別の局面において、本発明は、乳房-、前立腺-、膀胱-、白血病-、リンパ腫-、中枢神経系-、肺-、結腸直腸-、黒色腫、子宮-、腎細胞-、膵-、卵巣-、及び胸膜の癌からなる群より選択される癌に患者が罹患しているかどうかを診断する方法に関し、これは：

- a)患者から試料を単離する工程；
- b)試料が、本明細書の説明と合わせて読む問題のそれぞれの癌の各々に関する図面について説明された少なくとも1種の遺伝子の疾患プロファイルに一致する発現レベルを含むかどうかを決定する工程を含む。

好ましくは、試料は、組織であり、より好ましくは癌であることが疑われる組織である。

#### 【発明を実施するための最良の形態】

#### 【0042】

##### 好ましい形の詳細な説明

本明細書において、「生物学的試料」は、少なくとも核酸分子又はポリペプチド分子を含有する試料を意味する。このような試料は、好都合なことに細胞を給源とすることができる。次にこの細胞は、組織から抽出することができる。試料の前処理は、例えば、当該技術分野において周知の技術を用い、生試料から、DNA、RNA及び/又はポリペプチドを抽出することにより行うことができる。別の前処理は、例えば、試料の、DNase、RNase、及びプロテイナーゼを使用する予備消化により行うことができる。

#### 【0043】

「状態」は、癌、先天性疾患及び遺伝疾患からなる群より選択される疾患、障害又は体質である。「状態」は、疾患を引き起こさない体質も含む。「状態の素因」は、状態に対する遺伝的素因も含んでもよい。疾患を引き起こさない体質の非限定的例は、脱毛である

10

20

30

40

50

。

## 【0044】

進化コネクショニストシステム (Evolving connectionist system) ("ECOS") は、進化プロセス及び知識発見のモデル化を促進する、多モジュール式コネクショニストのアーキテクチャである (PCT国際公開公報第01/78003号)。

## 【0045】

ECOSは、多くの進化コネクショニストモジュールからなる。ECOSは下記に従い、時間内に連続操作し、並びにその環境との及び他のシステムとの連続相互作用を介してその構造及び機能性を適合するような、ニューラルネットワークシステムである：(i)システム操作時に変化を施すパラメータPのセット；(ii)未知の分布を伴う、入情報の連続流れ；(iii)本システムの性能を経時的に最適化するために適用される目標(合理的)判定基準(修飾も施される)。

10

進化コネクショニストシステムは、下記の特異的特徴を有する：

(1)これらは、オープンスペースで進化し、固定された次元(dimension)は必要としない。

(2)これらは、オンライン、パターンモード、増分学習、迅速学習し - これはおそらくデータ伝播のひとつのパスによる。

(3)これらは一生の(life-long)学習モードで学習する。

(4)これらは、個別のシステム、及び進化的な集団システムの両方として学習する。

(5)これらは、進化構造を有し、及び構築的学習を使用する。

20

(6)これらは、局所的(locally)及び局所パーティションに問題スペースを学習し、その結果迅速な適合及び経時的進化プロセスの追跡を可能にする。

(7)これらは、ほとんどは記憶ベースでの、統計的及び記号的知識を組合せた、様々な種類の知識を促進する。

## 【0046】

ECOS操作のふたつの個別の相が存在する。最初の学習相の間に、データベクトルが、このシステムへひとつずつそれらの公知の出力値と共に供給される。第二の相(想起(recall))において、新規ベクトルが、このシステムに提示され、これはこれに関する出力値を計算する。

PCT国際公開公報第01/78003 A1号で保護されたECOSには、異なるモデルが存在する。そのひとつである進化ファジィニューラルネットワーク(EFuNN)は、図1に示され、そのアルゴリズムは図2に示されている。EFuNNは、分類タスク及び予測タスクの両方に使用することができる。

30

## 【0047】

「ニューラルネットワークモジュール」は、前記本発明の局面の必要要件を満足するニューラルネットワークのいずれかを意味することができると同時に、ECOSニューラルネットワークの使用が好ましい。特に好ましいのは、PCT国際公開公報第01/78003号(本明細書に参照として組入れられている)に例示されたニューラルネットワークである。国際公開公報第01/78003号のニューラルネットワークを説明するアルゴリズムを以下に説明する：

40

## 【0048】

EFuNN学習アルゴリズム (PT0国際公開公報第01/78003 A1号)。

システムパラメータ：メンバシップ関数の数の初期値を設定する；初期感度閾値(デフォルト $S_j=0.9$ )；誤差閾値E；集計パラメータNagg - 各集計が行われた後の連続例の数；ブルーニングパラメータOLD an Pr；m値(m-of-nモード)；最大ラジウス限界Rmax；ルール抽出のための閾値 $T_1$ 及び $T_2$ 。

最初のルール節点 $r_0$ を、第一例(x,y)を記憶するために設定：

$$W1(r_0) = x_f \text{、及び} W2(r_0) = y_f \text{；} \quad (10)$$

新たな入力-出力対(x,y)のプレゼンテーションを超えたループ

{  $x_f$  間の局所正規化されたファジィ距離D及び現存するルール節点コネクションW1の評価(

50

式(1))

ルール節点層の活性化A1の計算。A1( $r_k$ )  $S_k$ (節点 $r_k$ の感度閾値)であるファジイ入力ベクトル $x_f$ に対し最も近傍のルール節点 $r_k$ (又はm-of-nモードの場合の最も近傍のmルール節点)の発見、

IFそのような節点が存在しないならば、( $x_f$ ,  $y_f$ )について新規ルール節点を作成する、さもなければ

ファジイ出力層 $A2=W2.A1(1-D(W1,x_f))$ の活性化、及び正規化された出力誤差 $Err= |y-y'|/Nout$ の発見。

IF、 $Err > E$ ならば、

現在の例( $x_f$ ,  $y_f$ )を収容するように新規ルール節点を作成、さもなければ

(2)及び(3)に従い、 $W1(r_k)$ 及び $W2(r_k)$ を更新(m-of-nシステムの場合、最高A1活性化を伴うnルール節点を全て更新)。

$N_{agg}$ 例の各群が提示された後の、ルール節点の集計手法の適用。ルール節点 $r_k$ パラメータ $S_k$ 、 $R_k$ 、 $Age(r_k)$ 、 $TA(r_k)$ の値の更新。

必要ならば、ブルーニングパラメータに定義されたように、ルール節点のブルーニング。ルール節点からのルール抽出}

【0049】

本発明の一部の局面に関する特に好ましいニューラルネットワークモジュールの別例は、ECFである。

進化分類機能(Evolving classification function)("ECF")を使用し、データを分類することができる。ECFの各反復の学習配置は、下記の工程で説明される：

1)全てのベクトルが入力されたならば、現在の反復を終了する；さもなければ、そのデータセットからのベクトルを入力し、及びベクトルと既に作成された全てのルール節点の間の距離を計算する；

2)全ての距離が、最大-ラジウスパラメータよりも大きいならば、新規ルール節点を作成され、新規ルール節点の位置は、入力データスペース内の現在のベクトルと同じであり、及びそのラジウスは、最小-ラジウスパラメータに設定され、次に工程1へ進む；さもなければ：

3)現在の入力ベクトルまでの距離がそのラジウスよりも少ない又は等しいルール節点が存在し、及びそのクラスが新規ベクトルのクラスと同じであるならば、何も変更せず、工程1へ進む；さもなければ：

4)入力ベクトルまでの距離がそのラジウスよりも少ない又は等しいルール節点が存在し、及びそのクラスが入力ベクトルのものと異なるならば、その影響フィールド(influence field)は、減少されなければならない、新規フィールドのラジウスは、その距離から最小-ラジウスを減算したのもよりも大きい値に、及び最小ラジウスに設定される；

5)入力ベクトルまでの距離が最大-ラジウスよりも少ない又は等しいルール節点が存在し、そのクラスがベクトルのものと同じであるならば、そのような拡大されたフィールドのみが異なるクラスのいずれか他のルール節点をカバーしない場合に、新規ラジウスとしての距離を採用することにより影響フィールドが拡大し；さもなければ、工程2と同じ方法で新規ルール節点を作成し、工程1へと進む。

【0050】

ECFにおける想起(新規入力ベクトルの分類相)は、下記の方法で行われる：

1)新規入力ベクトルが、ひとつのクラスに連想された1個又は複数のルール節点のフィールド内に置かれているならば、このベクトルはそのクラスに属する；

2)入力ベクトルが、異なるクラスに連想された1個又は複数のルール節点のフィールド内に置かれているならば、このベクトルは、最接近ルール節点に対応しているクラスに属するであろう；

3)入力ベクトルがいずれのフィールド内にも置かれていないならば、次にふたつの場合がある：(1)one-of-nモード：ベクトルは最接近ルール節点に対応するクラスに属する；(2)m-of-nモード：新規ベクトルルール節点によりmが最高に活性化されたmを採用し、ベク

10

20

30

40

50

トルから同じクラスの節点までの平均距離を計算すると；このベクトルは、最小平均距離に対応するクラスに属する。

【 0 0 5 1 】

分類に関する前述のECFは、使用したデータセットに応じて最適化されることが必要であるいくつかのパラメータを有する。これらは：

1)最大ラジウス、2)最小ラジウス、3)メンバシップ関数(mf)の数、4)m-of-n値、5)学習相の間のデータプレゼンテーションのための反復数。

これらのパラメータは、進化コンピューティング法、又は他の統計法の使用により最適化することができる。

【 0 0 5 2 】

ひとつの態様において、本発明は、遺伝子減少、モデル作成及びプロファイリングのための適応方法を使用する。この方法論はECOSを使用する。この方法論は、下記の主要相で構成される：

(1)入データに関する進化コネクショニストシステム(ECOS)の連続的訓練、その結果全ての利用可能なデータを収容する「マザー」システムの作成；

(2)「マザー」システムからの出力クラスに関連した特徴(遺伝子)の抽出；

(3)選択された特徴及び出力クラスをベースにしたモデルの作成。

(4)プロファイル(ルール)の抽出。

【 0 0 5 3 】

全ての入手可能なデータに関する教師あり訓練により作成されたECOSからのルールの抽出を通して、ECOSの適応構成要素における特徴選択が行われることが好ましい。ECOS訓練パラメータは、分類誤差が最小化され及びECOSがデータに存在する最も近い特徴をモデル化するように、最適化される。ECOSの隠れ層内の各節点は、類似試料のクラスタの中心を提示し、及びルールとして意味的に表現することができる。各ルールは、そのデータセットからの特定のクラスに属する1種又は複数の試料に関する入力特徴レベルのパターンに関連している。どのルールが、EFuNNから抽出された場合のように見えるかの例を以下に示す：

IF VAR1はLOW (0.80)及び

VAR3はHIGH (0.76)及び

VAR12はHIGH (0.91)及び

VAR25はLOW (0.80)及び

VAR31はLOW (0.87)及び . . . . .

THEN 、CLASS-Zが、VERY LIKELY(メンバシップ値0.92を伴う)。このルールにおいて収容された訓練例(ATE)は、50中10であり、このルールに関するクラスタのラジウスは0.15である。

【 0 0 5 4 】

その後、クラス間の識別において有意である変量のセットを同定するために、ルールを解析する。これは、Ramaswamiら；Dudoitらの論文に使用された標準のシグナル-ノイズ比法を用い、各クラスに関するルールのその重要性に従い各変量をランク付けすることにより実現される。

この方法を用い、各入力変量を、各クラスの-1~0の間に値に割当てて。各クラスについて、それらのランク値がセット閾値を超えた場合には、その後変量が選択される。この値は、入力特徴の最適セットを選択するために、変更される。1種よりも多いクラスにより選択されるいずれかの変数のみが、特徴(遺伝子)セットに一旦含まれ、その結果減少した遺伝子セットを示す。一旦特徴選択(遺伝子選択)相が完了したならば、当初のデータセットが、この特徴セットに存在しない特徴を除去することにより最小化される。次にこの新規データを用い、新規ECOSを訓練する。減少された特徴スペースにより、訓練時間は有意に短縮されるであろう。分類誤差が最小化され及びモデルの汎化能が最大化されるように、ECOSの性能は評価され、及び訓練パラメータは修飾されなければならない。

【 0 0 5 5 】

10

20

30

40

50

本発明は、複雑な遺伝子発現データを統合及び解釈することができ、並びに患者の診断評価、予後の評価及び臨床の管理に適合することができる、情報システムを提供する。この方法は、入力-出力データ対からのパターンを学習し、次に構造からのルールを抽出するために、教師有り学習ニューラルネットワークの学習能を使用する、知識-ベースのニューロ-コンピューティングをベースにしている。

一旦特定の疾患に関するプロファイルが抽出されたならば、このプロファイルに使用された遺伝子は、精度評価のための全ての入手可能なデータに関するリーブワンアウト法(Dudoitら、Shippらの論文参照)において訓練され及び検定された新規ECOSモデルの入力変量として使用することができる。

本発明者らは、国際公開公報第01/78003号に開示されたECOSが、遺伝子発現データを含む、様々な情報源をベースにした複雑な疾患プロファイリングに特に適していることを発見した。

#### 【0056】

ひとつの形において、本発明は、疾患分類、治療の予後の転帰、及び薬物反応の予後判定のための新規方法の開発に適用することができる。疾患プロファイルを、前述の方法を通して抽出することもできる。このネットワークから抽出されたプロファイルは、これらのルールを構成する遺伝子の機能の決定により解釈することができる、腫瘍亜群を定義することができる。これらの遺伝子は、腫瘍形成に寄与するか、又は腫瘍形成の結果である過剰発現された遺伝子及び過小発現された遺伝子に対応するであろう。例えばシグナル伝達経路もしくは細胞接着複合体のような、同じ機能単位に寄与する異なる遺伝子、又は例えば同じ表現型の転帰につながる平行シグナル伝達経路のような、重複する機能単位からの異なる遺伝子を含む、様々なルールが抽出されるであろう。

#### 【0057】

これらの一般的機能単位は、アポトーシスの回避、増殖シグナル伝達、血管新生、組織侵襲/転移及び複製能を含むことができる。

従ってこれらのルールにおいて遺伝子を同定及び機能するための解析は、腫瘍において決定的に混乱された機能単位及びこれらの混乱された機能の結果を同定するために推定されるであろう。

#### 【0058】

これらのルールを形成する遺伝子は、腫瘍組織-特異的パターンで発現された転写産物にほぼ対応する。これらは、例えば、インスリノーマのインスリン又は特定の分化マーカーのような、腫瘍組織において過剰発現されたが、疾患進行には寄与しない組織-特異的遺伝子であるか、もしくは腫瘍形成において役割を果たすが、腫瘍特異的パターンではこれらの遺伝子は先に同定されたもののサブセットであるような過剰発現及び過小発現された遺伝子であることができ、すなわちこれらの遺伝子は、複数の癌に共通の機能単位に寄与する遺伝子を除外するであろう。

#### 【0059】

本発明のひとつの形において、入力データは、生の遺伝子発現データを含む。この入力データは、最初に適当なフィルタリング、正規化又は対数変換プロセスを用い、先に説明されたようにフィルタがかけられる。標準の統計検定、例えば、T検定又は他の相関関係は、この予備処理したデータ、及び出力クラスを最良に識別するこのデータから得られた遺伝子に適用することができる。N<sub>1</sub>遺伝子は、N<sub>1</sub>は選択閾値を超えるN未満であるようなNの初期セットから選択することができる。

#### 【0060】

入力データは、前述のもののような、N選択された遺伝子の個々の発現値のような、N<sub>1</sub>遺伝子入力を含むことができる。これらの出力は、例えば、14種の異なる型の癌であることができる。図2を参照し先に説明されたように、ひとつ又は複数のこのモデルは、この入力及び出力データについて訓練されるであろう。その後このモデルは、N<sub>1</sub>の特徴の減少された入力データのセットについて検定される。

#### 【0061】

10

20

30

40

50

訓練後、各特徴、例えば高値 $G_h$ を定義するメンバシップ関数に属する遺伝子は、計算された値 $G_h$ 、 $l$ を基に、下記のように、各クラス $CI$ についてランク付けされる：

$$G_{h,l} = (G_h, m_i - G_h, m_a) (G_h, m_r - G_h, m_a)$$

$G_h, m_i$ は、遺伝子 $G$ がそのクラス $CI$ を支持する訓練されたネットワークからの全てのルール節点について高い場合の平均値であり、 $G_h, m_a$ は、遺伝子 $G$ が出力クラスを超えて訓練されたネットワークからの全てのルール節点について高い場合の平均値であり、並びに $G_h, m_r$ は、遺伝子 $G$ がそのクラス $CI$ 以外のクラスを支持する訓練されたネットワークからの全てのルール節点について高い場合の平均値である。

【0062】

閾値 $G_{thr}$ は適用され、及び特徴の新規セットが、この遺伝子の場合、閾値を上回る少なくとも1種のメンバシップ値を有する全ての遺伝子を含む各クラスについて選択される。全てのクラスの特徴は、 $N_2$ 遺伝子の新規特徴セットを形成することに組合せられている。

【0063】

その後ルールがこのモデルから抽出され、各ルールは、 $N$ -次元の遺伝子発現スペースからの群又はクラスタのプロファイルを表している。

場合によっては、入手可能な新規入力データがあり、並びにECOSの構造は、新規データを追加するための手段を提供し、その結果このモデルは新規入力データにより更新していき、向上した精度を得る可能性がある。

【0064】

#### 初期診断適用

例えば、血液、尿、腹膜洗浄液及び糞便抽出物を含む体液中の特異抗原の同定は、疾患及び他の状態を初期診断し、初期治療及び改善された予後判定につながる、価値のある方法を提供することができる。特異抗原は、疾患又は状態進行をモニタリングし、定められたコースを辿る手術、放射線療法及び化学療法の有効性に可能性をもたらす手段も提供することができる。しかし多くの疾患及び状態に関して、入手可能なマーカーは、不十分な感度及び特異性に悩まされている。

