

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2006-139762

(P2006-139762A)

(43) 公開日 平成18年6月1日(2006.6.1)

(51) Int. Cl.

G06T 11/20 (2006.01)

F I

G06T 11/20 120

テーマコード (参考)

5B080

審査請求 未請求 請求項の数 27 O L 外国語出願 (全 33 頁)

(21) 出願番号 特願2005-296120 (P2005-296120)  
 (22) 出願日 平成17年10月11日 (2005.10.11)  
 (31) 優先権主張番号 10/964524  
 (32) 優先日 平成16年10月12日 (2004.10.12)  
 (33) 優先権主張国 米国 (US)  
 (31) 優先権主張番号 11/124500  
 (32) 優先日 平成17年5月6日 (2005.5.6)  
 (33) 優先権主張国 米国 (US)

(71) 出願人 399117121  
 アジレント・テクノロジーズ・インク  
 AGILENT TECHNOLOGIES, INC.  
 アメリカ合衆国カリフォルニア州パロアルト  
 ページ・ミル・ロード 395  
 395 Page Mill Road  
 Palo Alto, California  
 U. S. A.

(74) 代理人 100087642

弁理士 古谷 聡

(74) 代理人 100076680

弁理士 溝部 孝彦

(74) 代理人 100121061

弁理士 西山 清春

最終頁に続く

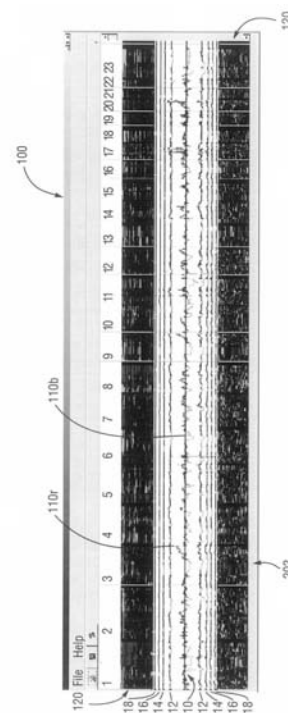
(54) 【発明の名称】 多くのグラフの集合についてフォーカス及びコンテキストを表示するためのシステム、ツール及び方法

(57) 【要約】 (修正有)

【課題】 データ（例えば、線グラフ群として表示し得るデータ）のより大きな集合をコンパクトなグラフの形態で視覚化する。

【解決手段】 コンテキスト及びフォーカスを提供するためにグラフの集合を視覚化するための方法で、少なくとも1つの圧縮されたグラフをズームして、圧縮されたグラフのスケールよりも第2の軸に沿ってより大きなスケールを有するその少なくとも1つのグラフを作成することができ、ズームされたグラフ（を圧縮されたグラフとともに表示することができる。この場合、そのズームされたグラフ（上記の1または複数）は圧縮されたグラフに対して同じ順序で、及び、ズームされる前の互いの順序と同じ順序で表示される。

【選択図】 図2



**【特許請求の範囲】****【請求項 1】**

コンテキスト及びフォーカスを提供するためにグラフの集合を視覚化する方法であって、各グラフを視覚化物(102)の1つの軸に沿って整列させることができ、

前記1つの軸に対して全てのグラフを整列させるステップと、

前記第1の軸に垂直な第2の軸に沿った方向において前記グラフの視覚化物を圧縮するステップと、

前記グラフの少なくとも1つを拡大して、前記圧縮されたグラフのスケールよりも前記第2の軸に沿ってより大きなスケールを有する前記少なくとも1つのグラフの視覚化物(110)を作成するステップと、

前記少なくとも1つの拡大されたグラフ(110)及び前記圧縮されたグラフ(120)を表示するステップであって、該少なくとも1つの拡大されたグラフを、前記圧縮されたグラフに対して同じ順序で、かつ、前記拡大するステップの前に存在していた時と同じ互いの順序で表示するステップ

を含む、方法。

**【請求項 2】**

前記拡大するステップが、選択したグラフ(110)を前記第2の軸に対して第1の倍率に拡大するステップと、該選択したグラフの少なくとも一方の側で該選択したグラフに隣接する少なくとも1つのグラフ(112)を、前記第2の軸に対して前記第1の倍率よりは低い倍率に拡大するステップを含む、請求項1に記載の方法。

**【請求項 3】**

前記選択したグラフの少なくとも一方の側に隣接する複数のグラフ(112、114、116、118)を拡大し、該選択したグラフ(110)に最も近いグラフ(112)を、圧縮されたままの状態のグラフに最も近い前記複数のグラフのうちの1つのグラフ(118)よりも大きく拡大する、請求項1又は2に記載の方法。

**【請求項 4】**

前記複数のグラフが、前記選択したグラフに最も近い前記グラフ(112)から圧縮されたままの状態の前記グラフに最も近い前記グラフ(118)へと倍率の度合が低くなる、請求項3に記載の方法。

**【請求項 5】**

前記少なくとも1つのグラフの1部分を拡大して、その1部分の視覚化物が、それぞれ、当該グラフの残りの部分よりも前記第1の軸に沿って大きなスケールを有するようにするステップをさらに含む、請求項1～4のいずれかに記載の方法。

**【請求項 6】**

1部分を拡大する前記ステップが、各グラフの同じ部分を前記第1の軸に対して拡大するステップを含む、請求項5に記載の方法。

**【請求項 7】**

前記グラフが線グラフを含む、請求項1～6のいずれかに記載の方法。

**【請求項 8】**

前記グラフがマイクロアレイデータに基づくプロットである、請求項1～7のいずれかに記載の方法。

**【請求項 9】**

前記グラフがaCGHデータに基づくプロットからなる、請求項1～7のいずれかに記載の方法。

**【請求項 10】**

移動平均、Z-スコア、及びp-値の少なくとも1つを計算することにより前記グラフを生成するステップをさらに含む、請求項1～9のいずれかに記載の方法。

**【請求項 11】**

前記グラフが散布図を含む、請求項1～7のいずれかに記載の方法。

**【請求項 12】**

10

20

30

40

50

グラフの集合を１つの全体的なビューで視覚化する一方で、同時に少なくとも１つのグラフについての詳細を表示する能力を提供するために使用されるユーザーインターフェースであって、

圧縮された形態でグラフの集合を視覚化するように構成された表示部（１００）と、  
圧縮された形態のグラフの集合における少なくとも１部の視覚化物（１０２）からグラフを選択するための機能を有し、

前記ユーザーインターフェースは、少なくとも前記選択されたグラフ（１１０）を拡大して、前記圧縮された視覚化物の圧縮軸に沿った前記圧縮されたグラフのスケールよりも前記圧縮軸に沿った大きなスケールを有する前記少なくとも選択されたグラフの視覚化物を作成し、少なくとも前記選択されたグラフは、圧縮された形態で存在していたときと同じ順序で表示されるが、拡大されなかったグラフ（１２０）は、拡大されたグラフに隣接して表示されることからなる、ユーザーインターフェース。

10

【請求項１３】

前記選択されたグラフ（１１０）が拡大されて、前記圧縮軸に対して第１の倍率で表示され、前記選択されたグラフの少なくとも一方の側において、前記選択されたグラフに隣接した少なくとも１つのグラフ（１１２）が、前記圧縮軸に対して前記第１の倍率よりも小さい倍率で拡大されて表示される、請求項１２に記載のユーザーインターフェース。

【請求項１４】

前記視覚化物をスクロールするための機能（４００，４０２）をさらに含むユーザーインターフェースであって、前記選択されたグラフ（１１０）を、前記グラフの１つから次の隣接するグラフへと２方向のうちのいずれかの方向にスクロールすることにより変化させ、前記機能は、前記グラフを連続的にスクロールし、各グラフがスクロールされるときにその各グラフを拡大するとともに、スクロールされない各グラフを再度圧縮するように構成されることからなる、請求項１２又は１３に記載のユーザーインターフェース。

20

【請求項１５】

前記選択されたグラフ（１１０）の１部分が、前記選択されたグラフの残りの部分のスケールに対して、第２の軸に沿ってより大きいスケールを有するように拡大するための機能をさらに含み、前記第２の軸が前記圧縮軸に対して垂直である、請求項１２～１４のいずれかに記載のユーザーインターフェース。

30

【請求項１６】

前記グラフの各々が前記視覚化物の１つの軸に沿って整列され、前記グラフの集合の全てのグラフが前記表示部上で単一の視覚化物中に表示される、請求項１２～１５のいずれかに記載のユーザーインターフェース。

【請求項１７】

前記グラフが染色体位置と相互に関連する生物学的データのプロットである、請求項１２～１６のいずれかに記載のユーザーインターフェース。

【請求項１８】

前記グラフを類似性により並べ替えるための機能をさらに有し、該並べ替え機能により生成された結果に従って新しい順序のグラフを表示する、請求項１２～１７のいずれかに記載のユーザーインターフェース。

40

【請求項１９】

前記表示部が生物の染色体の図式的視覚化物（５０６）を含むユーザーインターフェースであって、前記グラフの集合は視覚化された染色体との相関にしたがって分割され、前記インターフェースは、各染色体に隣接する圧縮されたグラフ（５０４）のサブ集合を視覚化し、各サブ集合は、該サブ集合の前記視覚化物に隣接した染色体と相互に関連するグラフをそれぞれ含む、請求項１２～１８のいずれかに記載のユーザーインターフェース。

【請求項２０】

前記各グラフにおけるデータが、該データと隣接して表示される染色体の位置と位置的に対応するように表示される、請求項１９に記載のユーザーインターフェース。

50

## 【請求項 2 1】

前記生物学的データがマイクロアレイデータを含む、請求項 1 2 ~ 2 0 のいずれかに記載のユーザーインターフェース。

## 【請求項 2 2】

前記マイクロアレイデータが a C G H データを含む、請求項 2 1 に記載のユーザーインターフェース。

## 【請求項 2 3】

選択された染色体の視覚化物の拡大された視覚化物 ( 5 0 6 m )、及び前記選択された染色体に対応するグラフのサブ集合の隣接する視覚化物 ( 5 0 4 e m ) を表示するための機能をさらに有する、請求項 1 9 ~ 2 1 のいずれかに記載のユーザーインターフェース。

10

## 【請求項 2 4】

前記圧縮された集合の視覚化物からグラフを選択するための前記機能が、前記拡大された視覚化物における前記サブ集合に対して動作可能である、請求項 2 3 に記載のユーザーインターフェース。

## 【請求項 2 5】

前記拡大された視覚化物 ( 5 0 4 e m ) 中の選択されたグラフの一部を拡大して、前記選択されたグラフの残りの部分のスケールに対して、第 2 の軸に沿ってより大きなスケールを有するようにするための機能をさらに有し、前記第 2 の軸が前記圧縮軸に対して垂直である、請求項 2 4 又は 2 5 に記載のユーザーインターフェース。

## 【請求項 2 6】

前記拡大された視覚化物をスクロールするための機能をさらに有し、選択されたグラフを前記グラフの 1 つから次の隣接するグラフへと 2 方向のうちのいずれかの方向にスクロールすることにより変化させ、前記機能は、前記グラフを連続的にスクロールして、各グラフがスクロールされるときにその各グラフを拡大すると共に、スクロールされない各グラフを再度圧縮するように構成される、請求項 2 3 ~ 2 5 のいずれかに記載のユーザーインターフェース。

20

## 【請求項 2 7】

コンテキスト及びフォーカスを提供するためにグラフの集合を視覚化するための 1 以上の命令のシーケンスを担持するコンピューター読み取り可能な媒体であって、各グラフを視覚化物の 1 つの軸に沿って整列させることができ、1 つ以上のプロセッサによる 1 つ以上の命令のシーケンスの実行により、前記 1 つ以上のプロセッサが以下のステップを実行するようになっている、コンピューター読み取り可能な媒体。

30

前記 1 つの軸に対して垂直な第 2 の軸に沿った方向に前記グラフの視覚化物を圧縮するステップと、

少なくとも 1 つの前記グラフを拡大して、前記圧縮されたグラフのスケールよりも前記第 2 の軸に沿ってより大きなスケールを有する前記少なくとも 1 つのグラフの視覚化物 ( 1 1 0 ) を作成するステップと、

前記少なくとも 1 つの拡大されたグラフ ( 1 1 0 ) 及び前記圧縮されたグラフ ( 1 2 0 ) を表示するステップであって、前記少なくとも 1 つの拡大されたグラフを、前記圧縮されたグラフに対して同じ位置で、かつ、前記拡大するステップの前に存在していた時と互いに同じ位置に表示するステップ。

40

## 【発明の詳細な説明】

## 【背景技術】

## 【0 0 0 1】

## 「相互参照」

本出願は、2 0 0 4 年 1 0 月 1 2 日に提出された係属中の米国特許出願第 1 0 / 9 6 4 , 5 2 4 の一部継続出願である。この米国特許出願第 1 0 / 9 6 4 , 5 2 4 は、2 0 0 4 年 4 月 3 日に提出された係属中の米国特許出願第 1 0 / 8 1 7 , 2 4 4 の一部継続出願である。本出願人は、これらの 2 つの出願について、米国特許法第 1 2 0 条に基づいて優先権を主張する。上記米国特許出願第 1 0 / 9 6 4 , 5 2 4 は、放棄された米国仮出願第 6

50

0 / 4 6 0 , 4 7 9 の利益を主張するものであるが、本出願人はこの利益もまた本出願について主張するものである。米国特許出願第 1 0 / 9 6 4 , 5 2 4 、同第 1 0 / 8 1 7 , 2 4 4 及び米国仮出願第 6 0 / 4 6 0 , 4 7 9 の内容全体は参照により本明細書に組み込まれるものとする。

#### 「発明の背景」

分子生物学の研究を支援する新しい実験技術の出現は、データの急増と生物学的測定データのタイプの多様性の急速な増加をもたらしている。かかる生物学的測定データタイプとしては、例えば、DNA マイクロアレイ又は Taqman (タクマン) 実験からの遺伝子発現情報、質量分析又はゲル電気泳動からのタンパク質同定情報、フローサイトメトリー (流動細胞計測法) からの細胞の局在化情報、臨床データ又はノックアウト (knockout) 実験からの表現型情報、会合研究及び DNA マイクロアレイ実験からの遺伝子型情報、比較ゲノムハイブリダイゼーション (Comparative Genomic Hybridization; CGH) データ、アレイベースの CGH データ (aCGH) データ等を挙げることができる。このデータは急速に変化しつつある。新しい技術は、しばしば新しいタイプのデータを生み出す。

#### 【0002】

アレイベースの CGH 技術が発達するにつれて、この技術を使用する研究には、データが生成される多数のアレイが一層含まれるようになる。コンテキスト中のデータの視覚的実験解析を容易にするために研究全体においてそのようなデータを視覚化する必要性が存在する。その他の分野においても、線グラフの形態で個別に表すことができる大きなデータ集合を視覚化するための解決法を適用できる同一又は類似の必要性があるであろう。

#### 【0003】

データを視覚化するための最近の技術は、通常、多くのアレイに対しては十分な拡張性を有さず、または、この技術が多数のアレイからのデータを表示するために拡張可能であっても、個別のアレイに関する十分な詳細は視覚化しない。例えば、標準的ヒートマップ型視覚化物 (または標準的ヒートマップ型視覚化) を、aCGH データを表すために使用することができる (例えば非特許文献 1 参照。この文献は参照によりその全内容を本明細書中に組み入れるものとする)。かかる表示方法は、通常、多数のアレイ / 実験に対し拡張可能であるが、ヒートマップの背景にある細部を調べることは難しい。同じ制約を受けるその他のソフトウェア製品としては、「dchip」 (<http://www.dchip.org> 参照)、「BioConductor」 (<http://www.biocconductor.org> 参照)、及び「GeneSpring」 (<http://www.silicongenetics.com> 参照) を挙げることができる。

#### 【0004】

CGH の視覚化に特に適合する視覚化ソフトウェア及びシステムは、染色体イデオグラム上に重ね合わせたデータを示す傾向がある。例えば、「Visualizing Expression Data on Chromosomal Graphic Schemes」と題する 2004 年 4 月 3 日に提出された現在係属中の米国特許出願第 10/817,244 号、及び、「System and Methods for Statistically Analyzing Apparent CGH Data Anomalies and Plotting Same」と題する 2004 年 10 月 12 日に提出された同時係属中の米国特許出願第 10/964,524 号を参照されたい。尚、これらの文献は参照によりその全内容を本明細書中に組み入れるものとする。このことは、CGH 研究については通常の内容であるが、かかる表示は、例えば、ディスプレイ上の数百の実験を同時に視覚化するためには十分な拡張性を有しない。

#### 【0005】

非常に大きなデータベース内に含まれるまばらなデータを表示するための視覚化ソフトウェア及びシステムが、「System and Methods for Navigating and Visualizing Multi-Dimensional Data」と題する 2004 年 8 月 13 日に提出された同時係属中の米国特許出願第 10/918,897 号に記載されている。尚、この文献は参照によりその全内容を本明細書中に組み入れるものとする。

#### 【0006】

【特許文献1】米国特許出願第10/817,244号明細書

【特許文献2】米国特許出願第10/964,524号明細書

【特許文献3】米国特許出願第10/918,897号明細書

【特許文献4】米国特許出願第10/964,524号明細書

【特許文献5】米国特許出願第10/817,244号明細書

【特許文献6】米国特許仮出願第60/460,479号明細書

【非特許文献1】Pollack他。“Microarray analysis reveals a major direct role of DNA copy number alteration in the transcriptional program of human breast tumors”, PNAS, Oct. 1, 2002, vol. 99, no. 20, 12963-12968

10

【発明の開示】

【発明が解決しようとする課題】

【0007】

データ（例えば、線グラフ群として表示し得るデータ）のより大きな集合をコンパクトなグラフの形態で視覚化するのを容易にする方法、ツール、及びシステムに対する継続的な必要性が存在する。

【課題を解決するための手段】

【0008】

コンテキスト及びフォーカスを提供するためにグラフの集合を視覚化するためのシステム、ツール、方法、及びコンピューター読み取り可能な媒体が提供され、この場合、各グラフを視覚化物の第1の軸に沿って整列させることができる。この第1の軸に対して垂直な視覚化物の第2の軸に沿った方向に全てのグラフを圧縮することにより、圧縮された視覚化物が形成される。この圧縮されたグラフのスケールよりも第2の軸に沿ってより大きなスケールを有する少なくとも1つのグラフの視覚化物を作成するために、圧縮されたグラフの少なくとも1つをズームすることができる。少なくとも1つのズームされたグラフを圧縮されたグラフと共に表示することができ、この場合、この少なくとも1つのズームされたグラフは、圧縮されたグラフに対して同じ順序で、かつ、ズームされる前の互いの順序と同じ順序で表示される。（尚、「ズーム」には一般に、拡大または縮小の意味があるが、本発明の実施例では「拡大」の意味で使用する）

20

30

【0009】

グラフの集合を全体的に見えるように視覚化する一方で、同時に少なくとも1つのグラフに関する詳細を表示するための能力を提供する際に使用するためのユーザーインターフェースが提供され、該インターフェースは、圧縮された形態のグラフの集合を視覚化するように構成された表示部、及び、圧縮された形態のグラフの集合の視覚化物からグラフを選択するための機能を含む。この場合、グラフが選択されると、該インターフェースは、少なくともその選択されたグラフを拡大して、圧縮された視覚化物の圧縮軸に沿ったその選択されたグラフのスケールが、その圧縮軸に沿った、ある圧縮されたグラフのスケールよりも大きくなるように、その選択されたグラフの視覚化物を生成する。また、少なくとも選択されたグラフを圧縮された形態で存在していたのと同じ順番で表示し、この場合、拡大されていないグラフは拡大されたグラフに隣接して表示される。

40

【0010】

本発明のこれら及びその他の利点及び特徴については、以下に、より十分に記載する方法、ツール、システム、及びコンピューター読み取り可能な媒体の詳細を読めば、当業者には明らかとなるであろう。

【実施例】

【0011】

「発明の詳細な説明」

本発明の方法、ツール、システム、及びコンピューター読み取り可能な媒体について記載する前に、本発明は、記載する特定の実施例に限定されるものでなく、当然に変更しう

50

るものであることを理解すべきである。また、本明細書中で使用する用語は、特定の実施形態のみを記載することを目的とするものであり、限定を意図するものではない。本発明の範囲は、添付の特許請求の範囲によってのみ限定されうるからである。

【0012】

ある範囲の値が与えられた場合、文脈上明確に別の値を指示していない限り、各介在値は、下限値の単位の1/10まで、その範囲の上限及び下限の間で具体的に示されることが理解される。ある既定された範囲における任意の指定値又は介在値の間の、より小さい各々の範囲、及びその指定された範囲における任意のその他の指定値若しくは介在値が、本発明の範囲内に包含される。これらのより小さい範囲における上限及び下限は独立してこの範囲に含まれる場合も、含まれない場合もあり、より小さい範囲内にいずれかの限界を含み、又はいずれも含まないか、又は、両方を含む各範囲もまた本発明の範囲内に包含されるが、指定された範囲内において任意の限界を具体的に指定して排除することができ、そのように指定されている場合は排除される。指定された範囲が1つ又は両方の限界を含む場合、含まれるこれらの限界のいずれか若しくは両方を除外する範囲もまた、本発明に包含される。

10

【0013】

別様に定義しない限り、本明細書中で使用する全ての技術的及び科学的用語は、本発明が属する技術分野の当業者に一般的に理解される同一の意味を有する。本明細書に記載されるものと類似又は均等な任意の方法及び材料を本発明の実施又は試験において使用することができるが、好ましい方法及び材料について以下に記載する。本明細書中で言及した全ての刊行物は、引用した刊行物と関連する方法及び/又は材料を開示及び記載するために、参照により本明細書に組み入れるものとする。

20

【0014】

この特許又は特許出願ファイルは、少なくとも1つの彩色されて作成された図を含む。彩色された図を伴うこの特許又は特許出願ファイルの複写は、申請及び必要とされる料金の納付を米国特許庁に行うことにより提供されるであろう。

【0015】

本明細書及び添付の特許請求の範囲において使用する場合、単数形の呼称(「ある」、「その」、「前記」など)、及び、「及び」は、文脈から明確に単数を指示していない限り、複数の対象も含むことに留意すべきである。したがって、例えば、単にグラフと言及した場合には複数のグラフを含み、また、単にアレイと言及した場合には、1以上のアレイ及び当業者に既知のその均等物などへの参照が含まれる。

30

【0016】

本明細書中において参照した刊行物は、本出願の出願日前のそれらの内容の開示のためにのみに提供される。本明細書において、先行の発明に基づいて本発明がそのような刊行物の内容に先行する資格を有しないことを認めるものとして解釈すべきものは全くない。さらに、示された出版日は、実際の出版日と異なるかもしれない、個別に確認する必要がある場合がある。

【0017】

「定義」

40

「異常性領域」とは、遺伝物質の顕著な増幅又は欠失を示すものとして同定されている染色体上の遺伝的データの連続的部分を意味する。

【0018】

「カラーコーディング(カラーコード化)」とは、色値に対して数値又はカテゴリ値(categorical value)をマッピングするソフトウェア技術であり、例えば、遺伝子発現又は増幅の変化の程度を表す色の濃淡/輝度を変化させながら、高レベルの遺伝子発現又は遺伝子増幅を赤みがかった色で表し、低レベルの遺伝子発現又は遺伝子欠失レベルを緑がかった色で表す。中間色を使用することもできる。例えば、欠失されていず、増幅もされていない遺伝物質を、第3の色(例えば、黒色)で表すことができる。カラーコーディングは、発現レベル又はCGHデータへの適用に限定されるものでなく、相対的に多くの量