【0065】

多くの疾患及び状態において、蛋白質は、悪性疾患を伴わない個人に比べ、上昇したレベルで体液中に存在することができ、並びに免疫検出を可能にするには不十分な安定性であり得る。体液は、血液、尿、痰、精液、胃液及び糞便を含む。このような体液が有用でない場合、疑わしい組織の生検を使用してもよい。過剰発現又は過小発現は、これらが、疑わしい組織から放出された細胞、細胞溶解液を含む場合に、液体中の核酸検出又は蛋白質検出技術のいずれかによっても検出することができる。

【0066】

関心のある蛋白質は、体液中に検出することもできる。例えば、関心のある全蛋白質ペプチドのいずれかに対するモノクローナル又はポリクローナル抗体を使用する免疫検出法である。関心のあるペプチドは、合成であるか又は*in vitro*もしくは*in vivo*システムにおいて発現されることができる。免疫検出技術は、ELISA/EIA、ラジオイムノアッセイ、比濁法、免疫濁度アッセイ、化学発光、免疫蛍光(顕微鏡又はフローサイトメトリーによる)、免疫組織化学及びウェスタンブロットを含むが、これらに限定されるものではない。その他の蛋白質検出法を使用することができることは、容易に理解することができる。

【0067】

#### キット

本発明の発見を基に、いくつかの種類の実験キットを作成することができる。第一に、検出分子(又は「捕獲試薬」)を予め搭載した検出器具を有するキットを、作成することができる。 $mRNA$ 検出の態様において、このような器具は、その上のオリゴヌクレオチドが、 $mRNA$ にハイブリダイズする捕獲試薬として検出される基板(例えば、ガラス、シリコン、石英、金属などの)を備えることができる。一部の態様において、 $mRNA$ の直接検出は、 $mRNA$ ( $cy3$ 、 $cy5$ 、放射標識又は他の標識物で標識された)が、基板上のオリゴヌクレオチドに

10

20

30

40

50

ハイブリダイズすることにより実現することができる。別の態様において、mRNAの検出は、最初に所望のmRNAに対する相補的DNA(cDNA)の作成により実現することができる。その後標識したcDNAを、基板上的オリゴヌクレオチドにハイブリダイズし、検出することができる。

使用した検出法に関わらず、被験発現の発現の標準測定との比較が望ましい。例えば、RNA発現は、全細胞DNA、構成的に発現されたRNA(例えば、リボソームRNA)の発現又は他の相対的に一定のマーカーに対し標準化することができる。

#### 【0068】

抗体も、捕獲試薬としてキットにおいて使用することができる。一部の態様において、基板(例えば、マルチウェルプレート)は、それに付着された特異的捕獲試薬を有することができる。一部の態様において、キットは、含まれた阻止試薬を有することができる。阻止試薬を用い、非特異的結合を減らすことができる。例えば、非特異的オリゴヌクレオチド結合は、サケ精子DNAのような検出のためのオリゴヌクレオチドを含まない、いずれか都合の良い給源由来の過剰なDNAを用い減少することができる。非特異的抗体結合は、血清アルブミンのような、過剰な阻止蛋白質を用い減少することができる。オリゴヌクレオチド及び蛋白質を検出する豊富な方法が当該技術分野において公知であり、並びに会合した分子を特異的に検出することができるいずれかの戦略を使用し及び本発明の範囲内と見なすことができることも理解することができる。

10

#### 【0069】

抗体検出に頼った態様において、関心のある蛋白質又はペプチドは、1個の細胞を基に、又は全細胞、組織もしくは液体蛋白質、液体容量、組織質量(重量)を基に、表わすことができる。加えて、血清中のXXXは、アルブミンのような、比較的高量の血清蛋白質を基に表わすことができる。

20

試験キットは、基板に加え、捕獲試薬(例えば、プローブ)、洗浄液(例えば、SSC、他の塩、緩衝液、界面活性剤など)、更には検出部分(例えば、cy3、cy5、放射性標識など)を備えることができる。キットは、使用説明書及び包装も含むことができる。

#### 【0070】

##### プロファイル中の蛋白質に対する抗体の作成法

これらの蛋白質に対する抗体を作成するために、当該技術分野において公知の方法のいずれかを使用することができる。例えば、ポリクローナル抗血清を、単離された蛋白質、ペプチド又は蛋白質及びペプチドの混合物の、適当な動物、例えばウサギへの注射により作成することができる。一部の態様において、アジュバントを使用し、免疫応答を増強することができる。適当な追加注射の後、血清を収集し、血清として使用するか、又は代わりの態様において、IgG画分を血清から作成することができる。

30

#### 【0071】

代わりの態様において、モノクローナル抗体を、当該技術分野において公知の常法を用い、蛋白質又はペプチドに対して作成することができる。簡単に述べると、単離された蛋白質又はペプチド調製物が、適当な動物(例えば、マウス)に注射され、及び免疫応答が惹起される。この動物の脾臓を摘出し、及び脾細胞を、黒色腫細胞に融合し、ハイブリドーマを作成する。標的蛋白質又はペプチドに対する抗体を産生するハイブリドーマを選択することができる。及び細胞培養物を増殖し、所望の量のモノクローナル抗体を作成する。抗体は更に、望ましい親和性について選択することができる。及び望ましい増殖特性及び抗体産生を有する細胞株が選択される。

40

#### 【0072】

ペプチドは、関心のある組織(例えば、胃の腫瘍組織)から単離した無傷の蛋白質から作成するか、又は化学的に合成することができる。他の態様において、蛋白質は、完全長標的蛋白質に特異的なcDNAから、又は標的蛋白質の一部をコードしている(インフレーム)オリゴヌクレオチドから、組換え法を用い作成することができる。

この抗体キットは、このプロファイルを特徴付ける最低数の蛋白質に対して生じた抗体からなるであろう。同じくアッセイを行うための全ての試薬も含む。

50

## 【0073】

RNAベースのキット

検出分子(又は「捕獲試薬」)を予め搭載した検出器具を有するキットを作成することができる。遺伝的プロファイル内の遺伝子からmRNAを検出する態様において、このような器具は、その上のオリゴヌクレオチドが、mRNAにハイブリダイズする捕獲試薬として検出される基板(例えば、ガラス、シリコン、石英、金属などの)を備えることができる。一部の態様において、mRNAの直接定量検出は、mRNA(cy3、cy5、放射性標識又は他の標識物で標識された)の基板上のオリゴヌクレオチドへのハイブリダイズにより実現することができる。別の態様において、mRNAの検出は、所望のmRNAに対する相補的DNA(cDNA)の最初の作成により実現することができる。その後、標識されたcDNAは、基板上のオリゴヌクレオチドにハイブリダイズし、及び検出することができる。cDNAは、Applied Biosystems社のTaqman(商標)手法を用い検出することもでき、これは、それが実時間PCRアッセイにおいて増幅されるにつれて、特異的蛍光標識されたオリゴヌクレオチドプローブが、標的cDNAにハイブリダイズされる。分子ビーコン及びPCR産物の SyBrグリーン標識を含むが、これらに限定されるものではない、他の定量的検出法も使用することができる。

10

## 【0074】

本明細書において有用なひとつの技術において、関連したRNAを、本発明の方法で同定された単独又は複数の遺伝子に保存された配列に特異的なオリゴヌクレオチドプライマーを用いる、RT-PCRにより検出することができる。総RNAを、標準技術により体液から抽出し、逆転写酵素を用いcDNAに転換し、及びその後PCRを用い増幅することができる。望ましいPCRプライマーは、イントロン性DNAに隣接し、ゲノムDNAのPCR増幅を防止する。結果は、例えば、ゲル電気泳動又は定量的実時間PCR技術などの、常法を用い検出されるであろう。実時間PCRは、蛍光-標識したTaqman(商標)プローブ又はPCR産物のSybrグリーンのような蛍光色素への直接結合のいずれかを用い行うことができる。しかし、関心のある核酸を検出する他の方法を使用することができることは、理解することができる。

20

試験キットは、基板に加えて、捕獲試薬(例えばプローブ)、洗浄液(例えば、SSC、他の塩、緩衝液、界面活性剤など)、更には検出部分(例えば、cy3、cy5、放射性標識など)を備えることができる。キットは、使用説明書及び包装も備えることができる。

## 【0075】

薬物標的バリデーション

一旦遺伝子セットが本発明の方法により同定されたならば、個々の遺伝子は、薬物標的として開発され、その後以下に説明されたようにバリデーションされる。薬物標的は、典型的には罹患した又は機能不全の細胞内で過剰発現された蛋白質である。その後薬物候補が、典型的にはこの薬物標的に対して製造される。モノクローナル抗体は、この蛋白質及びこの蛋白質が結合する合成的化学物質に対して作成される。従って本発明は、薬物標的を同定する方法及び薬物候補に拡大される。

30

## 【0076】

薬物標的の存在を決定するひとつの方法は、本発明の方法により、状態又は疾患に關与していると同定された遺伝子を、誘導性プロモーターにより、癌細胞株に安定してトランスフェクションすることである。その後この遺伝子により表された蛋白質が発現され、及びこの細胞株の生存度に対する過剰発現の作用が決定される。この細胞株における遺伝子の発現は、小型干渉性RNA(siRNA)又はこの蛋白質に対して生じた抗体のいずれかを用い、ノックアウトされる。細胞生存度に対する作用が決定される。問題の遺伝子が、細胞に対して、細胞殺傷又は正常型への細胞に回復を引き起こすかのいずれかにより、所望の作用を有するならば、これは、薬物標的である。

40

## 【0077】

疾患管理マーカー

本発明の方法は、疾患管理マーカーの作成を可能にする。このようなマーカーは、医師が、疾患が存在するかどうか、及び存在するのであればどのように治療に反応するかを決定することを可能にする。特異的疾患特性に反応する本発明の方法により同定されたプロ

50

ファイルは、以下のような疾患管理マーカーとして開発されるであろう：

疾患管理マーカーのための検出キットは、典型的にはEfuNN又はRNAレベルでの発現の試験が可能で、キットにより同定された遺伝子プロファイルに対して生じた抗体のパネルからなる。

#### 【0078】

抗体キットは、特定疾患又は状態を伴う患者において過剰又は過小発現された本発明の方法において同定された遺伝子を用いて開発することができる。同定された遺伝子によりコードされた蛋白質に対する抗体又は他の免疫分子は、当該技術分野における標準技術により作成される。罹患した又は異常な組織の生検試料が、患者から採取され、その後罹患した組織において過剰又は過小発現された蛋白質レベルに関する、ELISA/EIA、ラジオイムノアッセイ、比濁法、免疫濁度測定アッセイ、化学発光、免疫蛍光(顕微鏡又はフローサイトメトリーによる)、免疫組織化学及びウェスタンブロットを含むが、これらに限定されるものではない免疫検出技術により、抗体又は免疫分子を用い分析される。異常なレベルの問題の蛋白質の存在は、特異的疾患又は状態の存在の指標である。

10

同様の技術を用い、特異的疾患又は状態を伴う患者をモニタリングし、治療が望ましい作用を有するかどうかを決定することができる。

#### 【0079】

別の態様において、RNAベースのアッセイを利用してもよい。総RNA又はmRNAは、標準技術を用い、生検試料から抽出され、その後特異的疾患又は状態において過剰又は過小発現された本発明の方法により同定された、遺伝子の存在の指標である1種又は複数のアンチセンスマーカーにRNAを連結し、及び結合したRNAレベルを決定することにより、標準核酸キットにおいて適用されるであろう。試料中の結合したRNAレベルは、特異的疾患又は状態の非存在又は存在の指標であろう。この技術を用い、特異的疾患又は状態を伴う患者をモニタリングし、治療が望ましい作用を有するかどうかを決定することができる。

20

#### 【0080】

##### プロファイリングにおける蛋白質に対する抗体の作成法

これらの蛋白質に対する抗体を作成するために、当該技術分野において公知の方法を使用することができる。例えば、ポリクローナル抗血清を、蛋白質、ペプチド、又は蛋白質及びペプチドの混合物の、ウサギのような適当な動物への注射により作成することができる。一部の態様において、免疫応答を増強するためにアジュバントを使用することができる。適当なブースター注射後、血清を収集し、及び血清として使用するか、又は別の態様において、IgG画分を血清から作成することができるかのいずれかである。

30

#### 【0081】

別の態様において、当該技術分野において公知の標準方法を用い、モノクローナル抗体を、これらの蛋白質又はペプチドに対して作成することができる。簡単に述べると、単離された蛋白質又はペプチド調製物を、適当な動物(例えば、マウス)に注射し、及び免疫応答を惹起する。この動物の脾臓を摘出し、及び脾細胞を、黒色腫細胞と融合し、ハイブリドーマを作成することができる。標的蛋白質又はペプチドに対する抗体を産生するハイブリドーマが選択され、及び細胞培養物が増殖され、所望の量のモノクローナル抗体を作成する。更に望ましい親和性を有する抗体が選択され、及び望ましい増殖特性及び抗体産生を有する細胞株が選択される。

40

#### 【0082】

ペプチドは、関心のある組織(例えば、胃の腫瘍組織)から単離された無傷の蛋白質から作成することができるか、もしくは化学的に合成することができる。他の態様において、蛋白質は、完全長標的蛋白質に特異的なcDNAから又は標的蛋白質の一部をコードしている(インフレーム)オリゴヌクレオチドから組換え法により作成することができる。

この抗体キットは、このプロファイルを特徴とする最低数の蛋白質に対して生じた抗体からなるであろう。同じく、このアッセイを実施するための全ての試薬も含む。

#### 【0083】

##### RNAベースのキット

50

検出分子(又は「捕獲試薬」)が予め搭載された検出器具を有するキットを製造することができる。遺伝的プロファイル内の遺伝子からのmRNAを検出する態様において、このような器具は、その上にmRNAとハイブリダイズする捕獲試薬としてオリゴヌクレオチドが存在する基板(例えば、ガラス、シリコン、石英、金属など)を備えることができる。一部の態様において、mRNAの直接定量検出は、mRNA(cy3、cy5、放射性標識又は他の標識により標識された)の基板上的オリゴヌクレオチドへのハイブリダイズにより実現することができる。別の態様において、mRNAの検出は、所望のmRNAに対する相補的DNA(cDNA)の最初の作成により実現することができる。次に、標識されたcDNAを、基板上的オリゴヌクレオチドにハイブリダイズし、検出することができる。

【0084】

10

cDNAはApplied Biosystems Tagman(商標)法を用いて検出することもでき、これは、それが実時間PCRアッセイにおいて増幅されるにつれて、特異的蛍光-標識されたオリゴヌクレオチドプローブが標的cDNAにハイブリダイズされる。分子ビーコン及びPCR産物のSyBrグリーン標識を含むが、これらに限定されるものではない他の定量検出法も使用することができる。

【0085】

試験キットは、基板に加え、捕獲試薬(例えば、プローブ)、洗浄液(例えば、SSC、他の塩類、緩衝液、界面活性剤など)、更には検出部分(例えば、cy3、cy5、放射性標識など)を備えることができる。キットは、使用説明書及び包装も備えることができる。

本発明は、限定ではない実施例を参照し、以下に説明する：

20

【実施例】

【0086】

#### 実施例1

図1は、ニューラルネットワークモジュール22を示す。好ましい構造は、ファジイルールを実行するコネクショニスト構造であるファジイニューラルネットワークである。ニューラルネットワークモジュール22は、入力データを受け取るように配置されている1個又は複数の入力節点42を有する、入力層40を備える。この入力データは、単数又は複数のニューラルネットワークモジュールが指示された特定の適用によって左右されるであろう。

ニューラルネットワークモジュールは、1個又は複数の出力節点58を有する出力層56も備える。出力節点は、出力変数の実数値を示す。

30

入力節点及び出力節点は、このシステムから検索される情報の種類に応じて構成される。疾患の遺伝子発現プロファイリング及び分類に關与するひとつの適用において、入力データは、遺伝子発現データのような、關連する変数のセットであり、及び出力変数は、時間的尺度にわたる疾患カテゴリ又は予後の転帰である。

【0087】

疾患予後のプロファイル開発及び時間的尺度にわたる疾患予後判定の別の構成において、入力データは、選択された遺伝子の個別の発現値を含み、及び出力データは、様々な種類の疾患、例えば様々な種類の癌を含む。

薬物反応予後判定及びプロファイル開発に關係する更なる構成において、入力は、特定の治療様式が与えられた特定の患者群に關する遺伝子発現データとして構成され、及び出力データは予後判定である。

40

【0088】

ニューラルネットワークモジュール22は更に1個又は複数のファジイ入力節点46を有するファジイ入力層44を備えてもよい。このファジイ入力節点46は、このシステムの更なる使用のために、データを入力節点42から変換する。ファジイ入力節点46の各々は、例えば三角型メンバシップ関数、ガウス関数又はいずれか他の目的に適した公知の関数のような、それに付屬した様々なメンバシップ関数を有する。ファジイ入力節点46の主要目的は、入力値を、入力節点42から、その値がメンバシップ関数に屬するメンバシップ値へと変換することである。

【0089】

50

ニューラルネットワークモジュール22は更に、1個又は複数のファジイ出力節54を有するファジイ出力層42を備える。各ファジイ節54は、ファジイ入力層54のファジイ入力節46に類似した、出力変量のファジイ定量化を表わしている。好ましくは、重み付け総和入力関数及び飽和型線形活性化関数(saturated linear activation function)をこれらの節に用い、存在した入力ベクトルに連想された出力ベクトルが出力メンバシップ関数の各々に属するメンバシップ値を算出する。

ニューラルネットワークモジュール22は、1個又は複数の記憶節62を有する短期記憶層60を備えてもよい。短期記憶層60の目的は、入力データの構造的時間関係を記憶することである。この短期記憶層は、ルールベース層48から情報を受け取り及びこの層へ情報を送信するように配置されていることが好ましい。

10

#### 【0090】

国際公開公報第01/78003号により詳細に述べられているように、各ルール節50は、ファジイ入力スペースからのハイパー球面(hyper sphere)と、ファジイ出力スペースからのハイパー球面の間の連想を表わしている。各ルール節 $r_j$ は、感度閾値パラメータ $S_j$ を有し、これは新たな例からの新規入力ベクトル $x$ 又はこのルール節に連想されるとみなされる例に関する順位の入力 $(x, y)$ に対するルール節 $r_j$ の最小活性化閾値を定義している。新規入力ベクトル $x$ は、 $x$ が、最小活性化閾値を満たし及びその後このルール節と連想されると見なされる場合には、ルール節を活性化する。入力ハイパー球面のラジウスは、 $R_j = 1 - S_j$ と定義され、 $S_j$ は感度閾値パラメータである。

#### 【0091】

ファジイロジックルール及び知識の他の型は、容易にアクセス可能な形で訓練されたニューラルネットワーク又はニューラルネットワークの組合せから抽出することができる。訓練ネットワークから抽出されたファジイルールの一例は：

20

R1: IF [遺伝子1が度数0.9までHigh]and[遺伝子3が度数0.9までHigh]and[遺伝子8が度数0.8までLow]、{受取りフィールドのラジウス = 0.109}ならば、THEN疾患A{収容された訓練例 = 24中15}

#### 【0092】

これらの型のルールは、入力データと出力データの間関係を表わしている。これらは、そこから、疾患の分類又は予後のプロセスに関する知識を得ることができるプロファイルを提供する。これらのルールは、例えば様々な種類の癌のような、特定疾患と強力に関連している遺伝子プロファイル及び臨床情報を指摘し、並びに新規試験及び治療の開発に使用することができる。

30

#### 【0093】

##### 実施例2

図2は、疾患プロファイリングのひとつの方法を図示している。

入力データは、任意にそれについて行われる前処理100を受ける。このような前処理は、フィルタリング、相関評価、正規化、対数変換及びノイズ低下を含む。

入力データ及び望ましい出力データを用い、このモデルを訓練する110。この訓練は、教師あり学習アルゴリズム、例えば、国際公開公報第01/78003号に開示されたような教師あり学習アルゴリズムのひとつにより行われる。

40

#### 【0094】

このモデルは、減少された入力データについて検定される120。この例において、入力データは、198種の腫瘍試料を含み、及び個別のモデルが、各腫瘍試料について訓練され、198種のモデルを生じる。各モデルは、減少された入力データセット、例えば198種の試料中の197種について訓練され、その後減少された入力データセットに含まれない試料について検定される。その後誤差が評価され、平均分類誤差が算出される。

#### 【0095】

表1は、減少された入力データに関する検定から得た試料結果を並べている。これは、14クラスの癌データに関する遺伝子とクラスの間関係の評価するために、相関統計解析を用い、16,036個のマイクロアレイ遺伝子発現データから選択された139遺伝子について

50

訓練された、異なるパラメータ値を有するふたつの異なるニューラルネットワークモデルを示している。この結果は、本発明のニューラルネットワークは、特定の状態に連結した遺伝子を減少していく上で有効であることを示している。この結果は、癌転帰及び癌型の推定において、高度の精度が得られることも示している。

再度図2を参照し、減少された入力データに関する単独又は複数のモデルの検定後、最良の精度をもたらす減少された入力データセットが、最終分類システム開発及びプロフィール抽出のために選択される。

【0096】

【表1】

10

表1

メンバーシップ機能MFの番号	Errthr-誤り閾値	例の各数の後の凝集	活性化機能； 1-線形； 2-RBF	ルールの平均数	Leave-one-out法の制度
2	0.9	-	1	37.9	79.3
1	0.9	200	2	48.8	81.9

【0097】

20

次にこのモデルは、より詳細には様々な適用について以下に説明されるように、選択された減少された入力データを基に訓練される130。次に様々なルールを、このモデルから抽出してもよい140。新規入力データが利用可能であるならば150、この新規入力データを用い、単独又は複数のモデルを更に訓練することができる。

【0098】

図4は、60遺伝子210にわたる38ルール200のセットを図示している。癌のクラスを表わしている各試料群は、少なくとも1種のルールにより、及び多くの場合は複数のルールにより表わされ、各ルールは、別個のクラスターの試料を表わしている。ルールは、入力特徴、例えばクラスタについて特徴的である遺伝子プロフィールを表わしている。

同じクラスの全てのクラスタ間の共通パターンを見つけるために、各クラスについて単独のプロファイルを生じる集計手法が適用され、これはそれらのいずれかのメンバーシップ値に関する閾値よりもより高い値を有する遺伝子を選択するために、使用された閾値 $C_{thr}$ により左右される。これらのより高い値は、高い発現値又は低い発現値を含むであろう。

30

クラスCIのプロファイルは、 $G_n$ が閾値 $C_{thr}$ を超えるクラスCIを支持する全てのルール節点において出現する場合は、遺伝子Gは高 $G_n$ であり、 $G_n, l$ 値を伴ない、並びに $G_n, l$ 値は、これらのルール節点を超える全ての $G_n$ 値の平均として算出される。

【0099】

Shippらと同じ遺伝子発現データであるが、患者に関して入手可能である追加の臨床情報 - 生存期間及びIPI(国際予後指標)数値 - を伴うものの、モデリング、予後判定及びルール抽出に関する、適応、学習、進化コネクショニストシステム(ECOS)を基にした技術及び方法論が、以下に説明される。ふたつの例が示される。

40

【0100】

第一例において、Shippらと同じ11遺伝子セットを用い、予後モデルを進化させる。58試料(32治愈、及び26致命的)の全セットについてクロスバリデーション(リーブワンアウト法)を用い、90%予測率(各々、93%対14%)を得た。この実験は、下記のパラメータ値を有する進化ファジィニューラルネットワーク(EFuNN)を利用する：入力変数に関する3メンバーシップ関数、及び出力変数に関する非ファジィ表示；最大ラジウス=1；訓練の1回反復。

【0101】

このモデルをデータから進化させた後、サブクラスもしくはプロトタイプを形成するデ

50

ータの各クラスを次に、臨床情報 - ファジイ表示により、月数で表わされた生存期間により、マッピングする - 短期(1、12ヶ月未満)、中期(2、20~40ヶ月)、長期(40~60ヶ月)。これらのプロトタイプの各々を特徴付ける遺伝子発現のルール(プロファイル)が、抽出される。これらのプロファイルは、各転帰クラス及び各生存期間のカテゴリに関する遺伝子発現の異なるパターンを確定する。

これらの結果は、表2のように視覚化することができる。

【0102】

【表2】

表2

	遺伝子1	遺伝子2	遺伝子3	遺伝子4	遺伝子5	遺伝子6
名称	ジストロフィン関連タンパク質	プロテイン・キナーゼCガンマ	MINOR/NOR1	PDE4B	プロテイン・キナーゼCベータ-1	Zink-フィンガー・タンパク質C2H2-150
治癒クラス	0	58	58	58	58	43
致死クラス	58	4	54	23	1	57

10

20

【0103】

図34aは、致命的クラスのグレイスケールクラスタを、他方図34bは、治癒クラスのグレイスケールクラスタを示している。カテゴリ2及び3の致命的症例は、治癒に対し遺伝子発現プロファイルに関して近似していることが認められる。カテゴリ1の致命的症例とカテゴリ2及び3の致命的症例の間には有意差が存在し、これは遺伝子発現により検出し、及びこのモデルにより正確に予測することができる。このモデルは、新規症例の推定に加え、結果として、新規症例の生存期間に関する予後を示す高度に合致したプロトタイプをもたらす。

30

【0104】

例として、致命的クラス、カテゴリ1のプロファイルを定義する訓練されたEFuNNから抽出されたルールのひとつを以下に示す：

ルール4：IF 03が(Low 0.83)and04が(High 0.83)andG8が(Medium 0.91)andG9が(Low 0.83)andG11が(Low 0.83)であるならば、THENクラスは[致命的]である：

ここで、G1などは、Shippらの論文(全体が本明細書に参照として組入れられている)と同じ順番で列記された遺伝子である。

【0105】

第二の実験において、Shippらの論文と同じ58名の患者に関する6,817個の遺伝子の生のAffymetrix遺伝子発現データを用いたが、しかし各IPIクラスについては、同じEFuNN技術を異なるパラメータ値で使用し、個別のモデルを作成した。各モデルについて、最初の6,817種から、5種の遺伝子を、Shippらと同じシグナル対ノイズ比解析を用い、選択した。4種の各IPIクラスモデルに関する予後(治癒対致命的)の精度は以下のものである：IPI=1(低)について、 $91.9 \pm 3.9$ ；IPI=2(低中間)について、 $96.4 \pm 5.7$ ；IPI=3(高中間)について、 $90.4 \pm 6.1$ ；IPI=4について、全てのクラスがクラス「致命的」に属するので、予後は全てのクラスで100%である。

40

【0106】

これらの試料は、それらのIPIに従い階層化した：

IPI=低：26試料、その19種は治癒クラスに、及び7種は致命的クラスに属する。

IPI=低中間：11試料、その7種は治癒クラスに、及び4種は致命的クラスに属する。

50

IPI = 高中間：17試料、その4種は治癒クラスに、及び13種は致命的クラスに属する。

IPI = 高：2試料、その2種は致命的クラスに属する。更なる分類に、マイクロアレイデータは不要である。

IPI = 不明：2試料。これらの試料は個々では考察しなかった。

【0107】

3種のIPIカテゴリの各々について、遺伝子選択及び予後判定を、N-クロスバリデーションを基に行った。最初に、5種の遺伝子を、各クロスバリデーション試行において、各IPIについて選択した。分類は、下記パラメータ値での、EFuNNにより行った：メンバシップ関数MF=2、誤差閾値Errth=0.1、適用した集計なし；入力ベクトル間の類似性の測定のためのユークリッド距離；EFuNNモデルのルール節点に関するラジアルベースの活性化関数；各クロスバリデーションモデルにおける訓練データセットを3回反復。各IPI群に関する30 N-クロスバリデーション法を試行し、分類性能を、表3に示したように評価した。

【0108】

【表3】

表3. 3つのIPIモデルについての交差確認実験の試験精度

IPI	低	低中間	高中間
元群	91.9±3.9	96.4±5.7	90.4±6.1

【0109】

ベースラインとして、ランダムIPI選択した試料の数及び同じクラス割合の群を作成した。この場合、精度はより低く、及び変動はより高く、例えば、3種のモデルの各々に関する変動は、各々、±6.4、±9.9、±6.5であった。

EFuNTNモデルのクロスバリデーション後、全ての例を用い、先の工程において定義されたような5遺伝子の減少された遺伝子セットを使用する、最終の予後モデルを構築した(表4)。ルールは、訓練したEFuNNから抽出し、各ルールは、5次元空間からの試料の群(クラスター)のプロファイルを表わしている。遺伝子発現プロファイルを用い、単に臨床環境における3種のIPIカテゴリのDLBCL患者治療の転帰を予測するのみではなく、製薬会社が新規標的医薬品を直接検索することもできる。

【0110】

10

20

30

【表 4】

表 4. IPIカテゴリーの各々についての交差確認EFuNNモデリングにおける選択遺伝子

カテゴリー	選択遺伝子のセット
IPI=1(低)	'KIAA0278遺伝子、部分cds' 26 'コルチコトロピン関連因子のプレプロ形態をコードする遺伝子' 25 'DBH ドーパミン・ベーターヒドロキシラーゼ(ドーパミン・ベーターモノオキシゲナーゼ)' 25 'SP4 Sp4転写因子' 14 'CCAATボックス-結合転写因子1' 14
IPI=2(低中間)	'LIMドメイン・タンパク質CLP-36 mRNA' 10 'GES2カルボキシルエステラーゼ2(肝)' 9 'KIAA0197遺伝子、部分cds' 7 'LYZリゾチーム' 3 'GS2374Iに関するApM2 mRNA(脂肪組織に特異的な未知の産物)' 2
IPI=3(高中間)	'HPRp18 mRNA' 15 'KIAA0036遺伝子' 15 'D-アスパルテート・オキシダーゼ' 15 'ピロホスファターゼに関する部分mRNA' 12 'JNK活性化キナーゼ(JNKK1)mRNA' 5

10

20

## 【0111】

## 実施例3

本実施例において、Ramaswamyらにより公開された遺伝子発現データベース、及び公に入手可能なもの(<http://www-genome.wi.mit.edu/MPR/GCM.html>)を用いた。このデータベースは、本明細書に参照として組入れられている。このデータベースは、14種の一般的な腫瘍型からの90種の正常組織試料及び218種の腫瘍試料に関する遺伝子発現値を含んでいる。各試料は、16,063種の遺伝子の発現レベル及び発現された配列タグ(EST)を有する。Ramaswamyらは、20種の腫瘍試料を、識別不良で、満足できない分化を生じたものとして示した。本発明者らは、腫瘍型間を識別する遺伝子発現値のパターンを同定することを意図しているため、識別不良の試料を除いたデータベースからの腫瘍試料を用いた。

30

## 【0112】

ここで考察された14種の腫瘍クラスの各々における遺伝子発現値のパターン(プロファイル)を同定するシステムは、前記特徴選択法を適用することにより開発し、並びに症例研究データ(Ramaswamiら)の特徴スペースを16,063遺伝子から399遺伝子へと最小化した。その後このデータセットを用い、14種の腫瘍クラス各々において遺伝子発現パターンをモデル化するように、EFuNNを訓練した。特定のクラスについて遺伝子の発現レベルを説明するルールを、EFuNNから抽出した。クラスは、1種よりも多いルールで説明することができるが、各ルールはただひとつのクラスを説明する。

40

## 【0113】

その後各腫瘍クラスのパターンを、そのクラスに関係する全てのルールを組合せることにより作成した。遺伝子は、その発現レベルのメンバシップ値(HIGH/LOW)が、その特定のクラスに関係する全てのルールの所定の閾値を上回る場合には、そのパターンに含まれる。閾値は、そのパターンの特性を特徴付ける指標をもたらす。この場合、閾値は、そのパターンの遺伝子の重要性を示す。例えば閾値が0.9である場合に、遺伝子gが乳がんのプロファイルに存在するならば、恐らくこれは、このクラスの重要な遺伝子であろう。すなわち、このEFuNNから抽出されたルールにおける遺伝子のメンバシップ値は、このクラスに関係した全てのルールにおいて0.9を上回る。場合によっては、この閾値は、比較的低く

50

設定されなければならない。このことは、特性(遺伝子)は重要ではなく、単にこれは、例えそのクラスに関連した全てのルールに共通であっても、このパターンにおける強力な特性ではないことを意味するものではない。

#### 【0114】

図を図5a及び5bに示している。300種の異なる種類の癌のクラスプロファイルを、抽出している。各クラスのプロファイルは、クラスプロファイルを明らかにするために、それを上回る遺伝子は過剰発現されるか又は過小発現されるかのいずれかであるようなメンバシップ値を定義する閾値 $C_{i_{her}}$ により、修飾することができる。

目盛り付きスケール320は、大きい枠の左上部のグレイスケール340で示された遺伝子の過小発現のレベルを示しているのに対し、他方の目盛り付きスケール330は、大きい枠の左上部のグレイスケール350で示された遺伝子の過剰発現のレベルを示している。各クラスのプロファイルは、この遺伝子についてプロファイルを明らかにするために、それを上回る遺伝子はこのクラスの全てのルールにおいて過剰発現されるか又は過小発現されるかのいずれかであるようなメンバシップ値を定義する各個別のクラスについて調整した閾値 $C_{i_{her}}$ により修飾することができる。癌の各クラスは、過小発現されたか又は過剰発現されたかのいずれかであるこのクラスの全ての組織試料において、共通の遺伝子を表わすただひとつのパターンにより表わされる。この組合せたクラスパターンに加え、クラスについていくつかの亜群(クラスタ)を、その方法により同定することができ、及び同じクラスの各亜群の特徴的パターンを抽出することができる。下記表において、最初にクラスプロファイルが抽出され、第二に各クラスについて - 修飾することができる定義された閾値について、亜群プロファイルが同じく抽出される。閾値が低下するにつれて、より詳細なプロファイルが抽出される。

#### 【0115】

14種の腫瘍クラスの各々に関するパターンが、図示されている。枠350a中の直線は、遺伝子の高発現レベルを表わし、及び340の直線は、遺伝子の低発現レベルを表わしている。これらの色の相対強度は、遺伝子に関する平均メンバシップ値を表わし、そのルールへのその遺伝子の関与の強度を示している。更に図のインターフェースは、以後図(b)において示されるように、使用者が個々のルールを目視することを可能にしている。これは、特定のパターンを作成するルールを目視による解釈を可能にする。これは、特にパターンが、何らかの共通の遺伝子を示さないことが明らかな場合に有用であろう。個々のルールを目視する場合には、全てのルールにおいて明らかである必要はない関心のある遺伝子を同定することを可能するであろう。

#### 【0116】

精度は、リンパ腫症例において評価し、使用した閾値が0.6である場合、13遺伝子が選択された。このモデルの精度は、97%である。遺伝子の最小数を選択するために、13遺伝子から異なるモデルを作成し、1遺伝子から始め、それらの精度を評価し、その結果(遺伝子数)/精度の比は最適化される。例えば、遺伝子D64142の使用は精度64%をもたらしたが、追加の2種の別の遺伝子は精度(信頼)95%をもたらした。

本発明の範囲は前記実施例のみに限定されることは意図されない。当業者に理解されるように、(添付された「特許請求の範囲」に示されたような)本発明の範囲から逸脱しない限りは、多くの変更が可能である。