50

の値と少ない量の値とを区別するように定量化し得る任意のデータを区別するために使用することができる。これらの場合のいずれにおいても、第3の色を、相対的に中立又は中間の値に対して使用することができ、陰影付けを、色指示物のより連続的なスペクトルを提供するために使用することができる。

【0019】

「ヒートマップ」又は「ヒートマップ視覚化」とは、カラーコーディングを数値又は数値範囲を表示するために使用する、視覚的データ表示方法である。線グラフ中の数値又は数値範囲を、それぞれ、その数値又は範囲を表す色にコード化することができる。カラーコーディングは、1つの色から別の色（例えば、数値に対して、緑色から赤色、又は黄色から青色）へと連続的に行うことができる。

10

【0020】

「ダウンレギュレーション」という用語は、遺伝子発現に関して使用され、制御に対する、遺伝子の発現によって形成されるメッセンジャーRNA（mRNA）量の低下を意味する。

【0021】

「遺伝子」という用語は、タンパク質のアミノ酸配列を決定するために必要とされる情報を含むDNAの一部である、遺伝的情報の単位を意味する。

【0022】

「遺伝子発現」とは、タンパク質合成に先立ち、遺伝子が転写されてメッセンジャーRNA分子を形成するレベル（段階）を意味する。

20

【0023】

「遺伝子発現比」とは、遺伝子発現の相対的尺度であり、試験サンプルの発現レベルを参照サンプルの発現レベルと比較したものである。

【0024】

「遺伝子産物」とは、遺伝子から形成され得る生物学的実体（例えば、メッセンジャーRNA又はタンパク質）のことである。

【0025】

「CGHデータ」とは、「比較ゲノムハイブリダイゼーション（Comparative Genomic Hybridization）」測定から得られるデータを意味する。CGHはDNAの獲得又は喪失を測定する技術を含む。ゲノム中の特定DNA配列のレベルを測定するために高スループットマイクロアレイ測定を利用する「アレイCGH（aCGH）」などのより新しい技術が現れてきている一方で、いくつかの技術はCGHを染色体レベルで行う。aCGHデータに特に限定するものではないが、本発明を、アレイベースの遺伝子発現測定と類似の形態で得られるaCGHデータに適用することができる。

30

【0026】

「促進する」という用語は、生物学的因子又は生物学的プロセスの効果の増大を意味する。

【0027】

「タンパク質」とは、ペプチド結合によって連結されたアミノ酸サブユニットの1以上の配列を有する巨大ポリマーのことである。

40

【0028】

「タンパク質存在量」という用語は、サンプル中のタンパク質量の尺度を意味し、参照サンプルに対する相対的存在量の尺度を意味することが多い。

【0029】

「タンパク質/DNA相互作用」とは、タンパク質が（通常、促進因子又は抑制因子領域に結合することにより）遺伝子の発現を制御する生物学的プロセスを意味する。

【0030】

「タンパク質/タンパク質相互作用」とは、2以上のタンパク質が結合して複合体を形成する生物学的プロセスを意味する。

【0031】

50



「配列」とは、タンパク質を形成するアミノ酸の順序付けられた集合、又はDNA断片若しくはRNA分子を形成する核酸塩基の順序付けられた集合を意味する。

【0032】

「オーバーレイ」又は「データオーバーレイ」という用語は、1つのビューにおけるデータを、異なるビューのデータ上に重ね合わせるためのユーザーインターフェース技術（例えば、染色体ビューの上部に遺伝子発現比を重ね合わせること）を意味する。

【0033】

「スプレッドシート」とは、コンピューターソフトウェアアプリケーションによって電子的に模倣した特大の元帳（ledger sheet）のことであり、表形式のデータ構造を表すのにしばしば使用される。

10

【0034】

「アップレギュレーション」という用語は、遺伝子発現を説明するために使用されときは、制御に対する、遺伝子の発現によって形成されるメッセンジャーRNA（mRNA）量の増大を意味する。

【0035】

「ビュー」という用語は、データ集合に対する1つの視覚的見え方をグラフ表示したものを意味する。

【0036】

「視覚化」又は「情報の視覚化」という用語は、人間の知覚を利用する様々な技術を使用する実験的データ解析へのアプローチを意味し、該技術は、大量のデータについてのグラフ表示を含む場合があり、データの操作と調査を対話的に行うことを容易にする。

20

【0037】

本明細書中で使用する時、「オリゴマー」という用語は、複数のモノマーを含む化学的実体を示すために使用される。本明細書中で使用する時、「オリゴマー」及び「ポリマー」という用語は同義的に使用される。「オリゴマー」及び「ポリマー」の例としては、ポリデオキシリボヌクレオチド（DNA）、ポリリボヌクレオチド（RNA）、プリン若しくはピリミジン塩基のC-グリコシドであるその他の核酸、ポリペプチド（タンパク質）若しくは多糖類（デンプン若しくは多糖質）、並びに、同様の化学的構造体の反復単位を含むその他の化学的実体を挙げることができる。

【0038】

本明細書中で使用する時、「核酸」という用語は、ヌクレオチド（例えば、デオキシリボヌクレオチド、リボヌクレオチド）、又は、ワトソン-クリック塩基対合相互作用に参与し得る2つの天然核酸の場合と同様の配列特異的な様式で天然の核酸とハイブリダイズする（ハイブリダイゼーションを行う）ことができる、合成的に生成された化合物（例えば、米国特許第5,948,902号及び該特許文献中の引用文献に記載されているPNA）から構成されるポリマーを意味する。

30

【0039】

本明細書中で使用する時、「リボ核酸」及び「RNA」という用語は、リボヌクレオチドから構成されるポリマーを意味する。

【0040】

本明細書中で使用する時、「デオキシリボ核酸」及び「DNA」という用語は、デオキシリボヌクレオチドから構成されるポリマーを意味する。

40

【0041】

本明細書中で使用する時、「オリゴヌクレオチド」という用語は、長さが約10～100のヌクレオチド及び長さが最大200のヌクレオチドの1本鎖ヌクレオチド多量体を意味する。

【0042】

本明細書中で使用する時、「官能基化（または機能化）」という用語は、基材表面上に複数の官能基を提供するための固体基材の修飾に関連する。「官能化表面」とは、複数の官能基が存在するように修飾された基材表面を意味する。

50

## 【 0 0 4 3 】

「反応部位」、「反応性官能基」又は「反応基」という用語は、有機合成プロセスの開始点として使用され得るモノマー、ポリマー又は基材表面上の部分の意味する。この用語は、基材表面上にも存在し得る「不活性」親水性基（例えば、ポリエチレングリコール、ポリアミド等に関連する親水性部位）と対比される。

## 【 0 0 4 4 】

本明細書中で使用する時、「サンプル」という用語は、1以上の対象成分を含む、典型的には流体形態の（これは必須ではないが）物質又は物質の混合物に関する。

## 【 0 0 4 5 】

「ヌクレオシド」及び「ヌクレオチド」という用語は、既知のプリン及びピリミジン塩基ばかりでなく、修飾されているその他の複素環式塩基も含有する成分を含むことを意図するものである。かかる修飾物には、メチル化プリン若しくはピリミジン、アシル化プリン若しくはピリミジン、アルキル化リボース、又はその他の複素環式化合物が含まれる。さらに、「ヌクレオシド」及び「ヌクレオチド」という用語には、通常のリボース及びデオキシリボース糖ばかりでなく、その他の糖を含有する成分も含まれる。修飾されたヌクレオシド又はヌクレオチドには、糖部分の修飾物も含まれ、この場合、例えば、1以上の水酸基が、ハロゲン原子若しくは脂肪族基と置換されるか、又は、エーテル、アミン等として官能化される。

## 【 0 0 4 6 】

「固体支持体表面に結合したオリゴヌクレオチド」という表現は、あるフィーチャー（形態）すなわちスポット中の固体基材の表面上に固定化されたオリゴヌクレオチド又はその擬似体（例えば、PNA）を意味し、この場合、基材は様々な構成（例えば、シート、ビーズ、又はその他の構造）を有することができる。特定の実施形態において、本発明中で使用するオリゴヌクレオチドのフィーチャーの集合は、同一平面状支持体の表面上に、例えば、アレイの形態で存在する。

## 【 0 0 4 7 】

「アレイ」という用語は、「マイクロアレイ」という用語を包含し、核酸等に結合するために提供される順序付けされたアレイ（順序配列）を意味する。以下により詳細に記載するように、アレイは通常、複数の個別の又は異なるフィーチャーから構成される。本明細書中においては、「フィーチャー」という用語は、「複数のフィーチャー」、「フィーチャーエレメント」、「スポット」、「アドレス指定可能領域」、「異なる成分の領域」、「表面又は基材固定化エレメント」、及び「アレイエレメント」という用語と同義的に使用され、この場合、各フィーチャーは、固体支持体の表面に結合したオリゴヌクレオチドから構成され、基材固定化核酸とも称される。

## 【 0 0 4 8 】

「アレイ」には、核酸、特に、オリゴヌクレオチド又はその合成擬似体（すなわち、上に記に定義したオリゴヌクレオチド）等を担持する任意の1次元、2次元又は実質的に2次元（及び3次元）の構成のアドレス指定可能領域（すなわち、例えば、スポット形態のフィーチャー）が含まれる。アレイが核酸アレイの場合、核酸は、核酸鎖に沿った任意の点または任意の複数の点でアレイと、吸着、物理的吸着、化学的吸着、又は共有結合させることができる。

## 【 0 0 4 9 】

任意の所与の基材は、基材表面上に配置された1、2、4、又はそれ以上のアレイを担持することが可能である。用途に応じて、いずれかの又は全てのアレイは、互いに同じか又は異なってもよく、それぞれが複数のスポット（すなわち、フィーチャー）を含むことができる。典型的なアレイは、 $20\text{ cm}^2$ よりも小さい、又はさらに $10\text{ cm}^2$ よりも小さい、例えば、（約 $1\text{ cm}^2$ よりも小さい、または、約 $1\text{ mm}^2$ よりも小さい（例えば、 $100\text{ }\mu\text{m}^2$ 、又はさらに小さい）領域を含む）約 $5\text{ cm}^2$ よりも小さい領域に、1以上（例えば、2、10、100、1000、10000、又は100000よりも多く）のフィーチャーを含んでもよい。例えば、フィーチャーは、 $10\text{ }\mu\text{m} \sim 1.0\text{ cm}$ の範

10

20

30

40

50

囲の幅（すなわち、円いスポットに対する直径）を有することができる。その他の実施形態において、各フィーチャーは、 $1.0\ \mu\text{m} \sim 1.0\ \text{mm}$ 、一般には $5.0\ \mu\text{m} \sim 500\ \mu\text{m}$ 、より一般的には $10\ \mu\text{m} \sim 200\ \mu\text{m}$ の範囲の幅を有することができる。非円形フィーチャーは、上記幅（直径）範囲の円形フィーチャーの面積範囲と同等の面積範囲を有することができる。フィーチャーの少なくともいくつか又は全てにおいて、成分が異なっている（例えば、各フィーチャー成分の任意の繰り返しが除外される場合、残りのフィーチャーは、合計フィーチャー数の少なくとも5%、10%、20%、50%、95%、99%、又は100%を占め得る）。いずれの核酸（又はフィーチャーを構成するタイプのその他のバイオポリマー（生体高分子）若しくは化学成分）も担持していないフィーチャー間の領域が通常（本質的ではないが）存在し得る。かかるフィーチャー間領域は、通常、試薬の液滴沈着を伴う工程によってアレイが形成される場合には存在し、一方で、例えば、フォトリソグラフィーによるアレイ製造工程を利用した場合には存在しないであろう。しかしながら、フィーチャー間領域が存在する場合には、それらは様々な大きさと構成を取り得ることが理解されるであろう。