#### 【0117】

図6は、ここで考察した他の13種の癌のプロファイルに対する、乳がん遺伝子の遺伝子発現クラスプロファイルを示している。これらの遺伝子は、遺伝子名に対応する独自の番号である、それらのGenBank寄託番号により表わされる。下記の序列番号(order number)を伴うリスト上の番号付けされた遺伝子は、過剰発現されている：2,3。残りの遺伝子は、過小発現されている。過剰又は過小発現されたファジイメンバシップ値の最低レベルは、0.55である。図7aは、乳がんに関する10種の遺伝子発現亜群プロファイルを示し、過小発現された遺伝子を強調している。図7bは、同じデータを示しているが、過剰発現された遺伝子を強調している。

10

20

30

40

50

## 【0118】

図8は、ここで考察した他の13種の癌のプロファイルに対する、前立腺癌遺伝子の遺伝子発現クラスプロファイルを示している。これらの遺伝子は、それらのGenBank寄託番号により表わされる。下記の序列番号を伴うリスト上の番号付けされた遺伝子は、過小発現されている：1。残りの遺伝子は、過剰発現されている。過剰又は過小発現されたファジイメンバシップ値の最低レベルは、0.62である。図9a及び9bは、前立腺癌に関する5種の遺伝子発現亜群プロファイルを示し、各々、過小発現及び過剰発現された遺伝子を強調している。

## 【0119】

図10は、ここで考察した他の13種の癌のプロファイルに対する、肺腺癌遺伝子の遺伝子発現クラスプロファイルを示している。これらの遺伝子は、それらのGenBank寄託番号により表わされる。下記の序列番号を伴うリスト上の番号付けされた遺伝子は、過小発現されている：4、8、9、15。残りの遺伝子は、過剰発現されている。過剰又は過小発現されたファジイメンバシップ値の最低レベルは、0.6である。図11a及び11bは、肺腺癌に関する8種の遺伝子発現亜群プロファイルを示し、各々、過小発現及び過剰発現された遺伝子を強調している。

## 【0120】

図12は、ここで考察した他の13種の癌のプロファイルに対する、結腸直腸腺癌遺伝子の遺伝子発現クラスプロファイルを示している。これらの遺伝子は、それらのGenBank寄託番号により表わされる。下記の序列番号を伴うリスト上の番号付けされた遺伝子は、過剰発現されている：1、2、3、6、8、9、10、17、22。残りの遺伝子は、過小発現されている。過剰又は過小発現されたファジイメンバシップ値の最低レベルは、0.65である。図13a及び13bは、結腸直腸腺癌に関する5種の遺伝子発現亜群プロファイルを示し、各々、過小発現及び過剰発現された遺伝子を強調している。

## 【0121】

図14は、ここで考察した他の13種の癌のプロファイルに対する、リンパ腫遺伝子の遺伝子発現クラスプロファイルを示している。これらの遺伝子は、それらのGenBank寄託番号により表わされる。下記の序列番号を伴うリスト上の番号付けされた遺伝子は、過剰発現されている：3、6、8、9、10。残りの遺伝子は、過小発現されている。過剰又は過小発現されたファジイメンバシップ値の最低レベルは、0.6である。図15a及び15bは、リンパ腫に関する6種の遺伝子発現亜群プロファイルを示し、各々、過小発現及び過剰発現された遺伝子を強調している。

## 【0122】

図16は、ここで考察した他の13種の癌のプロファイルに対する、膀胱移行細胞癌遺伝子の遺伝子発現クラスプロファイルを示している。これらの遺伝子は、それらのGenBank寄託番号により表わされる。下記の序列番号を伴うリスト上の番号付けされた遺伝子は、過剰発現されている：1、2、3、5。残りの遺伝子は、過小発現されている。過剰又は過小発現されたファジイメンバシップ値の最低レベルは、0.6である。図17a及び17bは、膀胱移行細胞癌に関する7種の遺伝子発現亜群プロファイルを示し、各々、過小発現及び過剰発現された遺伝子を強調している。

## 【0123】

図18は、ここで考察した他の13種の癌のプロファイルに対する、黒色腫遺伝子の遺伝子発現クラスプロファイルを示している。これらの遺伝子は、それらのGenBank寄託番号により表わされる。下記の序列番号を伴うリスト上の番号付けされた遺伝子は、過剰発現されている：1、3、4、5、6、7。遺伝子2は、過小発現されている。過剰又は過小発現されたファジイメンバシップ値の最低レベルは、0.65である。図19a及び19bは、黒色腫に関する5種の遺伝子発現亜群プロファイルを示し、各々、過小発現及び過剰発現された遺伝子を強調している。

## 【0124】

図20は、ここで考察した他の13種の癌のプロファイルに対する、子宮腺癌遺伝子の遺伝

子発現クラスプロファイルを示している。これらの遺伝子は、それらのGenBank寄託番号により表わされる。下記の序列番号を伴うリスト上の番号付けされた遺伝子は、過剰発現されている：1、2、3、4、5、6、8、9、10。下記リストの残りの遺伝子は、過小発現されている。過剰又は過小発現されたファジイメンバシップ値の最低レベルは、0.65である。図21a及び21bは、子宮腺癌に関する4種の遺伝子発現亜群プロファイルを示し、各々、過小発現及び過剰発現された遺伝子を強調している。過剰又は過小発現されたファジイメンバシップ値の最低レベルは、0.7である。

【0125】

図22は、ここで考察した他の13種の癌のプロファイルに対する、白血病遺伝子の遺伝子発現クラスプロファイルを示している。これらの遺伝子は、それらのGenBank寄託番号により表わされる。下記の序列番号を伴うリスト上の番号付けされた遺伝子は、過剰発現されている：1、2、3、4、5、6、7、9。下記リストの残りの遺伝子は、過小発現されている。過剰又は過小発現されたファジイメンバシップ値の最低レベルは、0.65である。図23a及び23bは、白血病に関する5種の遺伝子発現亜群プロファイルを示し、各々、過小発現及び過剰発現された遺伝子を強調している。

10

【0126】

図24は、ここで考察した他の13種の癌のプロファイルに対する、腎細胞癌遺伝子の遺伝子発現クラスプロファイルを示している。これらの遺伝子は、それらのGenBank寄託番号により表わされる。下記の序列番号を伴うリスト上の番号付けされた遺伝子は、過剰発現されている：6、7。下記リストの残りの遺伝子は、過小発現されている。過剰又は過小発現されたファジイメンバシップ値の最低レベルは、0.63である。図25a及び25bは、腎細胞癌に関する6種の遺伝子発現亜群プロファイルを示し、各々、過小発現及び過剰発現された遺伝子を強調している。

20

【0127】

図26は、ここで考察した他の13種の癌のプロファイルに対する、膵臓腺癌遺伝子の遺伝子発現クラスプロファイルを示している。これらの遺伝子は、それらのGenBank寄託番号により表わされる。下記の序列番号を伴うリスト上の番号付けされた遺伝子は、過剰発現されている：1、2、3、4、5、6、8。下記リストの残りの遺伝子は、過小発現されている。過剰又は過小発現されたファジイメンバシップ値の最低レベルは、0.6である。図27a及び27bは、膵臓腺癌に関する6種の遺伝子発現亜群プロファイルを示し、各々、過小発現及び過剰発現された遺伝子を強調している。

30

【0128】

図28は、ここで考察した他の13種の癌のプロファイルに対する、卵巣腺癌遺伝子の遺伝子発現クラスプロファイルを示している。これらの遺伝子は、それらのGenBank寄託番号により表わされる。下記の序列番号を伴うリスト上の番号付けされた遺伝子は、過剰発現されている：9。下記リストの残りの遺伝子は、過小発現されている。過剰又は過小発現されたファジイメンバシップ値の最低レベルは、0.6である。図29a及び29bは、卵巣腺癌に関する7種の遺伝子発現亜群プロファイルを示し、各々、過小発現及び過剰発現された遺伝子を強調している。過剰又は過小発現されたファジイメンバシップ値の最低レベルは、0.65である。

40

【0129】

図30は、ここで考察した他の13種の癌のプロファイルに対する、胸膜中皮腫遺伝子の遺伝子発現クラスプロファイルを示している。これらの遺伝子は、それらのGenBank寄託番号により表わされる。下記の序列番号を伴うリスト上の番号付けされた遺伝子は、過剰発現されている：2、13、14、16、17、18、19、20、21。下記リストの残りの遺伝子は、過小発現されている。過剰又は過小発現されたファジイメンバシップ値の最低レベルは、0.7である。図31a及び31bは、胸膜中皮腫に関する4種の遺伝子発現亜群プロファイルを示し、各々、過小発現及び過剰発現された遺伝子を強調している。

【0130】

図32は、ここで考察した他の13種の癌のプロファイルに対する、中枢神経系癌遺伝子の

50

遺伝子発現クラスプロファイルを示している。これらの遺伝子は、それらのGenBank寄託番号により表わされる。下記の序列番号を伴うリスト上の番号付けされた遺伝子は、過剰発現されている：6、7、8、10、11、12、13、14、15。下記リストの残りの遺伝子は、過小発現されている。過剰又は過小発現されたファジイメンバシップ値の最低レベルは、0.8である。図33a及び33bは、中枢神経系癌に関する3種の遺伝子発現亜群プロファイルを示し、各々、過小発現及び過剰発現された遺伝子を強調している。過剰又は過小発現されたファジイメンバシップ値の最低レベルは、0.7である。

【0131】

産業上の利用可能性

本発明のニューラルネットワーク及び方法は、診断、疾患の管理、療法の有効性の評価、疾患又は状態を診断するための薬剤及び試験キットの製造、薬物標的の評価、並びに様々な疾患及び状態の治療のための薬物の開発において有用である。

10

【0132】

下記刊行物を本明細書に援用する：

【化 1】

S. Dudoit, J. Fridlyand, and T. P. Speed, Comparison of Discrimination Methods for the Classification of Tumors Using Gene Expression Data, *Journal of the American Statistical Association*, vol.97, no.457, March, pp.77-87, 2002.

T.R. Golub, D. K. Slonim, P. Tamayo, C. Huard, M. Gaasenbeek, J. P. Mesirov, H. Coller, M. L. Loh, J. R. Downing, M. A. Caligiuri, C. D. Bloomfield, and E. S. Lander, Molecular classification of cancer: class discovery and class prediction by gene expression monitoring, *Science*, vol.286, 15 October, pp.531-537, 1999.

10

J. Khan, J.S.Wei, M. Ringner, L. H. Saal, M. Ladanyi, F. Westermann, F. Berthold, M. Schwab, C. R. Antonescu, C. Peterson, and P. S. Meltzer, Classification and diagnostic prediction of cancers using gene expression profiling and artificial neural networks, *Nature Medicine*, Nature Publishing Group, vol.7, no.6, pp.673-679, 2001.

20

N. Kasabov, *Evolving Connectionist Systems: Methods and Applications in Bioinformatics*, *Brain Study and Intelligent Machines*, Springer Verlag, 2002.

N. Kasabov, Evolving Connectionist Systems for Adaptive On-line Knowledge-Based Learning, *IEEE Transactions of Systems, Man and Cybernetics – part B: Cybernetics*, vol.31, no.6, pp.902-918, 2001.

30

N.Kasabov, Adaptive method and system, PCT, WO 01/78003 A1.

S. Ramaswamy, P. Tamayo, R. Rifkin, S. Mukherjee, C-H. Yeang, M. Angelo, C. Ladd, M. Reich, E. Latulippe, J. P. Mesirov, T. Poggio, W. Gerald, M. Loda, E. S. Lander, and

【化 2】

T. R. Golub, Multiclass cancer diagnosis using tumor gene expression signatures, *Proceedings of the National Academy of Sciences*, vol.98, no.26, pp.15149-15154, 2001.

F. P. Roth, Bringing out the best features of expression data, *Genome Research*, Cold Spring Harbor Laboratory Press, vol.11, no.11, pp.1878-1887, 2001.

10

M. A. Shipp, K. N. Ross, P. Tamayo, A. P. Weng, J. L. Kutok, R. C. T. Aguiar, M. Gaasenbeek, M. Angelo, M. Reich, G. S. Pinkus, T. S. Ray, M. A. Koval, K. W. Last, A. Norton, T. A. Lister, J. Mesirov, D. S. Neuberger, E. S. Lander, J. C. Aster, and T. R. Golub Diffuse large B-cell lymphoma outcome prediction by gene expression profiling and supervised machine learning, *Nature Medicine*, vol.8, no. 1, pp.68-74, 2002.

T. D. Wu, Analysing gene expression data from DNA microarrays to identify candidate genes, *Journal of Pathology*, JohnWiley and Sons, Ltd., vol.195, pp.53- 65, 2001.

20

C-H. Yeang, S. Ramaswamy, P. Tamayo, S. Mukherjee, R. M. Rifkin, M. Angelo, M. Reich, E. Lander, J. Mesirov, and T. Golub, Molecular classification of multiple tumor types, *Bioinformatics*, Oxford University Press, vol.17, Suppl.1, pp.S316-S322, 2001.

30

【0133】

疾患プロファイリングの方法及びシステムの好ましい形は、添付図面を参照し、ここに説明する：

【図面の簡単な説明】

【0134】

【図1】図1は、本発明において有用なニューラルネットワークモジュールの概略的ブロック略図である。

【図2】図2は、本発明の好ましい方法を例示する流れ図である。

【図3】図3は、訓練されたEFuNNから抽出された各癌クラスに関連するルールを示す。

【図4a】図4aは、EFuNNに関連したように同定された遺伝子の各々に関する遺伝子の遺伝子発現レベルを示す。影付き領域は、過小発現領域を示す。

40

【図4b】図4bは、EFuNNに関連したように同定された遺伝子の各々に関する遺伝子の遺伝子発現レベルを示す。影付き領域は、過小発現領域を示す。

【図5a】図5aは、過小発現される遺伝子に注目する本発明の方法を適用することによる、抽出された14種の癌に関するクラスプロファイルを示す。

【図5b】図5bは、過剰発現される遺伝子に注目する図5aの14種の癌に関するクラスプロファイルを示す。

【図6】図6は、乳癌遺伝子の遺伝子発現クラスプロファイルを示す。

【図7a】図7aは、過小発現された遺伝子に注目する乳癌の10種の遺伝子発現亜群プロファイルを示す。

50

【図7b】図7bは、過剰発現された遺伝子に注目する乳腺癌の10種の遺伝子発現亜群プロファイルを示す。

【図8】図8は、前立腺癌遺伝子の遺伝子発現プロファイルを示す。

【図9a】図9aは、過小発現された遺伝子に注目する前立腺癌の5種の遺伝子発現亜群プロファイルを示す。

【図9b】図9bは、過剰発現された遺伝子に注目する前立腺癌の5種の遺伝子発現亜群プロファイルを示す。

【図10】図10は、肺腺癌遺伝子の遺伝子発現プロファイルを示す。

【図11a】図11aは、過小発現された遺伝子に注目する肺腺癌の8種の遺伝子発現亜群プロファイルを示す。

【図11b】図11bは、過剰発現された遺伝子に注目する肺腺癌の8種の遺伝子発現亜群プロファイルを示す。

【図12】図12は、結腸直腸腺癌遺伝子の遺伝子発現プロファイルを示す。

【図13a】過小発現及び過剰発現された遺伝子に注目する結腸直腸腺癌の5種の遺伝子発現亜群プロファイルを示す。

【図13b】過小発現及び過剰発現された遺伝子に注目する結腸直腸腺癌の5種の遺伝子発現亜群プロファイルを示す。

【図14】図14は、リンパ腫遺伝子の遺伝子発現プロファイルを示す。

【図15a】過小発現及び過剰発現に注目するリンパ腫の6種の遺伝子発現亜群プロファイルを示す。

【図15b】過小発現及び過剰発現に注目するリンパ腫の6種の遺伝子発現亜群プロファイルを示す。

【図16】図16は、膀胱移行細胞癌遺伝子の遺伝子発現プロファイルを示す。

【図17a】過小発現及び過剰発現に注目する膀胱移行細胞癌の7種の遺伝子発現亜群プロファイルを示す。

【図17b】過小発現及び過剰発現に注目する膀胱移行細胞癌の7種の遺伝子発現亜群プロファイルを示す。

【図18】図18は、黒色腫遺伝子の遺伝子発現プロファイルを示す。

【図19a】過小発現及び過剰発現に注目する黒色腫の5種の遺伝子発現亜群プロファイルを示す。

【図19b】過小発現及び過剰発現に注目する黒色腫の5種の遺伝子発現亜群プロファイルを示す。

【図20】図20は、子宮腺癌遺伝子の遺伝子発現プロファイルを示す。

【図21a】過小発現及び過剰発現に注目する子宮腺癌の4種の遺伝子発現亜群プロファイルを示す。

【図21b】過小発現及び過剰発現に注目する子宮腺癌の4種の遺伝子発現亜群プロファイルを示す。

【図22】図22は、白血病遺伝子の遺伝子発現プロファイルを示す。

【図23a】過小発現及び過剰発現に注目する白血病の5種の遺伝子発現亜群プロファイルを示す。

【図23b】過小発現及び過剰発現に注目する白血病の5種の遺伝子発現亜群プロファイルを示す。

【図24】図24は、腎細胞癌遺伝子の遺伝子発現プロファイルを示す。

【図25a】過小発現及び過剰発現に注目する腎細胞癌の6種の遺伝子発現亜群プロファイルを示す。

【図25b】過小発現及び過剰発現に注目する腎細胞癌の6種の遺伝子発現亜群プロファイルを示す。

【図26】図26は、膵臓腺癌遺伝子の遺伝子発現プロファイルを示す。

【図27a】過小発現及び過剰発現に注目する膵臓腺癌の6種の遺伝子発現亜群プロファイルを示す。

10

20

30

40

50

【図 27 b】過小発現及び過剰発現に注目する膵臓腺癌の6種の遺伝子発現亜群プロファイルを示す。

【図 28】図 28は、卵巣腺癌遺伝子の遺伝子発現プロファイルを示す。

【図 29 a】過小発現及び過剰発現に注目する卵巣腺癌の7種の遺伝子発現亜群プロファイルを示す。

【図 29 b】過小発現及び過剰発現に注目する卵巣腺癌の7種の遺伝子発現亜群プロファイルを示す。

【図 30】図 30は、胸膜中皮腫遺伝子の遺伝子発現プロファイルを示す。

【図 31 a】過小発現及び過剰発現に注目する胸膜中皮腫の4種の遺伝子発現亜群プロファイルを示す。

【図 31 b】過小発現及び過剰発現に注目する胸膜中皮腫の4種の遺伝子発現亜群プロファイルを示す。

【図 32】図 32は、中枢神経系遺伝子の遺伝子発現プロファイルを示す。

【図 33 a】過小発現及び過剰発現に注目する中枢神経系癌の3種の遺伝子発現亜群プロファイルを示す。

【図 33 b】過小発現及び過剰発現に注目する中枢神経系癌の3種の遺伝子発現亜群プロファイルを示す。

【図 34 a】致命的状態及び治癒状態に注目する本発明の方法を用いる、様々な癌型の予後の転帰を示す。

【図 34 b】致命的状態及び治癒状態に注目する本発明の方法を用いる、様々な癌型の予後の転帰を示す。

10

20

【図 1】

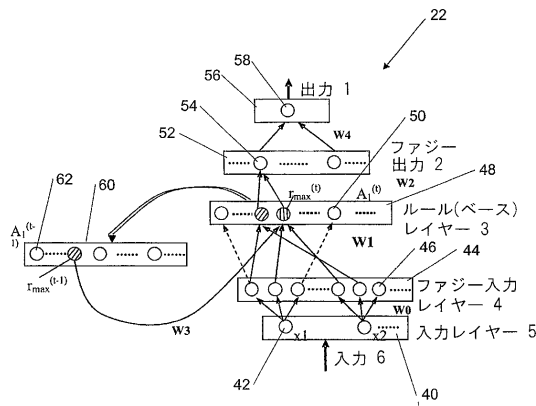


FIGURE 1

【図 2】

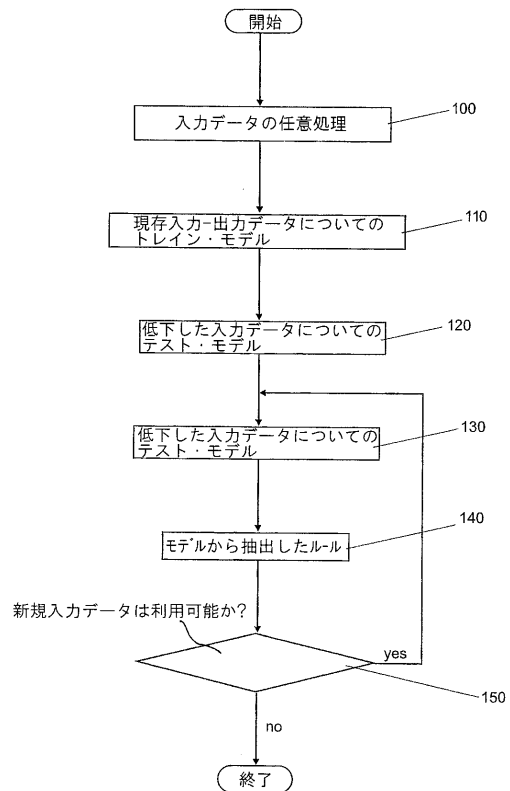


FIGURE 2

【 図 3 】

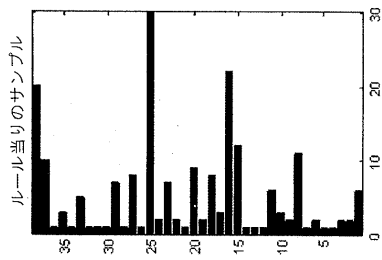


FIGURE 3

【 図 4 b 】

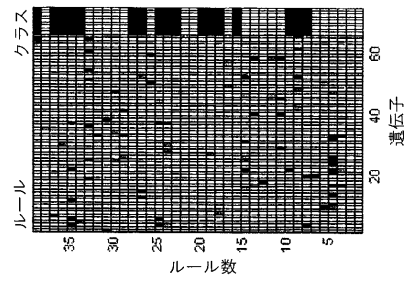


FIGURE 4b

【 図 4 a 】

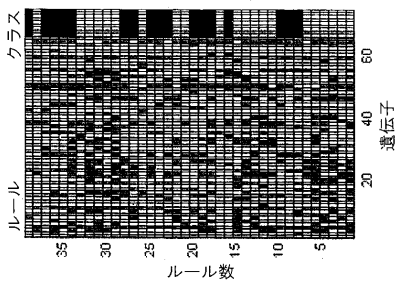


FIGURE 4a

【 図 5 a 】

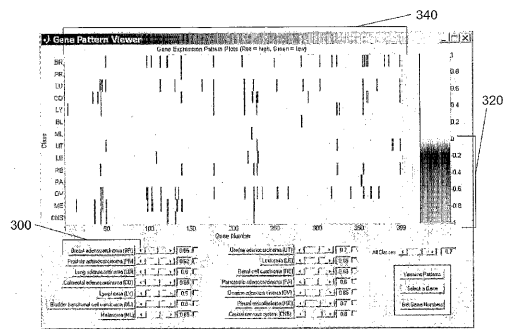


FIGURE 5a

【 図 5 b 】

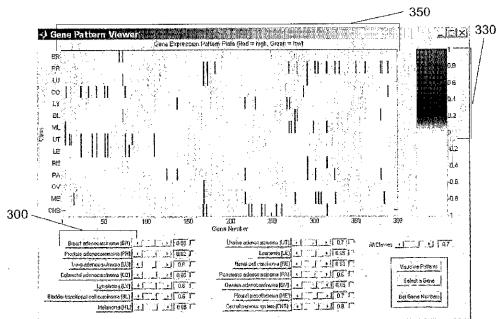


FIGURE 5b

【 図 7 a 】

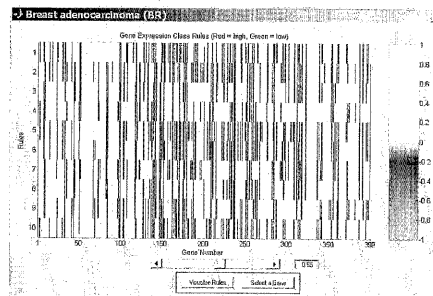


FIGURE 7a

【 図 6 】

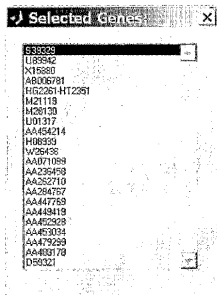


FIGURE 6

【 図 7 b 】

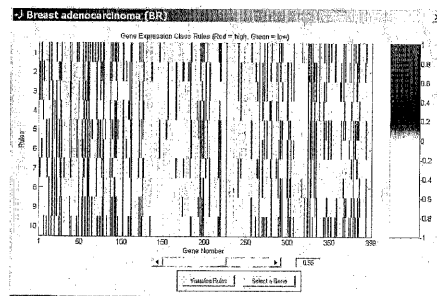


FIGURE 7b

【 図 8 】

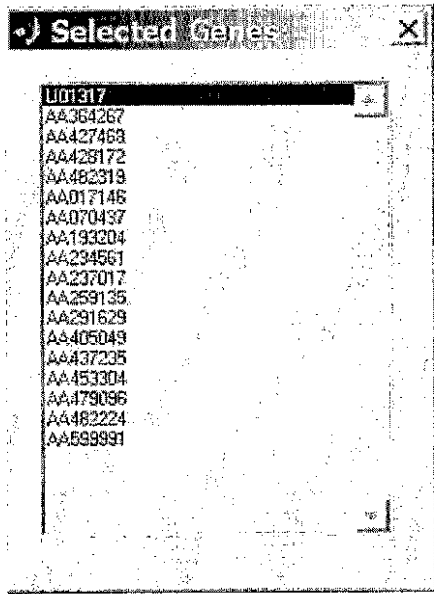


FIGURE 8

【 図 10 】

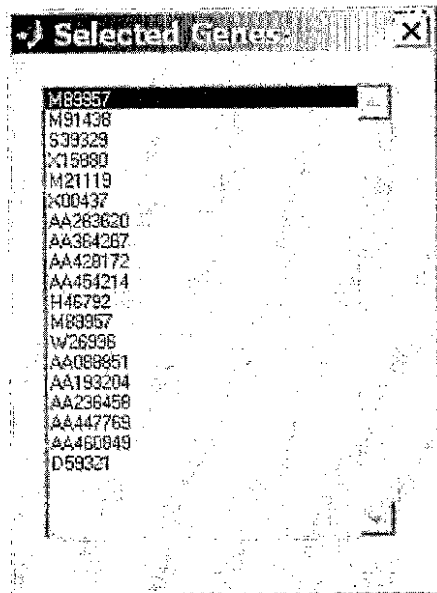


FIGURE 10

【 図 9 a 】

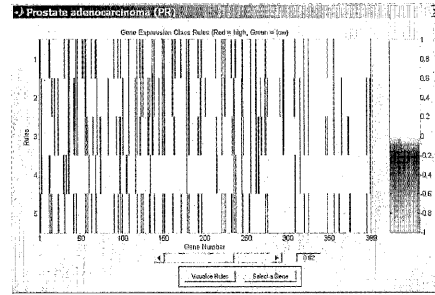


FIGURE 9a

【 図 9 b 】

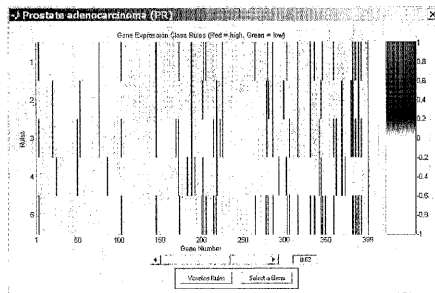


FIGURE 9b

【 図 11 a 】

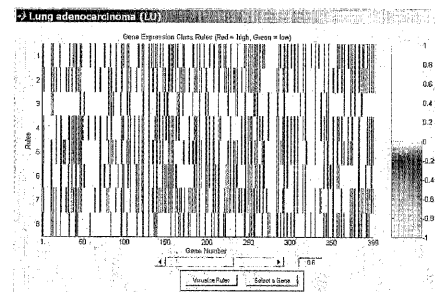


FIGURE 11a

【 図 11 b 】

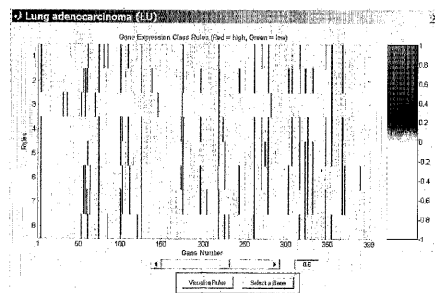


FIGURE 11b

【 1 2 】

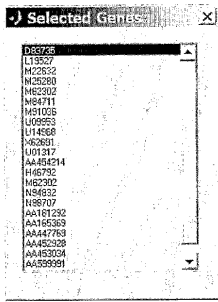


FIGURE 12

【 1 3 b 】

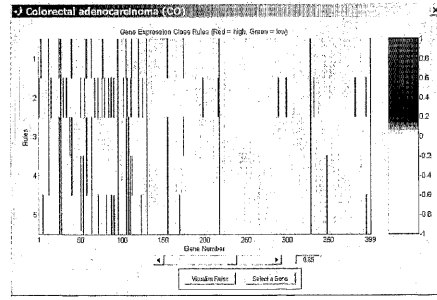


FIGURE 13b

【 1 3 a 】

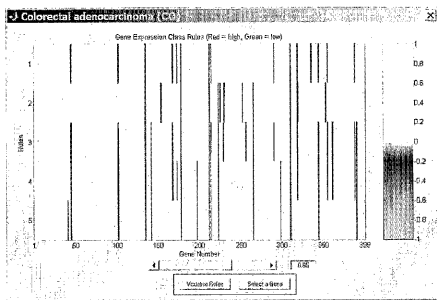


FIGURE 13a

【 1 4 】

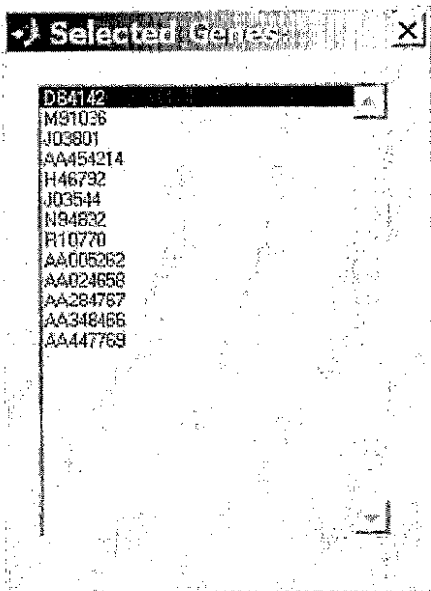


FIGURE 14

【 1 5 a 】

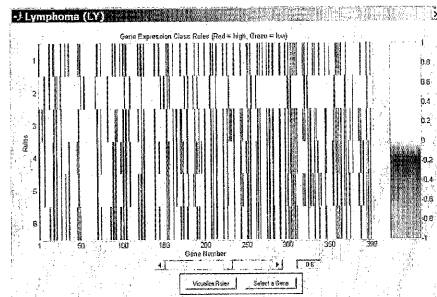


FIGURE 15a

【 1 5 b 】

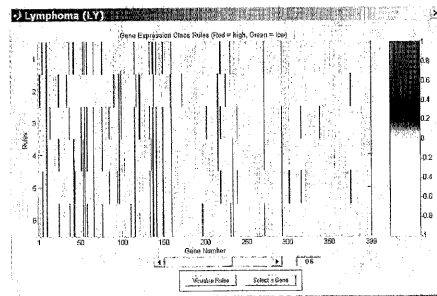


FIGURE 15b

【 図 16 】

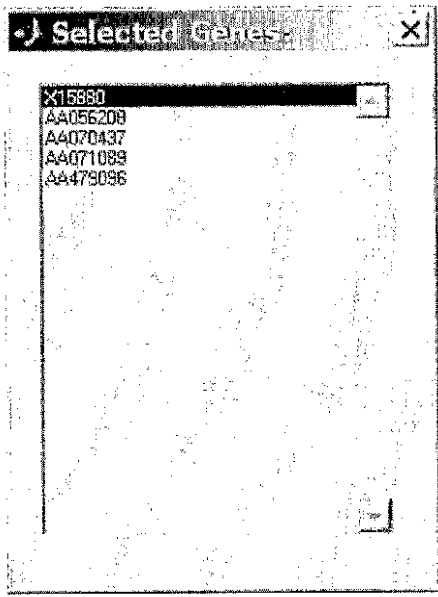


FIGURE 16

【 図 17 a 】

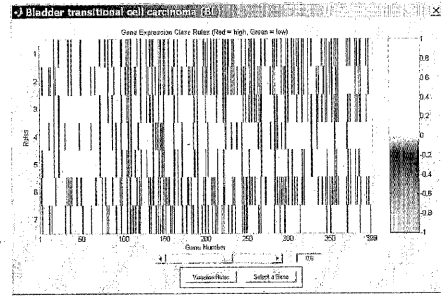


FIGURE 17a

【 図 17 b 】

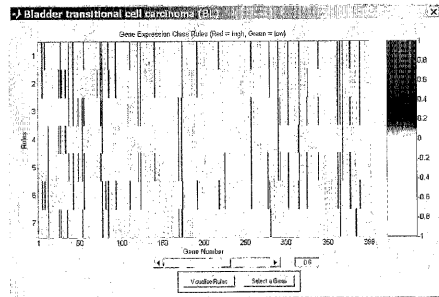


FIGURE 17b

【 図 18 】

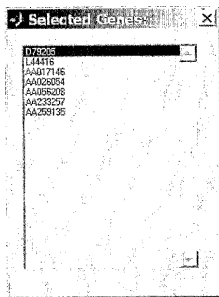


FIGURE 18

【 図 19 b 】

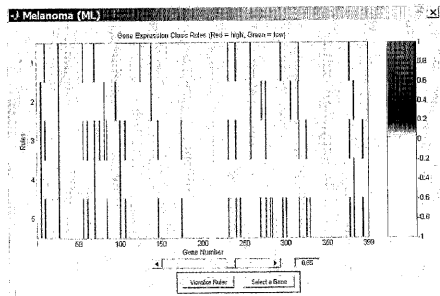


FIGURE 19b

【 図 19 a 】

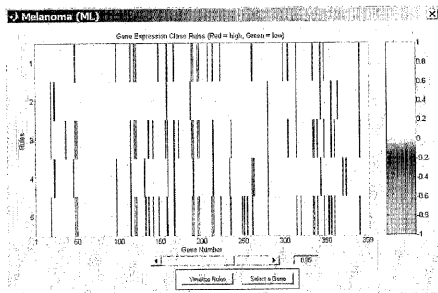


FIGURE 19a

【 2 0 】

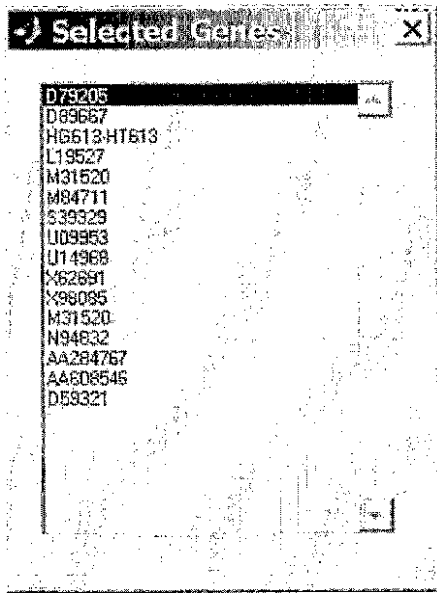


FIGURE 20

【 2 1 a 】