10

#### 【0050】

各アレイは、 $200\ \text{cm}^2$ より小さい領域、又は $50\ \text{cm}^2$ 、 $5\ \text{cm}^2$ 、 $1\ \text{cm}^2$ 、 $0.5\ \text{cm}^2$ 、若しくは $0.1\ \text{cm}^2$ よりさらに小さい領域をカバーすることができる。特定の実施形態において、1以上のアレイを担持する基材は通常、（他の形も可能であるが） $4\ \text{mm} \sim 150\ \text{mm}$ 、一般的には $4\ \text{mm} \sim 80\ \text{mm}$ 、より一般的には $20\ \text{mm}$ より短い長さを有し、 $4\ \text{mm} \sim 150\ \text{mm}$ 、一般的には $80\ \text{mm}$ より小さい、より一般的には $20\ \text{mm}$ より小さい幅を有し、及び $0.01\ \text{mm} \sim 5.0\ \text{mm}$ 、一般的には $0.1\ \text{mm} \sim 2\ \text{mm}$ 、より一般的には $0.2\ \text{mm} \sim 1.5\ \text{mm}$ 、例えば、約 $0.8\ \text{mm} \sim$ 約 $1.2\ \text{mm}$ の厚さを有する矩形の固体として形成される。蛍光を検出することにより読み取られるアレイとともに、基材は、励起光の照射時に弱い蛍光を発する材料であってもよい。さらにこの場合、集光レーザービームが照射領域をあまりに遅く走査する場合の入射照射レーザー光の吸収と続く加熱を低減させるために、基材を比較的透明性のものとすることができる。例えば、基材は、かかる照射光のスペクトルを集約した全体にわたって、あるいは $532\ \text{nm}$ 又は $633\ \text{nm}$ において測定したときに、表面に対する照射光入射の少なくとも20%、又は50%（あるいは、少なくとも70%、90%、又は95%）を透過し得る。

20

#### 【0051】

アレイは、その場で（*in situ*）作成する場合における核酸前駆体単位（例えば、モノマー）、あるいは予め得た核酸のいずれかによるパルス・ジェットからの液滴沈着を用いて、製造することができる。かかる方法は、例えば、米国特許第6,242,266号、同第6,232,072号、同第6,180,351号、同第6,171,797号、同第6,323,043号、Carénらによる1999年4月30日に出願された米国特許出願第09/302,898号など今までに引用されている文献、ならびにそれらの中で引用されている文献中に詳細に記載されている。すでに述べたように、これらの文献を参照により本明細書中に組み入れるものとする。その他の液滴沈着方法を、本明細書中で前述したような製造のために使用することができる。さらに、液滴沈着法の代わりに、フォトリソグラフィーによるアレイ製造方法を利用することもできる。アレイが上記特許文献中に記載されているようなフォトリソグラフィーによる方法によって作られるときには、フィーチャー間領域は特に存在する必要はない。

30

40

#### 【0052】

ある特定の実施形態では、その場で（*in situ*）作成されるアレイが使用される。その場で（*in situ*）作成されるオリゴヌクレオチドアレイ（例えば、核酸アレイ）は、フィーチャーとフィーチャー間領域との間で大きく異なる基材表面特性を有することにより特徴付けることができる。具体的には、かかるアレイは、高表面エネルギー・親水性のフィーチャーと、疎水性・低表面エネルギー・疎水性のフィーチャー間領域を有することができる。基材の所与の領域（例えば、フィーチャー又はフィーチャー間領域）が高表面エネルギー又は低表面エネルギーを有するかについては、当該技術分野で既知であり、また同

50

時係属出願第10/449,838号(この文献の開示内容は参照により本明細書中に組み入れるものとする)中にさらに記載されているように、水との「接触アングル」を該領域について調べることにより容易に判定することができる。本発明の特定の実施形態において、そのような特別興味のあるアレイの形態を作成するその場で(in situ)作成されたアレイのその他の特徴としては以下のものがある(但し、これらに限定するものではない)。フィーチャー密度、各フィーチャー内のオリゴヌクレオチド密度、フィーチャー均一性、低フィーチャー内バックグラウンド、低フィーチャー間バックグラウンド(例えば、疎水性のフィーチャー間領域による)、個々のフィーチャーを構成するオリゴヌクレオチドエレメントの忠実度、アレイ/フィーチャーの再現性などである。上記のその場(in situ)作成のアレイの利点は、高度に複雑なサンプルへの対応に必要とされるストリン

10

#### 【0053】

アレイは、複数の異なる部分の領域、すなわち、複数のフィーチャーを有する時(例えば、それぞれが異なるオリゴヌクレオチド配列から構成されている時)、「アドレス指定可能」であり、これにより、アレイ上の特定の所定位置(すなわち、「アドレス」)の領域(すなわち、アレイの「フィーチャー」又は「スポット」)が、特定の液相核酸配列を検出する。アレイのフィーチャーは、必ずしも必要ではないが、通常、中間に存在する空間によって分け隔てられている。

#### 【0054】

20

任意の所与の基材は、基材表面上に配置された1、2、4、又はそれ以上のアレイを担持することが可能であり、アレイの用途に応じて、いずれかの又はすべてのアレイは互いに同じか又は異なっているとしてもよく、それぞれが複数のスポット又はフィーチャーを含むことができる。通常、1以上のアレイは表面の一部のみをカバーしており、該表面の反対側の面、先端、及び後端に隣接する表面領域はアレイによってカバーされていない。各アレイを、試験サンプル、参照サンプル、それらの組み合わせ、又は生体高分子(例えば、ポリヌクレオチド)の既知の混合物であろうとなかろうと、任意の種類のサンプルに対して試験するために設計することができる。基材は、上記したように、任意の形状であってもよい。

#### 【0055】

30

上記したように、アレイは、オリゴマー(例えば、ポリヌクレオチドの形態、特にオリゴヌクレオチドの形態で)の複数のスポット又はフィーチャーを含む。上記したように、全てのフィーチャーは異なっているとしてもよく、又は、いくつか若しくは全てが同じであってもよい。フィーチャー間領域は、様々な大きさ及び構造であってもよい。各フィーチャーは、所定のオリゴマー、例えば、所定のポリヌクレオチド(これはポリヌクレオチド混合物である可能性を含む)を担持する。アレイ表面と第1のヌクレオチドとの間に任意の既知の種類のリンカー分子が存在し得る事が理解されるであろう。

#### 【0056】

基材は、例えば、基材上に印刷されたバーコード等の形態、接着剤若しくは任意の便利な手段によって付着されたラベルの形態の識別コードを有していてもよい。識別コードは、基材上に位置するアレイ(複数の場合もある)に関する情報を含むことができ、この場合、該情報は、限定するものではないがアレイの識別表示(すなわち、アレイ(複数の場合もある)に関するレイアウト情報等)を含んでもよい。

40

#### 【0057】

本出願におけるアレイの場合、「標的」とは、様々な領域において基材と結合している「プローブ」により検出される移動相(典型的には流体)中の成分を意味する。

#### 【0058】

「スキャン領域」とは、既に定義した対象とするアレイスポット又はフィーチャーが見出される又は検出される連続的(好ましくは、矩形の)領域を意味する。蛍光標識を使用する場合、スキャン領域は、照明された全領域の部分であり、生じた蛍光は、その部分が

50

ら検出され記録される。その他の検出プロトコルを使用する場合、スキャン領域は、照会された全領域の部分であり、生成された信号は、その部分から検出され記録される。本発明の目的のため及び蛍光検出実施形態に関して、スキャン領域は、たとえ対象となるフィーチャーを有しない介在領域が存在していても、最初の対象フィーチャーと最後の対象フィーチャーとの間に、レンズを通過することにスキャンされるスライド（滑動体）の全領域を含む。

#### 【0059】

「アレイレイアウト」とは、フィーチャーの1以上の特性を意味する（例えば、基材上のフィーチャーの位置、1以上のフィーチャーの大きさ、及び所与の位置における成分の表示）。核酸に関して「ハイブリダイゼーションを行う」と「結合する」とは同義的に使用される。

10

#### 【0060】

「離れた位置」とは、アレイが存在し、且つハイブリダイゼーションが起こっている位置以外の位置を意味する。例えば、離れた位置は、同じ都市内の別の位置（例えば、オフィス、研究所等）、異なる都市内の別の位置、異なる州内の別の位置、異なる国内の別の位置等であってもよい。このように、1つの要素がもう一方の要素から「離れている」と示される場合、これは、2つの要素が少なくとも異なる部屋又は建造物中にあり、少なくとも1マイル、10マイル、または少なくとも100マイル離れている場合があり得ることを意味する。情報を「伝達する」とは、適切な伝達チャネル（例えば、構内ネットワークまたは公衆網）上の電気信号として、その情報を表すデータを伝達することを意味する。ある要素を「転送する」とは、その要素を物理的に移動させるか、または（可能であれば）その他の方法により、その要素を1つの位置から次の位置へ移動させる任意の手段を意味し、少なくともデータの場合には、データを運搬するかまたはデータを伝達する媒体を物理的に移動させることを含む。アレイ「パッケージ」は、アレイとアレイが配置された基材のみであってもよいが、他の機能（例えば、チャンバを有するハウジング）を含むこともできる。「チャンバ」とは密閉された容積部を意味する（但し、チャンバは、1以上のポートを通じてアクセス可能であってもよい）。また、「上部」、「上側」、及び「下側」等の用語は相対的な意味においてのみ使用されることは、本願明細書を通じて理解されるであろう。

20

#### 【0061】

本明細書中で使用する時、「ストリンジェントなアッセイ条件（厳格な試験条件）」という用語は、所望のレベルの特異性を分析時に提供するのに十分な相補性を有する核酸（例えば、表面結合した核酸及び液相核酸）の結合対を形成するために適合している一方、所望のレベルの特異性を提供するのに相補性が不十分な結合メンバー間で結合対を形成するための適合性が低い条件を意味する。ストリンジェントなアッセイ条件とは、ハイブリダイゼーション条件と洗浄条件の両方を加えたもの又はそれらの組み合わせ（合計）である。

30

#### 【0062】

核酸ハイブリダイゼーション（例えば、アレイにおけるサザン（Southern）又はノーザン（Northern）ハイブリダイゼーション）の状況において「ストリンジェントなハイブリダイゼーション」及び「ストリンジェントなハイブリダイゼーション洗浄条件」は、配列依存的であり、異なる実験パラメーターの下では異なる。本発明の範囲内の核酸を同定するために使用することができるストリンジェントなハイブリダイゼーション条件には、例えば、50%ホルムアミド、5×SSC、及び1%SDSからなるバッファー（緩衝液）中42におけるハイブリダイゼーション、又は、5×SSC及び1%SDSからなるバッファー（緩衝液）中65におけるハイブリダイゼーションを含むことができ、共に、0.2×SSC及び0.1%SDSの65での洗浄を含む。例示的なストリンジェントなハイブリダイゼーション条件には、40%ホルムアミド、1MNaCl、及び1%SDSからなるバッファー中、37におけるハイブリダイゼーション、及び、1×SSCにおける45での洗浄もまた含まれる場合がある。あるいは、0.5M N