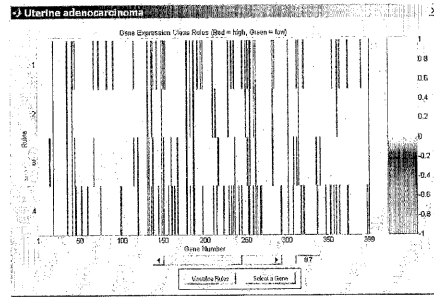


FIGURE 21a

【 2 1 b 】

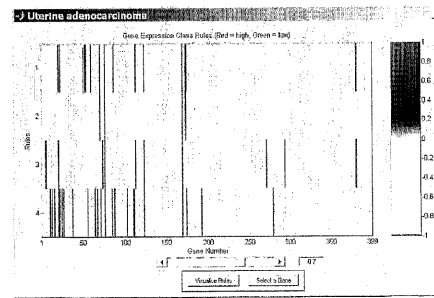


FIGURE 21b

【 2 2 】

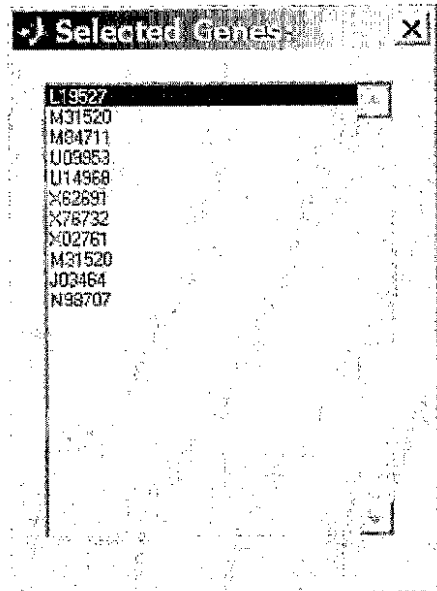


FIGURE 22

【 2 3 a 】

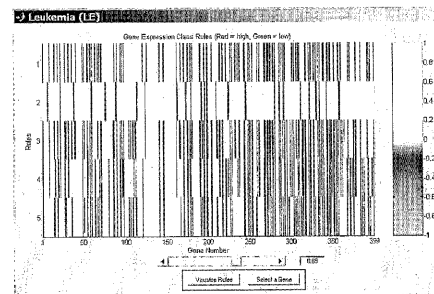


FIGURE 23a

【 2 3 b 】

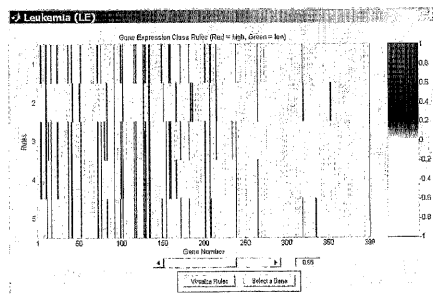


FIGURE 23b

【 2 4 】

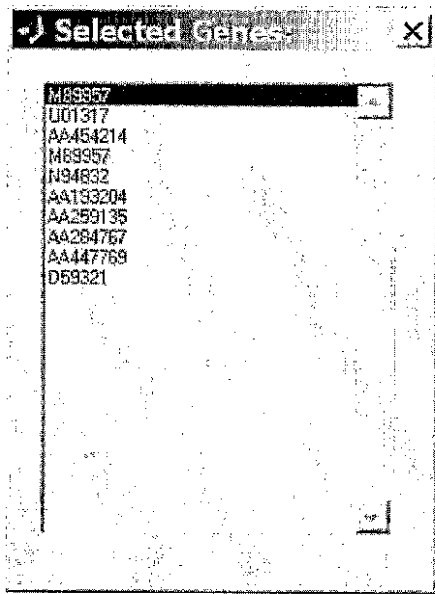


FIGURE 24

【 2 6 】

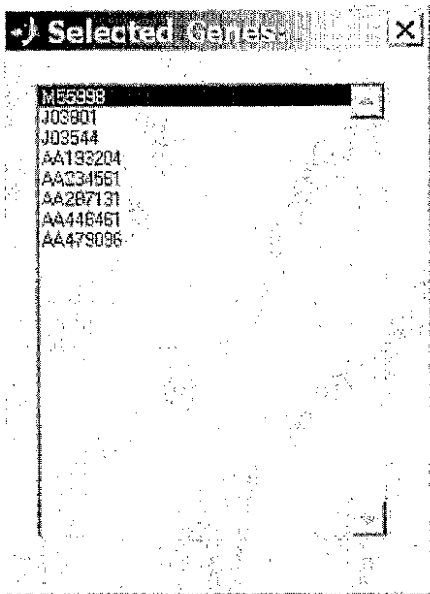


FIGURE 26

【 2 5 a 】

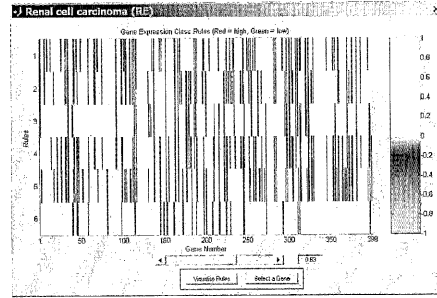


FIGURE 25a

【 2 5 b 】

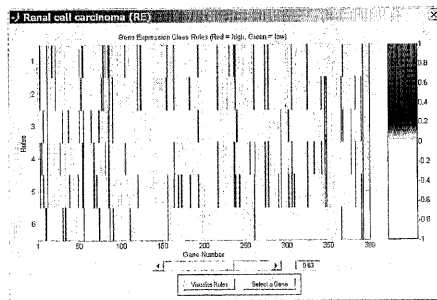


FIGURE 25b

【 2 7 a 】

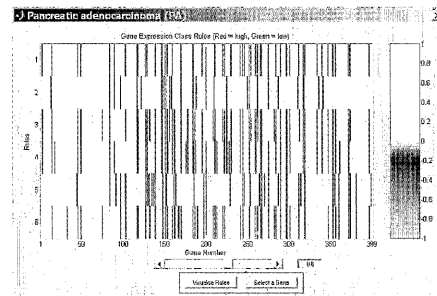


FIGURE 27a

【 2 7 b 】

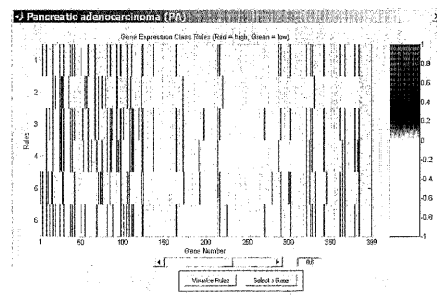


FIGURE 27b

【 図 28 】

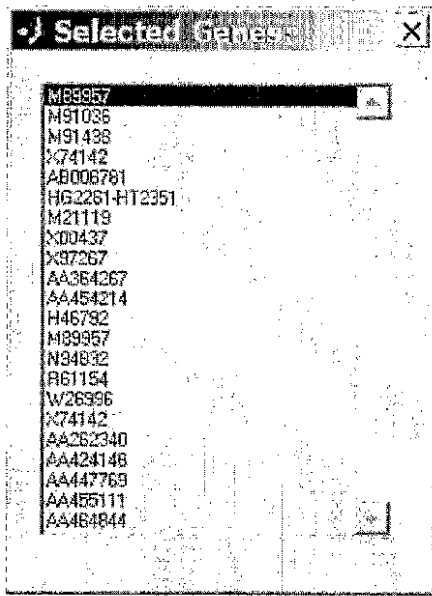


FIGURE 28

【 図 30 】

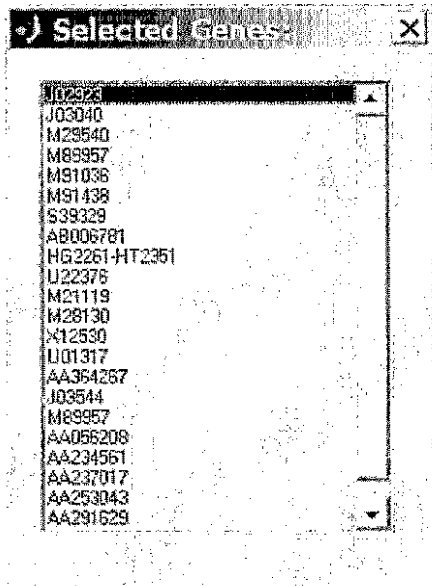


FIGURE 30

【 図 29 a 】

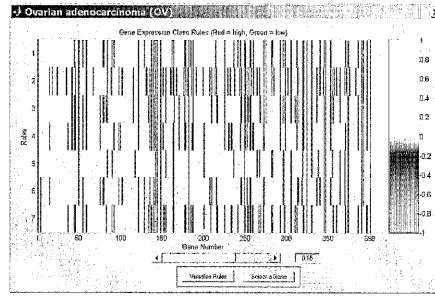


FIGURE 29a

【 図 29 b 】

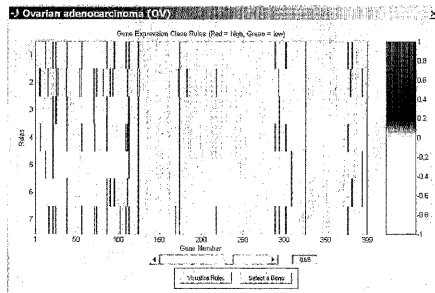


FIGURE 29b

【 図 31 a 】

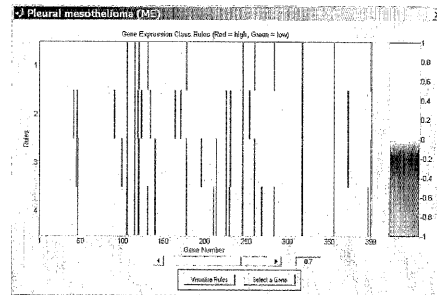


FIGURE 31a

【 図 31 b 】

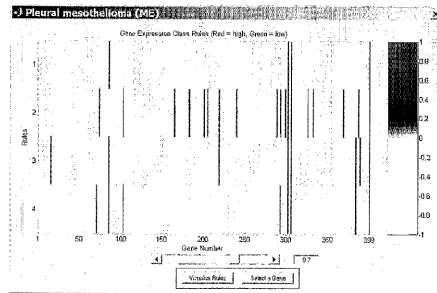


FIGURE 31b

【 図 3 2 】

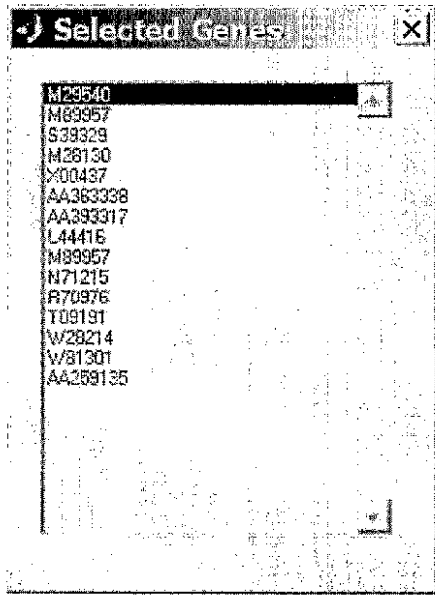


FIGURE 32

【 図 3 3 a 】

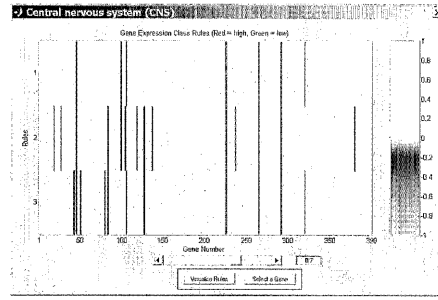


FIGURE 33a

【 図 3 3 b 】

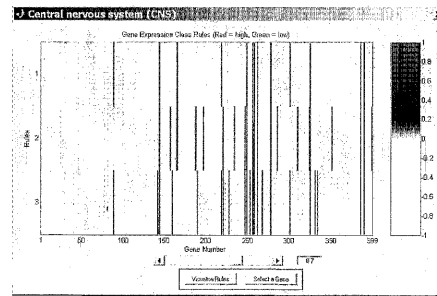


FIGURE 33b

【 図 3 4 a 】

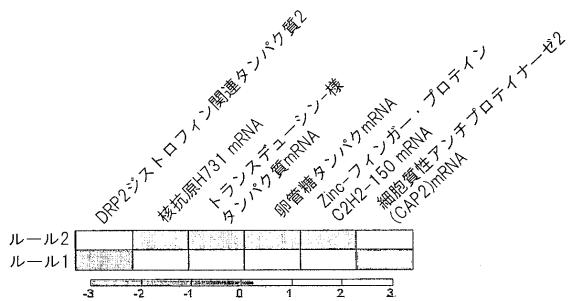


FIGURE 34a

【 図 3 4 b 】

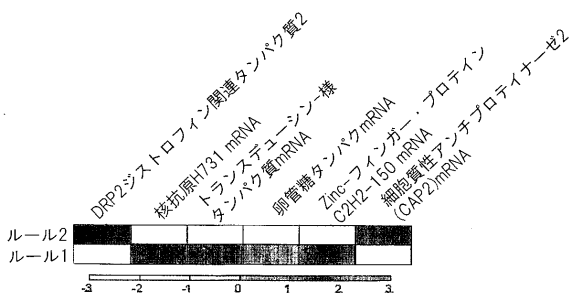


FIGURE 34b

## 【手続補正書】

【提出日】平成16年11月16日(2004.11.16)

## 【手続補正1】

【補正対象書類名】特許請求の範囲

【補正対象項目名】全文

【補正方法】変更

【補正の内容】

【特許請求の範囲】

## 【請求項1】

遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む入力層；1個又は複数のルール節点を含むルールベース層；1種又は複数の状態を出力するために構成された1個又は複数の出力節点を含む、出力層；及び、遺伝子発現データと1種又は複数の状態の間の関係を表わしているルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備える、ニューラルネットワークモジュール。

## 【請求項2】

遺伝子発現データと1種又は複数の状態の間の関係を決定する汎用方法であり、少なくとも：

- a) 1種又は複数の予め決定された状態のクラスにカテゴリ化された遺伝子発現データのセットを提供する工程；
- b) 該遺伝子発現データ及び該1種又は複数の予め決定された状態のクラスについて、ニューラルネットワークモジュールを訓練する工程であり、ここでこのニューラルネットワークモジュールが、遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の状態のクラスを出力するために構成された1個又は複数の出力節点を含む、出力層；及び、遺伝子発現データと1種又は複数の状態のクラスの間関係を表わすルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備える工程；並びに
- c) ルールベース層からルールを抽出する工程であり、該ルールが、遺伝子発現データと1種又は複数の状態のクラスの間関係を表わす工程を含む、方法。

## 【請求項3】

遺伝子発現データと1種又は複数の状態の間の関係を決定するシステムであり：

- a) 1種又は複数の予め決定された状態のクラスにカテゴリ化された遺伝子発現データのセットを受け取ることが可能な、入力；
- b) 該遺伝子発現データ及び該1種又は複数の予め決定された状態のクラスについて訓練可能なニューラルネットワークモジュールであり、ここでこのニューラルネットワークモジュールが、遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の状態のクラスを出力するために構成された1個又は複数の出力節点を含む、出力層；及び、遺伝子発現データと1種又は複数の状態のクラスの間関係を表わすルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備える、モジュール；並びに
- c) ルールベース層からルールを抽出する手段であり、該ルールが、遺伝子発現データと1種又は複数の状態のクラスの間関係を表わしている手段を具備する、システム。

## 【請求項4】

患者の状態を診断する汎用方法であり、該患者の生物学的試料から抽出された遺伝子発現データが、その遺伝子発現データと1種又は複数の状態のクラスの間関係を表わすルールを満足するかどうかを決定する工程を含み、該ルールが、請求項2記載の方法により決定される、方法。

## 【請求項5】

患者の状態を診断する汎用システムであり、該患者の生物学的試料から抽出された遺伝子発現データが、その遺伝子発現データと1種又は複数の状態のクラスの間関係を表わすルールを満足するかどうかを決定する手段を備え、該ルールが、請求項2記載の方法に

より決定される、システム。

【請求項6】

1種又は複数の状態に連結された過剰発現された又は過小発現された遺伝子を識別するセットを選択する方法であり：

- a) 1種又は複数の予め決定された状態にカテゴリ化された遺伝子発現データの1種又は複数のセットを提供する工程；
- b) 該遺伝子発現データ及び1種又は複数の予め決定された状態について、ニューラルネットワークモジュールを訓練する工程であり、このニューラルネットワークが、遺伝子発現データを受け取るために構成された、1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の状態を出力するために構成された1個又は複数の出力節点を含む、出力層；及び、遺伝子発現データと1種又は複数の状態の間の関係を表わすルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備え、該適応構成要素が、入力データを基にしたルールベース層中の選択された2個又はそれよりも多いルール節点を集合するように配置されている、工程；
- c) 適応構成要素に、ルールベース層内の選択された2個又はそれよりも多いルール節点を集合させる工程；
- d) ルールベース層からルールを抽出する工程であり、該ルールが、遺伝子発現データと1種又は複数の状態の間の関係を表わしている、工程；
- e) 抽出されたルールから、過剰発現された又は過小発現された遺伝子を同定する工程であり、この遺伝子は、1種又は複数の状態に連結された過剰発現された又は過小発現された遺伝子を識別するセットを表わしている工程を含む、方法。

【請求項7】

1種又は複数の状態に連結された過剰発現された又は過小発現された遺伝子を識別するセットを選択するシステムであり：

- a) 1種又は複数の予め決定された状態へカテゴリ化された遺伝子発現データの1種又は複数のセットを受け取ることが可能な、入力；
- b) 該遺伝子発現データ及び1種又は複数の予め決定された状態を受け取るように適合された、ニューラルネットワークモジュールであり、このニューラルネットワークは、遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の状態を出力するために構成された1個又は複数の出力節点を含む、出力層；及び、遺伝子発現データと1種又は複数の状態の間の関係を表わしているルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備え、該適応構成要素が、入力データを基にしたルールベース層中の選択された2個又はそれよりも多いルール節点を集合するように配置されている、モジュール；
- c) ルールベース層からルールを抽出するための、抽出構成要素であり、該ルールが、遺伝子発現データと1種又は複数の状態の間の関係を表わしている、要素；及び
- d) 抽出されたルールから過剰発現された又は過小発現された遺伝子を同定するための、識別子であり、これらの遺伝子が、1種又は複数の状態に連結された過剰発現された又は過小発現された遺伝子を識別するセットを表わしている、識別子を備える、システム。