40

50

a H P O<sub>4</sub>、7 %ドデシル硫酸ナトリウム ( S D S )、1 m M E D T A 中、6 5 において、フィルター結合 D N A (filter-bound DNA) に対しハイブリダイゼーションし、0 . 1 x S S C / 0 . 1 % S D S において 6 8 で洗浄する方法を採用し得る。さらなるストリンジェントなハイブリダイゼーション条件には、6 0 以上、3 x S S C ( 4 5 0 m M 塩化ナトリウム / 4 5 m M クエン酸ナトリウム ) におけるハイブリダイゼーション、又は 3 0 %ホルムアミド、1 M N a C l、0 . 5 %サルコシンナトリウム ( sodium sarcosine )、5 0 m M M E S、p H 6 . 5 含有溶液中での 4 2 におけるインキュベーションが含まれる。同様のストリンジェント ( 厳格さ ) な条件を提供するために、代替的であるが同等のハイブリダイゼーション及び洗浄条件を利用できることを当業者は容易に認識するであろう。

10

#### 【 0 0 6 3 】

特定の実施形態において、洗浄条件のストリンジェンシー ( 厳格さ ) は、ある核酸が、表面に結合された核酸に対し特異的にハイブリダイズされるかどうかを判定する条件を表す。核酸を同定するために使用される洗浄条件は、例えば、「塩濃度が p H 7 で約 0 . 0 2 M ( モル )、及び温度が少なくとも約 5 0 、又は約 5 5 ~ 約 6 0 」、又は、「塩濃度が約 0 . 1 5 M N a C l、7 2 にて約 1 5 分間」、あるいは、「塩濃度が約 0 . 2 x S S C、温度が少なくとも約 5 0 、又は約 5 5 ~ 約 6 0 にて約 1 5 ~ 約 2 0 分間」、あるいは又、「ハイブリダイゼーション複合体を塩濃度が 0 . 1 % S D S を含有する 2 x S S C の溶液で室温、1 5 分間の洗浄を 2 回、次いで 0 . 1 % S D S を含有する 0 . 1 x S S C により 6 8 、1 5 分間の洗浄を 2 回」、又は、それらと同等の条件を含み得る。ストリンジェント ( 厳格 ) な洗浄条件は又、例えば 0 . 2 x S S C / 0 . 1 % S D S、4 2 とすることができる。

20

#### 【 0 0 6 4 】

ストリンジェントなアッセイ条件の特定の例は、( 例えば、2 0 0 0 年 9 月 5 日に提出された米国特許出願第 0 9 / 6 5 5 , 4 8 2 ( この開示内容を参照により本明細書に取り込むものとする ) に記載されているように ) 1 ユニタリオン合計濃度が 1 . 5 M である塩ベースのハイブリダイゼーション緩衝液中での 6 5 における回転ハイブリダイゼーションと、それに続く、室温における、0 . 5 x S S C 及び 0 . 1 x S S C の洗浄である。

#### 【 0 0 6 5 】

ストリンジェントなアッセイ条件は、少なくとも上記の代表的条件と同様の厳格性を有するハイブリダイゼーション条件であり、その場合、上記の特異的な条件と比較して、所与の条件のセット中で、望まれる特異性を提供するための十分な相補性を欠く付加的な結合複合体が実質的に生成されない場合には、その所与の条件セットは少なくとも厳格であるとみなされる。この場合、「実質的にそれ以上でない」とは、約 5 倍未満であることを意味し、典型的には 3 倍未満であることを意味する。その他のストリンジェント ( 厳格 ) なハイブリダイゼーション条件は当該技術分野において既知であり、適宜採用されている。

30

#### 【 0 0 6 6 】

感度とは、サンプル中の所与の分析物 ( 被検体 ) ( 例えば、対象としている核酸種 ) を検出するための所与のアッセイ ( 分析法 ) の能力を示すために使用される用語である。例えば、サンプル中の低濃度の被検分子を検出できる場合、アッセイは高感度を有する。逆に言えば、サンプル中の高濃度の被検分子 ( すなわち、対象としている特定の溶液相の核酸 ) しか検出できない場合、所与のアッセイの感度は低い。所与のアッセイの感度は、採用する試薬の特異性 ( 例えば、ラベル ( または同位体トレーサー ) のタイプ、結合分子のタイプ等 )、採用されるアッセイ条件、採用される検出プロトコール及びその他を含むパラメーターの数に依存する。本発明におけるようなアレイハイブリダイゼーションアッセイの状況において、所与のアッセイの感度は、表面に固定化した核酸の性質、ハイブリダイゼーション及び洗浄条件の性質、ラベリングシステムの性質、検出システムの性質等のうちの 1 つ以上に依存し得る。

40

#### 【 0 0 6 7 】

50

液体クロマトグラフィーは、成分分離、典型的には溶媒中に溶解したイオンや分子を分離するのに有用な分析的クロマトグラフィー技術である。この技術では、成分（例えば、被検体）をまず溶媒中に溶解し、次いで数センチメートルから数メートルの範囲とすることができ、クロマトグラフィーカラムを通過して流れるようにする。カラムは、使用する溶媒に適合し、吸着により被検体と結合する固相のクロマトグラフィー材料で充填される。次いで、付加的な別の溶媒を、濃度を高めながら（例えば滑らかな勾配をなす増加、又は段階的な増加により）流れの中に混合する。被検体中の各混合物（または各化合物）は、付加的な溶媒の特定の濃度において固相から解放され、次いでカラムから流出して、被検体中に含まれる混合物の連続的な分離をもたらす。流出液中の混合物の存在を同定するための各種の検出器が、各種の異なる検出原理に基づき過去30年以上にわたり開発されてきている。典型的には、クロマトグラフィー検出器からの信号強度を溶出時間（クロマトグラム）の関数としてプロットすることができ、ピークを成分の同定に使用する。クロマトグラフィーカラムにおける特徴的な保持（滞留）時間のようなその他の技術も成分同定に適用することができる。この適用において質量分析計は、複数混合物の存在を同時に検出し、かつ検出された混合物同士を区別できる非常に高感度の多重化検出器として機能する。

#### 【0068】

液体クロマトグラフィー/質量分析（LC/MS）は、複雑な生物学的サンプル中のタンパク質、ペプチド及び/又は代謝産物の全体的な同定及び定量化のために広く使用されている技術である。この技術では、液体クロマトグラフィーは、所与の時間に質量分析計に供される成分数を低減させる目的で、質量検出の前に成分をクロマトグラフィーにより分離するために、質量分析計とインラインで使用される。

#### 【0069】

本発明は、線グラフ、散布図（scatter plot）、又は線グラフの表現に変換可能なその他の任意のデータの非常に大きな集合をコンパクトなグラフ形式で視覚化し、その一方で、単一の又は複数の線グラフを、グラフの非常に大きな集合に関して詳細に検査する能力を依然として維持するためのシステム、ツール、方法、及びコンピューター読み取り可能な媒体を提供する。このような能力は、コンテキストとフォーカス（すなわちコンテキストとは、グラフ間全体の関係が認識できるような表示、例えば、全グラフ表示に対応し、フォーカスとは、注目したいグラフの局所の細部表示、例えばズーム表示に対応する）と呼ばれる場合がある。例えば、散布図を圧縮する場合、散布図中のポイントは、各々が2つのポイントをつなぐ勾配線により接続され得る。これらは直線勾配又はスプライン・フィット又はその他の補間法を表す勾配であり得る。プロットの複数のポイントがX軸上の同じ位置に存在する場合、その結果これらの値をX軸位置ごとに平均するか又は1つの代表データ値に結合することができる。なぜなら、ほとんどの圧縮された表示画面（ビュー）において、各X座標位置におけるポイントは、1つのY座標値によってのみ表すことができるからである。また、完全に圧縮した時には、画面は基本的にヒートマップビューであり、その場合Y座標値は色符号化（カラーコーディング）により表され得る。本発明に従って表示し、操作することができるデータの例は、「アレイ上で行われる実験から得られるaCGHデータの集合（例えば、数百から数千までのそのような実験をコンピューターディスプレイ等の1画面上に表示することができる）」又は「紫外線スペクトルの集合、可視光スペクトルの集合、赤外線スペクトルの集合、核磁気共鳴スペクトルの集合、マス（質量）スペクトルグラフの集合、又はその他の任意の関連するスペクトルの集合、若しくは線グラフの形態で表せるその他の任意のデータなどの関連するスペクトルの集合」を含む（但し、これらに限定するものではない）。このような集合の表示内におけるデータの詳細の調査も可能である。

#### 【0070】

ここで、図1に、本システムにより実現される表示部100の例を完全に圧縮された形態で示す。完全に圧縮された形態において、表示部100は、aCGHデータを表示するために現在なされているいくつかの視覚化に類似している。図1において表示されている

データは、Pollackらにより行なわれ“Microarray analysis reveals a major direct role of DNA copy number alteration in the transcriptional program of human breast tumors”で報告された乳癌の研究に由来している。尚、この文献は参照により本明細書に組み込まれるものとする。示されているビュー（画面）中に、82のマイクロアレイ実験の集合から得られたデータが圧縮された形式で表示されている。ビュー100は、数百又は数千にも及ぶそれらの実験を、標準的なコンピュータディスプレイ上に単一の視覚化物として表示できることを理解できるようなスケールで示されている。さらに、付加的な（すなわち、視覚化物上に一度に表示可能な最大数より大きい）実験が存在する場合には、付加的な実験を視覚化するために表示をスクロールするためのユーザーインターフェースが装備される。

10

#### 【0071】

82のマイクロアレイ（いくつかは遺伝子発現用であり、いくつかはaCGH用である）についてのデータをシステムに取り込み、1アレイに対し20の測定値からなる連続するグループ毎に単純移動平均を計算した。連続的な測定値は染色体に沿った連続した位置から取得された測定値であり、同一測定（値）からのデータは、染色体の位置に関連して同一の順番で表示される。（癌のサンプルと癌でないサンプルとの）割合値の色勾配マッピングを表示するが、この場合、例えば1未満の割合を第1の色でカラーコーディングし、1より大きい割合値を第2の色でカラーコーディングすることができる。さらに、1に等しい割合値を第3の色でカラーコーディングすることができる。例えば、連続的な非直線的なシグモイド状（S字型）のカラーグレーディング方式を採用してもよく、その場合、1よりはるかに小さな割合値には明るい緑色を割り当て、1に近づくにつれて連続的に暗くなる緑色を割り当て、1の値には黒色を割り当て、1より大きな値には赤色を割り当て、1をわずかに超える値には暗い赤を、そして値が大きくなるにつれて連続的に明度が増加するように割り当てる。プロットされたCGH値について、ユーザーは、高い割合値を増幅の可能性があると解釈し、低い割合値を欠失の可能性があると解釈し得る。あるいは、Z-スコア又はその他の統計的異常値をCGHデータに関してプロットし、本明細書中に記載のようにプロットすることもできる。このようなZ-スコアプロットの例は米国特許出願第10/964,524号明細書に記載されているが、それは、本明細書中に記載のようにプロットされる。

20

30

#### 【0072】

最もコンパクトにデータを表現できるようにするために、マイクロアレイを表す各線グラフの各線セグメントがカラーコーディングされた（例えば、典型的な遺伝子発現ヒートマップにおいてヒートマップカラーリングが行われるのと同じか又は類似の方法で）。典型的なヒートマップカラーコーディングの例としては、Kincaidによる、“Vista Clara: an interactive visualization for exploratory analysis of DNA microarrays”、Proceedings of the 2004 ACM symposium on Applied computing、ACM Press, 2004, pp 167-174及び、“Methods and System for Simultaneous Visualization and Manipulation of Multiple data Types”と題する2003年3月31日に提出された同時係属中の米国特許出願第10/403,762号を参照されたい（両者の全文を参照により本明細書中に組み入れるものとする）。カラーコーディングは、欠失及び増幅（及び/又は発現レベルにおける顕著な増加及び減少）などの対象となる領域を、さほど詳細にでなくとも色表現によって依然として知らせることができるようにして、圧縮形態での線グラフの図式的表現を可能にする。このような圧縮されたビューは、複数の実験における傾向を視覚的に識別するのに非常に強力であり得る。例えば、図1において、多くの実験は赤色でカラーコーディングされた大きなブロックで示されているように、図1中の102rにより識別される染色体1の領域の増幅を示すことが見て取れる。欠失を示す複数のアレイの例を染色

40

50



体 2 3 において観察することができる。参照符号 1 0 4 g を参照されたい。

【 0 0 7 3 】

図 1 に示すような圧縮ビューは複数の実験における傾向を識別するのに有用であり得るが、上述したように、このような視覚化は、圧縮された個々のグラフに関する詳細を示すものではない。各グラフはグラフあたりわずか約 1 ~ 2 ピクセルのスケール（このスケールは、図 1 における縦軸のスケールすなわち高さであるが、横軸のスケールすなわち幅とすることもできる）に圧縮されるため、圧縮された形態では目盛に沿った大きさの変動は視覚化されないからである。従って、図 1 に示される形態における視覚化物のような圧縮された視覚化物を見た場合には、任意の特定のアレイ又は染色体にわたる挙動に関する非常に細かい細部を識別することはできない。本システムは、圧縮された視覚化物の選択された部分をズーム又は拡張する能力を提供することにより、このような画面表示の汎用性を向上させる。例えば、図 1 の表示部において対象領域を「クリックする」又はその他の方法で選択することにより、その領域の圧縮が解除され、拡張され、または、ズームされて、図 2 に示すようなデータの視覚的な表現が現れる。

10

【 0 0 7 4 】

クリックされた場所は拡大されて、表示部中で最大にズームされた、すなわち、最大に拡大された表現となる。すなわち、視覚化物 1 0 2 上の選択された線グラフ 1 1 0、又はその選択部分に最も近い線グラフ 1 1 0 が、図 2 で示される拡大された視覚化物 2 0 2 において最大の大きさで表示される線グラフである。システムは、選択された線グラフ 1 1 0 をズームすなわち拡大するのみならず、近傍の線グラフ 1 1 2 も拡大されたビュー中に表示する。ただし、この場合、縦軸方向の拡大倍率すなわちスケールは、線グラフ 1 1 0 が表示されているものよりは小さい。図示の例では、線グラフ 1 1 0 の縦軸スケールに対して、線グラフ 1 1 2、1 1 4、1 1 6 及び 1 1 8 が次第に減少する縦軸スケールにおいて表示され、ズームされた線グラフ 1 1 0 から完全に圧縮された線グラフ 1 2 0 まで小さくなっていく視覚的印象を伝える。ズームされた線グラフを、最大にズームされた線グラフ 1 1 0 から両方向に離れるように縦軸スケールを次第に減少させながら表示することは有用ではあるが、本発明にとって必須ではないことに留意されたい。その他の表示方式を代わりに採用することもできるからである。例えば、縦軸スケールは線グラフ 1 1 0 が表示されるものよりは小さいが、線グラフ 1 1 2、1 1 4、1 1 6 及び 1 1 8 を全て同じ縦軸スケールで表示してもよい。あるいは、より多くの、若しくはより少ない数の線グラフを線グラフ 1 1 0 に隣接してズームされたスケールで表示することもできる。又は、線グラフを、選択された線グラフ 1 1 0 の片側方向（上又は下のいずれか）にのみ隣接して拡大してもよい。もちろん、ここに挙げた方法の代わりとして利用できるその他の多くの表示方式が存在するが、それらは、当業者には容易に明らかなものであろう。

20

30

【 0 0 7 5 】

近傍の線グラフ 1 1 2、1 1 4、1 1 6、1 1 8 をズームされたスケールで表示することにより（線グラフ 1 1 0 と同じスケールであっても、線グラフ 1 1 0 から離れる方向に徐々に縦軸スケールが小さくなるものであっても、それ以外の表示方式によるものであっても）、観察者は、選択された実験（例えば、線グラフ 1 1 0）の近傍の詳細を観察する機会を与えられる。従ってこの特徴は、近傍の線グラフ / 実験結果の詳細を比較するために有用である。この特徴は、わずかな操作のエラーを修正するのにさらに有用である。すなわち、完全に圧縮された形態 1 0 2 にあるときは、ユーザーは圧縮した形態の行間で選択エラーを起こす可能性がある。例えば、図 1 ~ 2 に示す例において、ユーザーは実際には、図 2 における線グラフ 1 1 0 の上方に表示されている線グラフ 1 1 4 の詳細を観察することに最大の興味を有しているという場合があり得る。この場合、図 2 の視覚化物が提示された時には、ユーザーは単純にクリックするか又はその他の方法で、より詳細な観察のための選択が本来意図されていた線グラフ 1 1 4 を選択可能である。システムは、この場合、図 2 に示されるものと類似の視覚化物が得られるように線グラフを再表示するが、図 2 における線グラフ 1 1 0 の上方に示された線グラフ 1 1 4 が今や最大の縦軸スケールを有するように表示され、線グラフ 1 1 4 の下方に縦軸スケールが徐々に小さくなるよう

40

50

にして線グラフ 1 1 2、1 1 0、1 1 2 及び 1 1 4 が表示され、線グラフ 1 1 6、1 1 8、及び線グラフ 1 1 8 のすぐ上に他の 2 つの線グラフ（図 2 ではラベル付けされていないが）、縦軸スケールが除々に小さくなるようにして表示される。

#### 【0076】

図 2 を再度参照するが、観察者を支援するための視覚的な手がかりとして、拡大された線グラフ 1 1 0、1 1 2、1 1 4、1 1 6 及び 1 1 8 のヒートマップカラーリング（ヒートマップの配色）が、圧縮されたグラフ中で同じ色領域を有する拡大後の線グラフの高値と、圧縮されたグラフ中で同じ色領域を有する拡大後の線グラフの低値との関連においてそれぞれ維持されることに留意されたい。例えば、線グラフ 1 1 0 のセグメント 1 1 0 r は、図示のズームされたビューにおいても、発現レベルにおける増幅又は増加を示すために赤色にカラーコーディングされるが、このことは、セグメント 1 1 0 r がベースライン 1 1 0 b 上に位置するという事実から容易に理解され得る。同様に、セグメント 1 1 0 g は、発現の欠失又は減少を示すために緑色でカラーコーディングされる。ベースライン 1 1 0 b は、また、選択された行であった行 / 線グラフを示しているが、ズーム方式は、その選択された行に基づいて視覚化を行っている。

10

#### 【0077】

同じ線グラフ 1 1 0 を 2 度目にクリック又はその他の方法で選択することにより、さらなるレベルの詳細にアクセスすることができる。すなわち、選択された線グラフ 1 1 0 が図 2 に示されているズームビューで表示されている間に、ユーザーは次いで、ユーザーが横軸方向に拡張したいと希望する線グラフ 1 1 0 の領域上をクリック又はその他の方法で選択することができる。それに応じて、視覚化物（表示物）の横軸スケールが、ユーザーにより選択されたグラフ 1 1 0 の領域において拡大される。拡大された領域は、典型的には、予め定義された視覚化物の列であり、それは、染色体や、例えばその他の予め定義された領域又は測定を表すことができる。図示の例において、番号 1 7（染色体 1 7 を指す）により識別される列において、ユーザーが線グラフ 1 1 0 を選択した場合を考える。図 2 において、染色体 1 7 に対応するデータは、その他のいくつかの染色体（例えば、1、2、3 等）に関する表示よりもかなり短い長さに沿って表示されている。染色体 1 7（列 1 7）領域における線グラフ 1 1 0 上をクリックすることにより、図 3 に示すように、染色体 1 7 におけるデータが拡大されるため、列 1 7 の長さが長くされ、したがって、全ての線グラフについて染色体 1 7 におけるデータの横軸スケールが拡大されるので、視覚化物 3 0 2 がその線グラフの詳細を示し、染色体 1 7 に対応する領域をより容易に識別できるようになる。

20

30

#### 【0078】

ズームの焦点を変更する、すなわち最大の縦軸スケールにより拡大されるべき別の線グラフを選択するために、異なる線グラフをクリックすることに対する別の方法として、本システムは視覚化物のスクロールを可能にする。スクロールバー 4 0 0（図 4 参照）又はその他の機能又はツールをユーザによる対話的な使用のために提供して、ユーザが、選択されるべき線グラフまでスクロールすることにより、拡大されるべき線グラフの選択を変更できるようにすることができる。ボタン 4 0 2 を上方向又は下方向のいずれかに縦方向にドラッグすることにより、ベースライン 1 1 0 b がカーソルとして作用して、視覚化物における線グラフ上を縦方向である上方向又は下方向のいずれかに移動し、その場合、そのベースライン 1 1 0 b の動きがユーザーによるボタン 4 0 2 の動きに対応する。スクロールの間に各線グラフをベースライン 1 1 0 b が通過する時、それは、図 2 におけるグラフ 1 1 0 により示されたものと同じような拡大されたグラフへと拡大されまたはズームされる。さらに、隣接したグラフも、スクロールの間に、例えば図 2 に示す程度まで拡大することができる。この意味では、スクロールは、そのグラフに対する倍率バー（例えば、小さい文字を読むための矩形の拡大鏡）をスライドさせるのに幾分似た視覚的效果を与え、その場合、「倍率バー」の中央部に位置する、すなわちベースライン 1 1 0 b が重ね合わさっていることにより規定されるグラフが最大の大きさに拡大される。ボタン 4 0 2 を選択された位置に「ドロップする」ことにより、対応する線グラフ（ベースライン 1 1 0

40

50

b が重ね合わさるグラフ) が選択され、図 2 における線グラフ 1 1 0 に関して図示し説明したように拡大された視覚化物の形態で表示される。記載したように、スクロールの機能は、ベースライン 1 1 0 b が線グラフ上を通過するとき、ベースライン 1 1 0 b が通過する各線グラフのズームング (拡大) を行うことができるのでさらに利点となり得る。このように、ユーザーには、さらに詳細な観察のためのズームされた視覚化物を提供するためにどの線グラフを選択すべきかを決定する際に有用であり得る、データの詳細を迅速に「スキャンする」能力が提供される。さらに、このようなスクロール機能は、グラフ間の傾向、類似性、及び / 又は相違をさらに識別するために使用され得る。

#### 【 0 0 7 9 】

対象とするデータの横軸スケールの大きさを変更するための代替機能によって、ユーザーは視覚化物中の 1 以上の列分離線 (column separator) 3 1 0 をクリック又は選択 (例えば、ドラッグ及びドロップ) することができる。例えば、図 3 における列分離線 3 1 0 をドラッグ及びドロップすることにより、ユーザーは、上述のように線グラフ 1 1 0 を染色体 1 7 上の領域において 2 度目にクリックすることで得られる結果と同様のビューを表示できる。しかし、この機能は、ユーザーが 1 以上の分離線をドラッグ及びドロップでき、また、1 以上の列を横方向に拡大する距離を変化させることができるため、より汎用性に富む。例えば、ユーザーは、図 3 の染色体 1 7 における情報がさらに詳細に表示されることを望む場合があるが、その場合、ユーザーは分離線 3 1 0 をよりさらに右方向へドラッグして、より拡大された新しい位置へとドロップすることができる。あるいは、分離線 3 1 0 を図 3 に示されている位置から左側の位置にドロップすることにより、染色体 1 7 におけるデータを、図 3 に示す程度よりは少ないが、幾分横方向に拡大することができるだろう。またさらに、拡大される領域を定義する片側又は両側の分離線 3 1 0 を動かすことにより、2 つ以上の領域を横方向に拡大することができる。