【請求項8】

請求項7記載のように1種又は複数の状態に連結された過剰発現された又は過小発現された遺伝子を識別するセットを選択するための、システムであり、ここで抽出されたルールが、特異的遺伝子発現を、特定の状態に連結させる、システム。

【請求項9】

遺伝子発現セットを減少する方法であり：

- a) 1種又は複数の状態によりカテゴリ化された遺伝子発現データの1種又は複数のセットを提供する工程；
- b) 該遺伝子発現データ及び1種又は複数の状態のクラスについてニューラルネットワークモジュールを訓練する工程であり、このニューラルネットワークが、遺伝子発現データを

受け取るために構成された1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の状態を出力するために構成された1個又は複数の出力節点を含む、出力層；及び、遺伝子発現データと1種又は複数の状態の間の関係を表わしているルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備え、該適応構成要素が、入力データを基にしたルールベース層中の選択された2個又はそれよりも多いルール節点を集合するように配置されている、工程；  
c) 適応構成要素に、ルールベース層内の選択された2個又はそれよりも多いルール節点を集合させる工程；  
d) ルールベース層からルールを抽出する工程であり、該ルールが、遺伝子発現データと1種又は複数の状態の間の関係を表わしている、工程；  
e) 抽出されたルールから、遺伝子を同定する工程であり、この遺伝子が、1種又は複数の状態に連結された減少された遺伝子発現セットを表わしている工程を含む、方法。

【請求項10】

遺伝子発現セットを減少するシステムであり：

- a) 1種又は複数の状態によりカテゴリ化された遺伝子発現データの1種又は複数のセットを受け取る、入力手段；
- b) 該遺伝子発現データ及び1種又は複数の状態のクラスについて訓練可能なニューラルネットワークモジュールであり、このニューラルネットワークが、遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の状態を出力するために構成された1個又は複数の出力節点を含む、出力層；及び、遺伝子発現データと1種又は複数の状態の間の関係を表わしているルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備え、該適応構成要素が、入力データを基にしたルールベース層中の選択された2個又はそれよりも多いルール節点を集合するように配置されている、モジュール；
- d) ルールベース層からルールを抽出するように適合された、ルール抽出手段であり、該ルールが、遺伝子発現データと1種又は複数の状態の間の関係を表わしている、手段；
- e) 抽出されたルールから遺伝子を同定するように適合された識別子であり、この遺伝子は、1種又は複数の状態に連結された減少された遺伝子発現セットを表わす識別子を備える、システム。

【請求項11】

遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の状態を出力するために構成された1個又は複数の出力節点を含む、出力層；及び、遺伝子発現データと1種又は複数の状態の間の関係を表わしているルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備え、該適応構成要素が、遺伝子発現データを基にしたルールベース層中の選択された2個又はそれよりも多いルール節点を集合するように配置されているものを備える、ニューラルネットワークモジュール。

【請求項12】

請求項2記載の方法に従いニューラルネットワークから抽出された場合の、遺伝子発現データと1種又は複数の状態の間の関係を表わしている、ルール。

【請求項13】

状態を診断するためにニューラルネットワークを訓練する方法であり：

- a) 1種又は複数の状態によりカテゴリ化された遺伝子発現データを提供する工程；
- b) 遺伝子発現データ及び1種又は複数の状態についてニューラルネットワークモジュールを訓練する工程であり、ここでこのニューラルネットワークモジュールが、遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の状態を出力するために構成された1個又は複数の出力節点を含む、出力層；及び、遺伝子発現データと1種又は複数の状態の間の関係を表わしているルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備える、工程；

- c) リーブワンアウト法を使用し、データを検定する工程；
- d) 遺伝子発現データの入力を減少し、最良の検定精度を提供する工程；
- e) その入力層のように、減少された遺伝子発現データを受入れるために、ニューラルネットワークモジュールを修飾する工程；
- f) 修飾されたニューラルネットワークモジュールを訓練する工程；並びに
- g) 適応構成要素からルールを抽出する工程を少なくとも含む、方法。

【請求項14】

更に、減少工程からの方法を反復する工程を含む、請求項13記載の状態を診断するニューラルネットワークを訓練する方法。

【請求項15】

状態を診断するために、ニューラルネットワークを訓練するシステムであり：

- a) 1種又は複数の状態によりカテゴリ化された遺伝子発現データを提供する、受け取り可能な入力；
- b) 遺伝子発現データ及び1種又は複数の状態について訓練可能なニューラルネットワークモジュールであり、ここでこのニューラルネットワークモジュールが、遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の状態を出力するために構成された1個又は複数の出力節点を含む、出力層；及び、遺伝子発現データと1種又は複数の状態の間の関係を表わしているルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備える、モジュール；
- c) 最良の検定精度を提供するために、遺伝子発現データを減少するように操作可能な、入力遺伝子発現データ減少子；
- d) ニューラルネットワークモジュールを、その入力層のように減少された遺伝子発現データを受入れるように修飾するための、ニューラルネットワーク修飾子；並びに
- e) ルールを適応構成要素から抽出するための、ルール抽出手段を備える、システム。

【請求項16】

抽出されたルールは、人間が読取り可能な形で表わされる、請求項15記載の状態を診断するためにニューラルネットワークを訓練するシステム。

【請求項17】

ニューラルネットワークモジュールが更に、1種又は複数の状態への十分な連結を示さないルールベース層内の節点をプルーニングするために配置されているプルーニングアルゴリズムを含む、請求項15又は16記載の状態を診断するためのニューラルネットワークを訓練するシステム。

【請求項18】

遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の予後の転帰を出力するために構成された1個又は複数の出力節点を含む、出力層；及び、遺伝子発現データと1種又は複数の予後の転帰の間の関係を表わしているルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備える、ニューラルネットワークモジュール。

【請求項19】

遺伝子発現データと予後の転帰の間の関係を決定する、汎用方法であり：

- a) 予め決定された予後の転帰により分類された遺伝子発現データのセットを提供する工程；
- b) 該遺伝子発現データ及び予後の転帰について、ニューラルネットワークモジュールを訓練する工程であり、ここでこのニューラルネットワークモジュールが、遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の予後の転帰を出力するために構成された1個又は複数の出力節点を含む、出力層；及び、遺伝子発現データと1種又は複数の予後の転帰の間の関係を表わすルールベース層から1種又は複数のルールを抽出するために構成

された、適応構成要素を備える工程；

c) ルールベース層からルールを抽出する工程であり、該ルールが、遺伝子発現データと1種又は複数の予後の転帰の間の関係を表わす工程を少なくとも含む、方法。

【請求項20】

該患者の生物学的試料から抽出された遺伝子発現データを、遺伝子発現データと予後の転帰の間の関係を表わすルールと相関するために適合された相関エンジンを備え、該ルールが、請求項19記載の方法により決定される、患者の状態の予後判定をする汎用システム。

【請求項21】

遺伝子発現データと予後の転帰の間の関係を決定する汎用システムであり：

a) 予め決定された予後の転帰により分類された遺伝子発現データのセットを受け取るための、入力；

b) 該遺伝子発現データ及び予後の転帰について訓練可能なニューラルネットワークモジュールであり、ここでこのニューラルネットワークモジュールが、遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の予後の転帰を出力するために構成された1個又は複数の出力節点を含む、出力層；及び、遺伝子発現データと1種又は複数の予後の転帰の間の関係を表わすルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備えるモジュール；

c) ルールベース層からルールを抽出するために適合されたルール抽出子であり、該ルールが、遺伝子発現データと1種又は複数の予後の転帰の間の関係を表わしている抽出子を備える、システム。

【請求項22】

患者の状態の予後判定のための汎用方法であり、該患者の生物学的試料から抽出された遺伝子発現データを、遺伝子発現データと予後の転帰の間の関係を表わしているルールと相関することを含み、該ルールが、請求項19記載の方法により決定されている、方法。

【請求項23】

患者の状態の予後判定のための汎用システムであり、該患者の生物学的試料から抽出された遺伝子発現データを、遺伝子発現データと予後の転帰の間の関係を表わしているルールと相関するために適合された相関手段を備え、該ルールが、請求項19記載の方法により決定される、システム。

【請求項24】

1種又は複数の予後の転帰に連結された過剰発現又は過小発現された遺伝子を識別するセットを選択する方法であり：

a) 1種又は複数の予め決定された予後の転帰へカテゴリ化された遺伝子発現データの1種又は複数のセットを提供する工程；

b) 該遺伝子発現データ及び1種又は複数の予め決定された予後の転帰についてニューラルネットワークモジュールを訓練する工程であり、このニューラルネットワークが、遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の予後の転帰を出力するために構成された1個又は複数の出力節点を含む、出力層；及び、遺伝子発現データと1種又は複数の予後の転帰の間の関係を表わすルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備え、該適応構成要素が、入力データを基にしたルールベース層内の選択された2個又はそれよりも多いルール節点を集合するために配置されている、工程；

c) 適応構成要素に、ルールベース層内の選択された2個又はそれよりも多いルール節点を集合させる工程；

d) ルールベース層からルールを抽出する工程であり、該ルールが、遺伝子発現データと1種又は複数の予後の転帰の間の関係を表わしている、工程；

e) 抽出されたルールから、過剰発現された又は過小発現された遺伝子を同定する工程であ

り、この遺伝子が、1種又は複数の予後の転帰に連結された過剰発現された又は過小発現された遺伝子を識別するセットを表わしている工程を含む、方法。

【請求項25】

1種又は複数の予後の転帰に連結された過剰発現された又は過小発現された遺伝子を識別するセットを選択するシステムであり：

- a) 1種又は複数の予め決定された予後の転帰へカテゴリ化された遺伝子発現データの1種又は複数のセットを受け取ることが可能な、入力；
- b) 該遺伝子発現データ及び1種又は複数の予め決定された予後の転帰について訓練可能な、ニューラルネットワークモジュールであり、このニューラルネットワークは、遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の予後の転帰を出力するために構成された1個又は複数の出力節点を含む、出力層；及び、遺伝子発現データと1種又は複数の予後の転帰の間の関係を表わしているルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備え、該適応構成要素が、入力データを基にしたルールベース層中の選択された2個又はそれよりも多いルール節点を集合するように配置されている、モジュール；
- c) ルールベース層からルールを抽出するために適合された、ルール抽出子であり、該ルールが、遺伝子発現データと1種又は複数の予後の転帰の間の関係を表わしている、抽出子；並びに
- d) 抽出されたルールから過剰発現された又は過小発現された遺伝子を同定するための、識別子であり、この遺伝子が、1種又は複数の予後の転帰に連結された過剰発現された又は過小発現された遺伝子を識別するセットを表わしている、識別子を備える、システム。

【請求項26】

遺伝子発現セットを減少する方法であり：

- a) 1種又は複数の予後の転帰によりカテゴリ化された遺伝子発現データの1種又は複数のセットを提供する工程；
- b) 該遺伝子発現データ及び1種又は複数の状態のクラスについてニューラルネットワークモジュールを訓練する工程であり、このニューラルネットワークが、遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の予後の転帰を出力するために構成された1個又は複数の出力節点を含む、出力層；及び、遺伝子発現データと1種又は複数の予後の転帰の間の関係を表わしているルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備え、該適応構成要素が、入力データを基にしたルールベース層中の選択された2個又はそれよりも多いルール節点を集合するように配置されている、工程；
- c) 適応構成要素に、ルールベース層内の選択された2個又はそれよりも多いルール節点を集合させる工程；
- d) ルールベース層からルールを抽出する工程であり、該ルールが、遺伝子発現データと1種又は複数の予後の転帰の間の関係を表わしている、工程；
- e) 抽出されたルールから、遺伝子を同定する工程であり、この遺伝子が、1種又は複数の予後の転帰に連結された減少された遺伝子発現セットを表わしている、工程を含む、方法。

【請求項27】

遺伝子発現セットを減少するシステムであり：

- a) 1種又は複数の予後の転帰によりカテゴリ化された遺伝子発現データの1種又は複数のセットを受け取る、入力；
- b) 該遺伝子発現データ及び1種又は複数の状態のクラスについて訓練可能なニューラルネットワークモジュールであり、このニューラルネットワークが、遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の予後の転帰を出力するために構成された1個又は

複数の出力節点を含む、出力層；及び、遺伝子発現データと1種又は複数の予後の転帰の間の関係を表わしているルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備え、該適応構成要素が、入力データを基にしたルールベース層中の選択された2個又はそれよりも多いルール節点を集合するように配置されている、工程；

d)ルールベース層からルールを抽出するように適合された、ルール抽出子であり、該ルールが、遺伝子発現データと1種又は複数の予後の転帰の間の関係を表わしている、抽出子；

e)抽出されたルールから遺伝子を同定するように適合された同定手段であり、この遺伝子は、1種又は複数の予後の転帰に連結された減少された遺伝子発現セットを表わす手段を備える、システム。

【請求項28】

遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の予後の転帰を出力するために構成された1個又は複数の出力節点を含む、出力層；及び、遺伝子発現データと1種又は複数の予後の転帰の間の関係を表わしているルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備え、該適応構成要素が、遺伝子発現データを基にしたルールベース層中の選択された2個又はそれよりも多いルール節点を集合するように配置されている、ニューラルネットワークモジュール。

【請求項29】

請求項26記載の方法でニューラルネットワークから抽出された場合に、遺伝子発現データと1種又は複数の予後の転帰の間の関係を表わしているルール。

【請求項30】

予後の転帰を提供するニューラルネットワークを訓練する方法であり：

- a)1種又は複数の予後の転帰によりカテゴリ化された遺伝子発現データを提供する工程；
- b)遺伝子発現データ及び1種又は複数の予後の転帰についてニューラルネットワークモジュールを訓練する工程であり、ここでこのニューラルネットワークモジュールが、遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の予後の転帰を出力するために構成された1個又は複数の出力節点を含む、出力層；及び、遺伝子発現データと1種又は複数の予後の転帰の間の関係を表わしているルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備える、工程；
- c)リーブワンアウト法を使用し、データを検定する工程；
- d)遺伝子発現データの入力を減少し、最良の検定精度を提供する工程；
- e)その入力層のように、減少された遺伝子発現データを受入れるために、ニューラルネットワークモジュールを修飾する工程；
- f)修飾されたニューラルネットワークモジュールを訓練する工程；並びに
- g)適応構成要素からルールを抽出する工程を少なくとも含む、方法。

【請求項31】

更に、減少工程からの方法を反復する工程を含む、請求項30記載の予後の転帰を提供するニューラルネットワークを訓練する方法。

【請求項32】

予後の転帰を提供するために、ニューラルネットワークを訓練するシステムであり：

- a)1種又は複数の予後の転帰によりカテゴリ化された遺伝子発現データを受け取るための、入力；
- b)遺伝子発現データ及び1種又は複数の予後の転帰について訓練可能なニューラルネットワークモジュールであり、ここでこのニューラルネットワークモジュールが、遺伝子発現データを受け取るために構成された1個又は複数の入力節点を含む、入力層；1個又は複数のルール節点を含む、ルールベース層；1種又は複数の予後の転帰を出力するために構成された1個又は複数の出力節点を含む、出力層；及び、遺伝子発現データと1種又は複数の

予後の転帰の間の関係を表わしているルールベース層から1種又は複数のルールを抽出するために構成された、適応構成要素を備える、モジュール；

c)最良の検定精度を提供するために、入力遺伝子発現データを減少する、減少子；

d)ニューラルネットワークモジュールを、その入力層のように減少された遺伝子発現データを受入れるように修飾するために適合された、修飾子；並びに

e)ルールを適応構成要素から抽出するための、ルール抽出子を備える、システム。

【請求項33】

抽出されたルールは、人間が読取り可能な形で表わされる、請求項32記載の、予後の転帰を提供するために、ニューラルネットワークを訓練するシステム。

【請求項34】

ニューラルネットワークモジュールが更に、1種又は複数の予後の転帰への十分な連結を示さない、ルールベース層内の節点をプルーニングするために配置されているプルーニングアルゴリズムを含む、請求項32又は33記載の予後の転帰を提供するためのニューラルネットワークを訓練するシステム。

【請求項35】

プロファイリングされた疾患が、乳腺癌、前立腺腺癌、肺腺癌、結腸直腸腺癌、リンパ腫、膀胱移行細胞癌、黒色腫、子宮腺癌、白血病、腎細胞癌、膵臓腺癌、卵巣腺癌、胸膜中皮腫、及び中枢神経系癌からなる群より選択される、請求項32～34のいずれか1項記載の予後の転帰を提供するためのニューラルネットワークを訓練するシステム。