#### 【 0 0 8 0 】

移動平均 (moving average) の計算によりプロットされるデータに関して、本システムは、デフォルトのウィンドウサイズとは異なるウィンドウサイズに基づくように移動平均を再計算する能力を提供する。例えば、図 1 に表示されているデータは、20 の測定からなる連続的なグループ毎に計算された移動平均に基づいて表示されている。ユーザーはウィンドウサイズを減少させるか又は増加させる決定ができ、この場合、新たなウィンドウサイズを用いて計算された移動平均に基づいてデータを視覚化することができる。例えば、ユーザーは、15 の測定からなる連続的なグループ毎、または 30 の測定からなる連続的なグループ毎に計算された移動平均の計算に基づいてデータを再表示するようシステムに指示することができる。

#### 【 0 0 8 1 】

他の機能により、ユーザーは、個々のデータポイントがプロットされる散布図を視覚化することができる。さらに、個々のポイントを、移動平均値によって生成された線グラフとともに、個々のポイントの散布図を組み合わせた視覚化物を提供するために、線グラフと同じ軸上にプロットすることができる。より一般的には、折りたたまれた、すなわち圧縮された線グラフのビューは、本質的に 2 次元のチャートを 1 次元のヒートマップのビュー (すなわち、圧縮されたビュー中に示されない第 2 の次元における値を表すための、カラーコーディングされた位置を伴う直線) に低減する。2 次元プロットを表示するために圧縮されたプロットが「開かれている」又は「圧縮が解除されている」時、ユーザーがそのようなプロットと関連付けることを望む場合がある任意の他のタイプのデータを、2 次元の線グラフと並置して、重ね合わせて表示することもできる。

#### 【 0 0 8 2 】

線グラフ (及び / 又は散布図) をプロットすることのほかに、又はそれに加えて、本システムは異なる基準で値を計算し、これらの値を単独で、又は元の線グラフ及び / 又は散布図の視覚化物上に重ね合わせてプロットすることができる。表示される付加的な情報には、Z - スコア、p - 値、又はその他の生データから計算された値を含めることができる (但し、これらに限定するものではない)。

## 【 0 0 8 3 】

上述したように、本発明は、a C G Hデータ又はその他のマイクロアレイ由来のデータに限定されないが、順序付けられたグラフとして表示でき、及びそのような順序で見ることができるデータの任意の集合を表示するために使用することができる。しかし、集合における全てのグラフを、有意なやり方でひとつの軸（典型的には横軸）に沿って整列させることができなければならない。図1に関して説明した例では、染色体の位置は、視覚化された全てのアレイ（線グラフ）間で共通である。

## 【 0 0 8 4 】

さらに、プロットされたデータが散布図又はその他の不連続なプロットである場合であっても、上述したように、直線、スプライン、又はその他の補間技術を使用して色勾配を補間することにより、ポイント間で値を連続的に表すことが可能である。この特徴は、有用であって、これにより、圧縮された視覚化物102は圧縮時に適切に（補間されて）満たされる。例えば、本システムは、補間された値に基づいて色が割り当てられることとなる圧縮された値を決定するために、散布図の隣接するデータ値の間を補間することができる。したがって、空白が個々のポイント（点）の間に存在するのではなく、その空白が圧縮されたビューにおいて表示される代表ポイントである場合は、システムは、個々のデータポイントを表す線グラフ中に存在したであろう値を表す値を計算する。しかし、この要件は、有意なカラーマッピング（カラーコーディング）をヒートマップの欠落した領域に対して確立することができる状況などにおいては厳格に必要とされるものではない。例えば、データが欠落している領域を考慮しない方がより正確であるときには、欠落したデータに由来する1以上の不連続箇所を有する不連続な線グラフを、欠落しているデータ領域に対して補間を行って赤色又は緑色のいくらかの濃淡を与えるのではなく、その領域にはデータが存在しないことを示す灰色又はなにか他の別の色によるカラーコーディングで表示することができる。その他の例は、セントロメア（動原体）の位置上のデータが得られないC G Hデータについて生じるが、これについては、データの欠落を示すようにカラーコーディングすることができる。いずれの場合においても、上述の理由により、圧縮されたマップにおいて十分に視覚的な特徴を生じさせることが重要である。このような圧縮されたマップにより、ユーザーは、ビュー上内の圧縮された多くの構成物（実験、グラフ）間の傾向、類似性及び相違を識別することが可能になるだろう。

## 【 0 0 8 5 】

ここで留意すべきは、当業者には容易に明らかとなるように、位置（例えば染色体上の位置）はx軸中に示され、実験（例えばマイクロアレイ）結果はy軸に沿って示されているが、それらの軸を反転できることは明らかであり、その場合も同じ機能を提供できることである。このような例の1つを図5Aに示すが、この図では、情報がプロットされる軸を反転することに加えて、視覚化物がさらに個別化すなわち分離されて、染色体毎に表示されている。したがって、全体の視覚化物502におけるそれぞれの視覚化物504は、データがプロットされる軸を反転した後の、視覚化物102中に表示されたデータの列/染色体に対応する。もちろん、染色体のプロットを、個別の染色体視覚物を表示するための代替法として、データが視覚化物102中にプロットされるのと同じやり方でデータがプロットされるように水平方向に表示することができる。圧縮されたデータ表示504は、各位置に隣接した染色体の図式的表現506に隣接して表示される（その各位置において、そのデータにより表される遺伝子材料が該染色体上に配置されている）。染色体毎に染色体及び関連するデータの視覚化を行うことに関するさらに詳細な情報は、米国特許出願第10/817,244及び10/964,524（これらは参照により本明細書中に取り込むものとする）に記載されている。

## 【 0 0 8 6 】

ユーザーは、例えば視覚化物504内の特定の領域をクリックするなどにより、個々の視覚化物506、504のいずれも選択できる。この対話的な動作により、システムは、拡大ビュー506m、504emにおいて、選択された視覚化物を表示するのを促される。図5A中に示した例では、染色体17及びそれに関連するデータの視覚化は、選択され

て、拡大ビュー 17m において、拡大された染色体地図 506m 及び選択されたグラフを拡大ビュー中に表示するデータ拡大ビュー 504em として表示される。この場合、隣接するグラフは、拡大ビュー中に示されるが、圧縮されたグラフに向かういずれかの方向に一覧できるように水平方向の大きさ（倍率）が低減される。上述の任意のやり方で隣接するグラフを拡大するだけでなく、選択された視覚化物 504 を選択されたグラフに沿ってズームすることができる。もちろん、データがプロットされていた軸が反転されているために、この例では、選択されたグラフは縦方向に走っており、グラフの値の振幅は水平方向に拡大されている。

#### 【0087】

拡大された視覚化物は、ユーザが、他のグラフを最大の倍率で見るために、そのグラフをクリックすることにより、又は、前述のやり方でグラフをスクロールすることによってそのグラフを選択できるように対話的（インタラクティブ）でもある。ビュー 504em はユーザの選択に従って変更され、ビュー 504e も同様にして変更される。選択されたグラフ 510 をさらに選択する（すなわち、すでに最大倍率で表示されているグラフの一部を選択する）と、そのグラフのその一部のスケールを縦方向に拡大して、そのグラフの一部に沿ったさらなる詳細を示すことができる。この場合、関連する染色体の対応するグラフ 506m は、染色体地図上のさらなる詳細を示すために同じ量だけ拡大される。図 5B はこの機能の 1 例を示しているが、この例では、ラベル q12 の近傍で選択がなされた。図 5B における点線のボックス 512 により示されるように、選択がなされた場所の近傍における表示 504em 及び 506m の一部は縦方向に拡大された。

10

20

#### 【0088】

図 6A には、紫外線（UV）検出器を備えた 2 次元高性能液体クロマトグラフ（2-D HPLC, Agilent 1100LC, Agilent Technologies, Inc., Palo Alto, California）システムから生成されたプロットの視覚化物 102 の 1 例が、完全に圧縮された形態で示されている。図 6B は、図 2 に関して説明したのと同様に、図 6A の視覚化物 102 における領域を拡大のために選択したことにより生成された視覚化物 202 である。したがって、視覚化物 102 中でクリックされるか、又はその他の方法で選択された位置は、視覚化物 202 の中で最大限にズームされた、すなわち、拡大された表現 610 まで引き伸ばされる。すなわち、視覚化物 102 上で選択された、又はその選択箇所に最も近い線グラフ 610 は、図 6B 中に示されている拡大された視覚化物 202 における最大部分を表示する線グラフである。システムによって、選択された線グラフ 610 が拡大される、すなわち、ズームされるだけでなく、近傍の線グラフ 612 もまた拡大して表示される。但し、線グラフ 612 は、線グラフ 610 が表示されているものよりは縦方向の倍率すなわちスケールは小さい。図 2 の例と類似して、ズームされた線グラフ 610 から完全に圧縮された線グラフ 620 まで遷移する視覚的印象を伝えるために、線グラフ 612、614、616 及び 618 は、線グラフ 610 の縦軸スケールに対して、縦方向スケールを順次小さくしながら表示される。最大にズームした線グラフ 610 から両方向に離れるにしたがって縦軸スケールを除々に減少させながらズームした線グラフを表示することは、有用ではあるが、本発明にとって必須ではないことに留意されたい。他の表示方式をその代わりに採用することができるからである。例えば、線グラフ 612、614、616 及び 618 を、線グラフ 610 が表示されているものよりは縦軸スケールは小さいが、全て同じ縦軸スケールで表示することができる。あるいはより多くの又はより少ない線グラフを、線グラフ 610 に隣接してズームしたスケールで表示することもできる。あるいは線グラフを、選択された線グラフ 610 の片側のみ（上側又は下側のいずれか）に隣接して拡大することができる。もちろん、それらの記載したものの代替として採用することが可能な、（当業者には容易に明らかになるであろう）多くの他の表示方式が存在する。尚、図 6A 及び 6B のグラフがカラーコーディングされていないこと、したがって、本発明は白黒又はグレイスケールのグラフに対しても有用であることに留意されたい。

30

40

#### 【0089】

50

さらに、本システムのユーザーインターフェースは、グラフ（例えば、アレイからのデータ）を、線グラフの類似性又はその他のソート（並べ替え）基準に従って並べ替えるための、ユーザーにより選択可能なソート機能を提供することができる。例えば、グラフの配列順を、類似性ソートに加えて各種のクラスター化技術又はその他の基準により再順序付けすることができる。グラフが生成される間の時間に基づいてソートを行なうことができ、及び／又は、癌の異なるサブタイプ（例えば、良性、非悪性、悪性）、生存率、異なる薬物療法等のメタデータを分類する１以上の他のカテゴリーに基づいてソートを行なうことができる。類似性ソートに対しては、米国特許出願第１０／４０３，７６２号（これは参照により組み込まれるものとする）に記載されているベクター類似性ソート手順を、類似する結果を示すグラフが表示１０２、２０２、３０２、５０２をなすように共に並べ替えられるようにグラフを順序付けるために使用することができる。類似性ソート、クラスター化又はその他のタイプのソートを、グラフの全長にわたって、又は選択されたサブセットについて行うことができる。例えば、a C G Hデータに対して、選択されたサブセットは特定の染色体又は染色体群であり得る。さらにユーザーは、染色体の１つの選択された部分、又は、染色体の複数の選択された部分からなるグループなどのような、類似性ソートが行われる、グラフのカスタムメードのサブセットを選択することができる。