【手続補正2】

【補正対象書類名】明細書

【補正対象項目名】0133

【補正方法】変更

【補正の内容】

【0133】

本発明は更に以下を提供する：

a)S39329、U89942、X15880、AB006781、HG2261-HT2351、M21119、M28130、U01317、AA454214、H08939、W26436、AA071089、AA236458、AA262710、AA284767、AA447769、AA449419、AA452928、AA453034、AA479299、AA488178及びD59321からなる群より選択される、少なくとも1種の遺伝子に選択的に連結する少なくとも1種の核酸配列を含む、乳癌を診断するための診断キット。

b)U01317、AA364267、AA427468、AA428172、AA482319、AA017146、AA070437、AA193204、AA234561、AA237017、AA259135、AA291629、AA405049、AA437235、AA453304、AA479096、AA482224及びAA599991からなる群より選択される少なくとも1種の遺伝子に選択的に連結する少なくとも1種の核酸配列を含む、前立腺癌を診断するための診断キット。

c)M89957、M91438、S39329、X15880、M21119、X00437、AA283620、AA364267、AA428172、AA454214、H46792、M89957、W26996、AA088851、AA193204、AA236458、AA447769、AA460849及びD59321からなる群より選択される少なくとも1種の遺伝子に選択的に連結する少なくとも1種の核酸配列を含む、肺癌を診断するための診断キット。

d)D83715、L19527、M22632、M25280、M62302、M84711、M91036、U09953、U14968、X62691、U01317、AA454214、H46792、M62302、N94832、N98707、AA161292、AA165369、AA447769；AA452928、AA453034及びAA599991からなる群より選択される少なくとも1種の遺伝子に選択的に連結する少なくとも1種の核酸配列を含む、結腸直腸癌を診断するための診断キット。

e)D64142、M91036、J03801、AA454214、H46792、J03544、N94832、R10770、AA005262、AA024658、AA284767、AA348466及びAA447769からなる群より選択される少なくとも1種の遺伝子に選択的に連結する少なくとも1種の核酸配列を含む、リンパ腫を診断するための診断キット。

f)X15880、AA056208、AA070437、AA071089及びAA479096からなる群より選択される少なくとも1種の遺伝子に選択的に連結する少なくとも1種の核酸配列を含む、膀胱癌を診断する

ための診断キット。

g) D79205、L44416、AA017146、AA026054、AA056208、AA233257及びAA259135からなる群より選択される少なくとも1種の遺伝子に選択的に連結する少なくとも1種の核酸配列を含む、黒色腫を診断するための診断キット。

h) D79205、D89667、HG613-HT613、L19527、M31520、M84711、S39329、U09953、U14968、X62691、X98085、M31520、N94832、AA284767、AA608546及びD59321からなる群より選択される少なくとも1種の遺伝子に選択的に連結する少なくとも1種の核酸配列を含む、子宮癌を診断するための診断キット。

i) L19527、M31520、M84711、U09953、U14968、X62691、X76732、X02761、M31520、J03464及びN98707からなる群より選択される少なくとも1種の遺伝子に選択的に連結する少なくとも1種の核酸配列を含む、白血病を診断するための診断キット。

j) M89957、U01317、AA454214、M89957、N94832、AA193204、AA259135、AA284767、AA447769及びD59321からなる群より選択される少なくとも1種の遺伝子に選択的に連結する少なくとも1種の核酸配列を含む、腎臓癌を診断するための診断キット。

k) M55998、J03801、J03544、AA193204、AA234561、AA287131、AA446461及びAA479096からなる群より選択される少なくとも1種の遺伝子に選択的に連結する少なくとも1種の核酸配列を含む、膵癌を診断するための診断キット。

l) M89957、M91036、M91438、X74142、AB006781、HG2261-HT2351、M21119、X00437、X97267、AA364267、AA454214、H46792、M89957、N94832、R61154、W26996、X74142、AA262340、AA424146、AA447769、AA455111、及びAA464844からなる群より選択される少なくとも1種の遺伝子に選択的に連結する少なくとも1種の核酸配列を含む、卵巣癌を診断するための診断キット。

m) 胸膜癌 J02923、J03040、M29540、M89957、M91036、M91438、S39329、AB006781、HG2261-HT2351、U22376、M21119、M28130、X12530、U01317、AA364267、J03544、M89957、AA056208、AA234561、AA237017、AA253043及びAA291629からなる群より選択される少なくとも1種の遺伝子に選択的に連結する少なくとも1種の核酸配列を含む、胸膜癌を診断するための診断キット。

n) 中枢神経系 M29540、M89957、S39329、M28130、X00437、AA363338、AA393317、L44416、M89957、N71215、R70976、T09191、W28214、W81301及びAA259135からなる群より選択される少なくとも1種の遺伝子に選択的に連結する少なくとも1種の核酸配列を含む、中枢神経系癌を診断するための診断キット。

前記診断キットにおいて、このキットは更に、試料中に存在する遺伝子発現産物の量を定量測定することが可能な少なくとも1種の核酸配列の量を備える。

本発明は更に、以下を提供する：

a) S39329、U89942、X15880、AB006781、HG2261-HT2351、M21119、M28130、U01317、AA454214、H08939、W26436、AA071089、AA236458、AA262710、AA284767、AA447769、AA449419、AA452928、AA453034、AA479299、AA488178及びD59321からなる群より選択される少なくとも1種の遺伝子で表された核酸配列から発現されたペプチドに選択的に結合することが可能なりガンドを含む少なくとも1種の核酸配列を含む、乳癌を診断するための診断キット。

b) U01317、AA364267、AA427468、AA428172、AA482319、AA017146、AA070437、AA193204、AA234561、AA237017、AA259135、AA291629、AA405049、AA437235、AA453304、AA479096、AA482224及びAA599991からなる群より選択される少なくとも1種の遺伝子で表された核酸配列から発現されたペプチドに選択的に結合することが可能なりガンドを含む少なくとも1種の核酸配列を含む、前立腺癌を診断するための診断キット。

c) M89957、M91438、S39329、X15880、M21119、X00437、AA283620、AA364267、AA428172、AA454214、H46792、M89957、W26996、AA088851、AA193204、AA236458、AA447769、AA460849及びD59321からなる群より選択される少なくとも1種の遺伝子で表された核酸配列から発現されたペプチドに選択的に結合することが可能なりガンドを含む少なくとも1種の核酸配列を含む、肺癌を診断するための診断キット。

- d) D83735、L19527、M22632、M25280、M62302、M84711、M91036、U09953、U14968、X62691、U01317、AA454214、H46792、M62302、N94832、N98707、AA161292、AA165369、AA447769；AA452928、AA453034及びAA599991からなる群より選択される少なくとも1種の遺伝子で表された核酸配列から発現されたペプチドに選択的に結合することが可能なりガンドを含む少なくとも1種の核酸配列を含む、結腸直腸癌を診断するための診断キット。
- e) D64142、M91036、J03801、AA454214、H46792、J03544、N94832、R10770、AA005262、AA024658、AA284767、AA348466及びAA447769からなる群より選択される少なくとも1種の遺伝子で表された核酸配列から発現されたペプチドに選択的に結合することが可能なりガンドを含む少なくとも1種の核酸配列を含む、リンパ腫を診断するための診断キット。
- f) X15880、AA056208、AA070437、AA071089及びAA479096からなる群より選択される少なくとも1種の遺伝子で表された核酸配列から発現されたペプチドに選択的に結合することが可能なりガンドを含む少なくとも1種の核酸配列を含む、膀胱癌を診断するための診断キット。
- g) D79205、L44416、AA017146、AA026054、AA056208、AA233257及びAA259135からなる群より選択される少なくとも1種の遺伝子で表された核酸配列から発現されたペプチドに選択的に結合することが可能なりガンドを含む少なくとも1種の核酸配列を含む、黒色腫を診断するための診断キット。
- h) D79205、D89667、HG613-HT613、L19527、M31520、M84711、S39329、U09953、U14968、X62691、X98085、M31520、N94832、AA284767、AA608546及びD59321からなる群より選択される少なくとも1種の遺伝子で表された核酸配列から発現されたペプチドに選択的に結合することが可能なりガンドを含む少なくとも1種の核酸配列を含む、子宮癌を診断するための診断キット。
- i) L19527、M31520、M84711、U09953、U14968、X62691、X76732、X02761、M31520、J03464及びN98707からなる群より選択される少なくとも1種の遺伝子で表された核酸配列から発現されたペプチドに選択的に結合することが可能なりガンドを含む少なくとも1種の核酸配列を含む、白血病を診断するための診断キット。
- j) M89957、U01317、AA454214、M89957、N94832、AA193204、AA259135、AA284767、AA447769及びD59321からなる群より選択される少なくとも1種の遺伝子で表された核酸配列から発現されたペプチドに選択的に結合することが可能なりガンドを含む少なくとも1種の核酸配列を含む、腎癌を診断するための診断キット。
- k) M55998、J03801、J03544、AA193204、AA234561、AA287131、AA446461及びAA479096からなる群より選択される少なくとも1種の遺伝子で表された核酸配列から発現されたペプチドに選択的に結合することが可能なりガンドを含む少なくとも1種の核酸配列を含む、膵癌を診断するための診断キット。
- l) M89957、M91036、M91438、X74142、AB006781、HG2261-HT2351、M21119、X00437、X97267、AA364267、AA454214、H46792、M89957、N94832、R61154、W26996、X74142、AA262340、AA424146、AA447769、AA455111及びAA464844からなる群より選択される少なくとも1種の遺伝子で表された核酸配列から発現されたペプチドに選択的に結合することが可能なりガンドを含む少なくとも1種の核酸配列を含む、卵巣癌を診断するための診断キット。
- m) 胸膜癌 J02923、J03040、M29540、M89957、M91036、M91438、S39329、AB006781、HG2261-HT2351、U22376、M21119、M28130、X12530、U01317、AA364267、J03544、M89957、AA056208、AA234561、AA237017、AA253043及びAA291629からなる群より選択される少なくとも1種の遺伝子で表された核酸配列から発現されたペプチドに選択的に結合することが可能なりガンドを含む少なくとも1種の核酸配列を含む、胸膜癌を診断するための診断キット。
- n) 中枢神経系 M29540、M89957、S39329、M28130、X00437、AA363338、AA393317、L44416、M89957、N71215、R70976、T09191、W28214、W81301及びAA259135からなる群より選択される少なくとも1種の遺伝子で表された核酸配列から発現されたペプチドに選択的に結合することが可能なりガンドを含む少なくとも1種の核酸配列を含む、中枢神経系癌を診断するための診断キット。

このリガンドは、抗体又は免疫分子であることができる。

この免疫分子は、fab断片であることができる。

このリガンドは、特に関連するペプチドに結合することが可能である。

このキットは更に、試料中に存在するペプチドの量を定量測定することができるリガンドの量を備えている。

本発明は、患者が、乳房-、前立腺-、膀胱-、白血病-、リンパ腫-、中枢神経系-、肺-、結腸直腸- 黒色腫、子宮-、腎細胞-、膵-、卵巣-、及び胸膜癌からなる群より選択される癌に罹患しているかどうかを診断する方法であり：

a)患者から試料を単離する工程；


b)この試料が、本明細書の説明と合わせて読む問題のそれぞれの癌の各々についての図面に説明された、少なくとも1種の遺伝子の疾患プロファイルと一致する発現レベルを含むかどうかを決定する工程を含む方法も提供する。

この試料は、組織であることができる。

この組織は、癌であることが疑われる。

疾患プロファイリングの方法及びシステムの好ましい形は、添付図面を参照し、ここに説明する：

## 【 国際調査報告 】

INTERNATIONAL SEARCH REPORT		International application No. PCT/NZ03/00045		
<b>A. CLASSIFICATION OF SUBJECT MATTER</b>				
Int. Cl. 7: G06N 3/04, G06F 19/00, 159:00, C12N 15/11, C12Q 1/68, C07K 16/30				
According to International Patent Classification (IPC) or to both national classification and IPC				
<b>B. FIELDS SEARCHED</b>				
Minimum documentation searched (classification system followed by classification symbols) IPC: G06F, G06N				
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched				
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) WPAT: "gene expression" and "neural network" PubMed: "gene expression" and "neural network"				
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>				
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.		
P, Y	Ando, T. et al., Fuzzy Neural Network Applied to Gene Expression Profiling for Predicting the Prognosis of Diffuse Large B-cell Lymphoma, Jpn. J. Cancer Res. 93, 1207-1212, November 2002 See whole document	1 to 35		
P, Y	US-2002/0156587-A1 (WOOLF et al.) 24 October 2002 (24-10-02) See whole document	1 to 35		
Y	Kim, J. et al., HyFIS: Adaptive Neuro-Fuzzy Inference Systems and Their Application to Nonlinear Dynamical Systems, Neural Netw. 1999 November; 12(9): 1301-1319 See whole document	1 to 35		
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C <input checked="" type="checkbox"/> See patent family annex				
<p>* Special categories of cited documents:</p> <table border="0" style="width: 100%;"> <tr> <td style="width: 50%; vertical-align: top;"> <p>"A" document defining the general state of the art which is not considered to be of particular relevance</p> <p>"E" earlier application or patent but published on or after the international filing date</p> <p>"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)</p> <p>"O" document referring to an oral disclosure, use, exhibition or other means</p> <p>"P" document published prior to the international filing date but later than the priority date claimed</p> </td> <td style="width: 50%; vertical-align: top;"> <p>"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</p> <p>"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</p> <p>"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</p> <p>"&amp;" document member of the same patent family</p> </td> </tr> </table>			<p>"A" document defining the general state of the art which is not considered to be of particular relevance</p> <p>"E" earlier application or patent but published on or after the international filing date</p> <p>"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)</p> <p>"O" document referring to an oral disclosure, use, exhibition or other means</p> <p>"P" document published prior to the international filing date but later than the priority date claimed</p>	<p>"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</p> <p>"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</p> <p>"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</p> <p>"&amp;" document member of the same patent family</p>
<p>"A" document defining the general state of the art which is not considered to be of particular relevance</p> <p>"E" earlier application or patent but published on or after the international filing date</p> <p>"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)</p> <p>"O" document referring to an oral disclosure, use, exhibition or other means</p> <p>"P" document published prior to the international filing date but later than the priority date claimed</p>	<p>"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</p> <p>"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</p> <p>"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</p> <p>"&amp;" document member of the same patent family</p>			
Date of the actual completion of the international search 6 June 2003		Date of mailing of the international search report 19 JUN 2003		
Name and mailing address of the ISA/AU AUSTRALIAN PATENT OFFICE PO BOX 200, WODEN ACT 2606, AUSTRALIA E-mail address: pct@ipaustalia.gov.au Facsimile No. (02) 6285 3929		Authorized officer  MICHAEL HARDY Telephone No : (02) 6283 2547		

## INTERNATIONAL SEARCH REPORT

International application No. PCT/NZ03/00045
---

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	Ichimura, T., et al., Extraction of fuzzy rules using neural networks with structure level adaptation - verification to the diagnosis of hepatobiliary disorders, International Journal of Bio-Medical Computing 40 (1995) 139-146 See whole document	1 to 35
P, A	WO-02/47007-A2 (PHASE IT INTELLIGENT SOLUTIONS AG) 13 June 2002 (13-06-02) See whole document	

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/NZ03/00045

**Box I 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 : **36 to 50**  
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:  
Claims 36 to 50 include (or are dependent on claims that include) the defining phrase, "... as set forth in the Figures as read with the description herein." It is impossible to determine to what extent the characterising features of the neural network systems and methods of claims 1 to 35 are to be regarded as definitive of the invention(s) claimed in claims 36 to 50.
  
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 II 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:

1.  As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims
2.  As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
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.
- No protest accompanied the payment of additional search fees.

**INTERNATIONAL SEARCH REPORT**  
Information on patent family membersInternational application No.  
**PCT/NZ03/00045**

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

Patent Document Cited in Search Report		Patent Family Member	
US	2002015658	NONE	
WO	200247007	AU	200228000
END OF ANNEX			

## フロントページの続き

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

(74) 代理人 100082898

弁理士 西山 雅也

(72) 発明者 リーブ, アンソニー エドムンド

ニュージーランド国, ダニーディン, マオリ ヒル, コモ ストリート 2 2

(72) 発明者 フツィック, マチアス エルビン

ニュージーランド国, ダニーディン, マナー プレイス 5 2

(72) 発明者 サリバン, マイケル ジェームズ

ニュージーランド国, クライストチャーチ, コッツモア クローズ 9

(72) 発明者 カサボフ, ニコラ キリロフ

ニュージーランド国, オークランド, ミッション ベイ, ニヒル クレッセント 4 7 シー

(72) 発明者 ギルフォード, バリー ジョン

ニュージーランド国, ダニーディン, イースト テイエリ, リッカートン ロード イースト 3 8

F ターム(参考) 4B063 QA18 QA20 QQ42 QQ52 QR55 QR82 QS39

专利名称(译)	基因表达数据在自适应学习系统中的应用		
公开(公告)号	<a href="#">JP2005521138A</a>	公开(公告)日	2005-07-14
申请号	JP2003577211	申请日	2003-03-17
[标]申请(专利权)人(译)	环太平洋生物技术有限公司		
申请(专利权)人(译)	太平洋边缘生物科技Rimitido		
[标]发明人	リーブアンソニーエドモンド フツィックマチアスエルビン サリバンマイケルジェームズ カサボフニコラキリロフ ギルフォードパリージョン		
发明人	リーブ,アンソニー エドモンド フツィック,マチアス エルビン サリバン,マイケル ジェームズ カサボフ,ニコラ キリロフ ギルフォード,パリー ジョン		
IPC分类号	G01N33/53 C12N15/00 C12Q1/68 G01N33/574 G06F19/00 G06N3/00 G06N3/04		
CPC分类号	G06N3/0436 G06N3/004 G06N3/0427 G16B25/00		
FI分类号	G06N3/00.540 G01N33/53.D G01N33/574.Z C12N15/00.Z C12Q1/68.A		
F-TERM分类号	4B063/QA18 4B063/QA20 4B063/QQ42 4B063/QQ52 4B063/QR55 4B063/QR82 4B063/QS39		
代理人(译)	青木 笃 石田 敬 中村弘 西山雅也		
优先权	517817 2002-03-15 NZ		
外部链接	<a href="#">Espacenet</a>		

摘要(译)

提供了一种神经网络模块。它包括输入层，该输入层包括被配置为接收基因表达数据的一个或多个输入节点。它还包括输出层，该输出层包括一个或多个规则节点的规则基础层和被配置为输出一个或多个状态的一个或多个输出节点。它还包括自适应组件，其被配置为从表示基因表达数据与一个或多个状态之间的关系的规则基础层提取一个或多个规则。除了使用该模块的方法和系统之外，还公开了利用该系统的特定简档。