10

20

30

40

50

#### 【００９０】

図７は、本発明の１実施形態に従う典型的なコンピューターシステムを示す。コンピューターシステム７００は、１次記憶装置７０６（典型的にはランダムアクセスメモリ、すなわち、ＲＡＭ）、１次記憶装置７０４（典型的にはリードオンリーメモリ、すなわち、ＲＯＭ）を含む記憶装置に結合された任意の数のプロセッサ７０２（中央処理装置、すなわち、ＣＰＵとも呼ばれる）を含む。当業者に良く知られているように、１次記憶装置７０４はデータ及び命令を１方向的にＣＰＵに転送するよう動作し、１次記憶装置７０６は、典型的にはデータ及び命令を双方向的に転送するために使用される。これらの２つの１次記憶装置は、上記のような任意の好適なコンピューター読み取り可能な媒体を含むことができる。大容量記憶装置７０８もまた双方向的にＣＰＵ７０２と結合され、追加のデータ記憶容量を提供し、上記の任意のコンピューター読み取り可能な媒体を含むことができる。大容量記憶装置７０８は、プログラム、データ、及びそれに類するものを記憶するために使用されることができ、典型的には、１次記憶装置よりは速度の遅いハードディスク等の２次記憶媒体である。大容量記憶装置７０８内部に保持された情報を、適切な場合においては、仮想メモリのように１次記憶装置７０６の一部として標準的な様式で組み込むことができることが理解されよう。ＣＤ－ＲＯＭ又はＤＶＤ－ＲＯＭ７１４等の特定の大容量記憶装置もまた、データを１方向にＣＰＵへと送ることができる。

#### 【００９１】

ＣＰＵ７０２はまた、ビデオモニター、トラックボール、マウス、キーボード、マイクロフォン、タッチセンシティブ・ディスプレイ、トランスデューサー・カードリーダー、磁気又は紙テープリーダー、タブレット、スタイラス、音声又は手書き認識装置、又はもちろん、その他のコンピューター等のその他のよく知られた入力装置等の１以上の入力／出力装置を含むインターフェース７１０と結合される。最後に、ＣＰＵ７０２を、オプションとして、コンピューターに、又は７１２において大まかに示されているネットワーク接続を使用して電気通信ネットワークに結合することができる。このようなネットワーク接続により、上述の方法ステップを実行する過程でＣＰＵがネットワークから情報を受け取り、又はネットワークに対して情報を出力することが考慮されている。上述の装置及び材料は、コンピューターハードウェア及びソフトウェアの技術分野における当業者にはなじみ深いものであろう。

#### 【００９２】

上述のハードウェア要素は、本発明の処理を実施するために複数のソフトウェアモジュールの命令を実行することができる。例えば、移動平均を計算しプロットするための命令を、大容量記憶装置７０８又は７１４に保存し、１次メモリ７０６と協働してＣＰＵ７０８で実行することができる。

## 【 0 0 9 3 】

本発明を特定の実施形態を参照して説明したが、本発明の真の思想及び範囲を逸脱することなく各種の変更が可能であり、また均等物で置換できることが当業者には理解されよう。さらに、本発明の目的、思想及び範囲に対して、特定の条件、材料、物質組成、処理、1つ又はいくつかの処理工程を適合させるための多くの変更が可能である。このような変更の全ては本明細書に添付の特許請求の範囲内に含まれることが意図されている。

## 【図面の簡単な説明】

## 【 0 0 9 4 】

【図 1】ユーザーインターフェース上で視覚化された、圧縮された線グラフの表示を示す図である。

10

【図 2】図 1 に示された視覚化物の表示を示す図であって、選択された部分に関してより詳細を視覚化するために、第 1 の軸に沿った第 1 の方向に関して、圧縮された表示の選択された一部が、圧縮を解除、又は、拡大すなわちズームされている。

【図 3】図 2 の視覚化物に類似の視覚化物を示しているが、選択された部分が視覚化物の他方の軸に沿ってさらに拡大されている。

【図 4】グラフの集合における詳細なグラフを視覚的にスキャンする時に使用することができるスクロールツールを概略的に示す図である。

【図 5 A】グラフの集合の別の視覚的表示を示しており、グラフの部分が、対応する染色体のグラフに隣接して個々に示されている。

【図 5 B】図 5 A で示されたグラフの別のビューであるが、拡大されたビュー内に示されたグラフの縦方向（垂直方向）が、選択した領域に関してさらに縦方向に拡大されている。

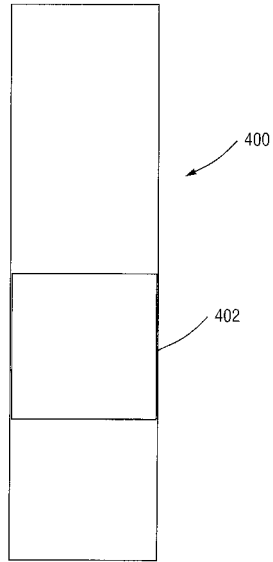
20

【図 6 A】2 次元高性能液体クロマトグラフから得られたプロットの視覚化物の 1 例を示す。

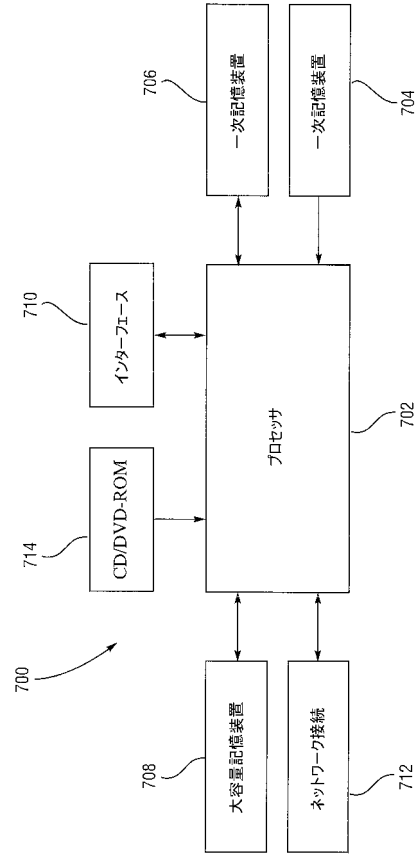
【図 6 B】図 6 A の視覚化物におけるある領域を拡大のために選択することにより生成された視覚化物である。

【図 7】本発明の 1 実施形態に従う典型的なコンピューターシステムを示す。

【 図 4 】

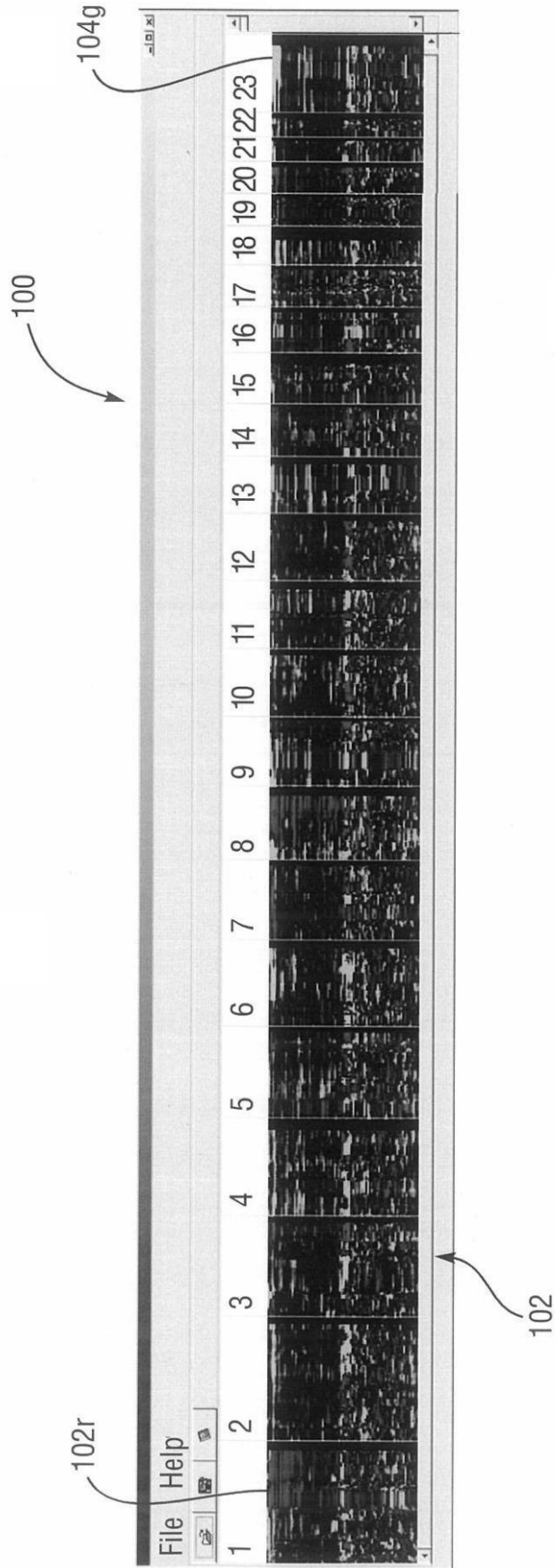


【 図 7 】

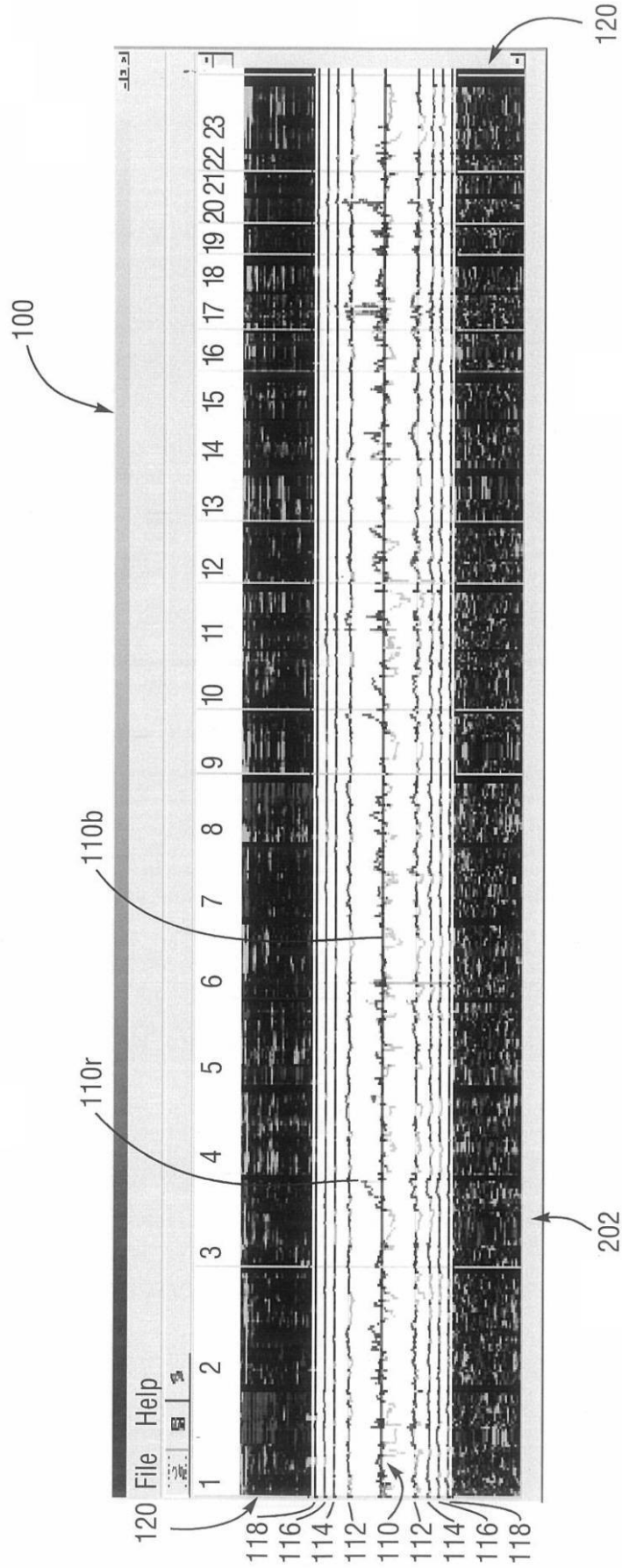




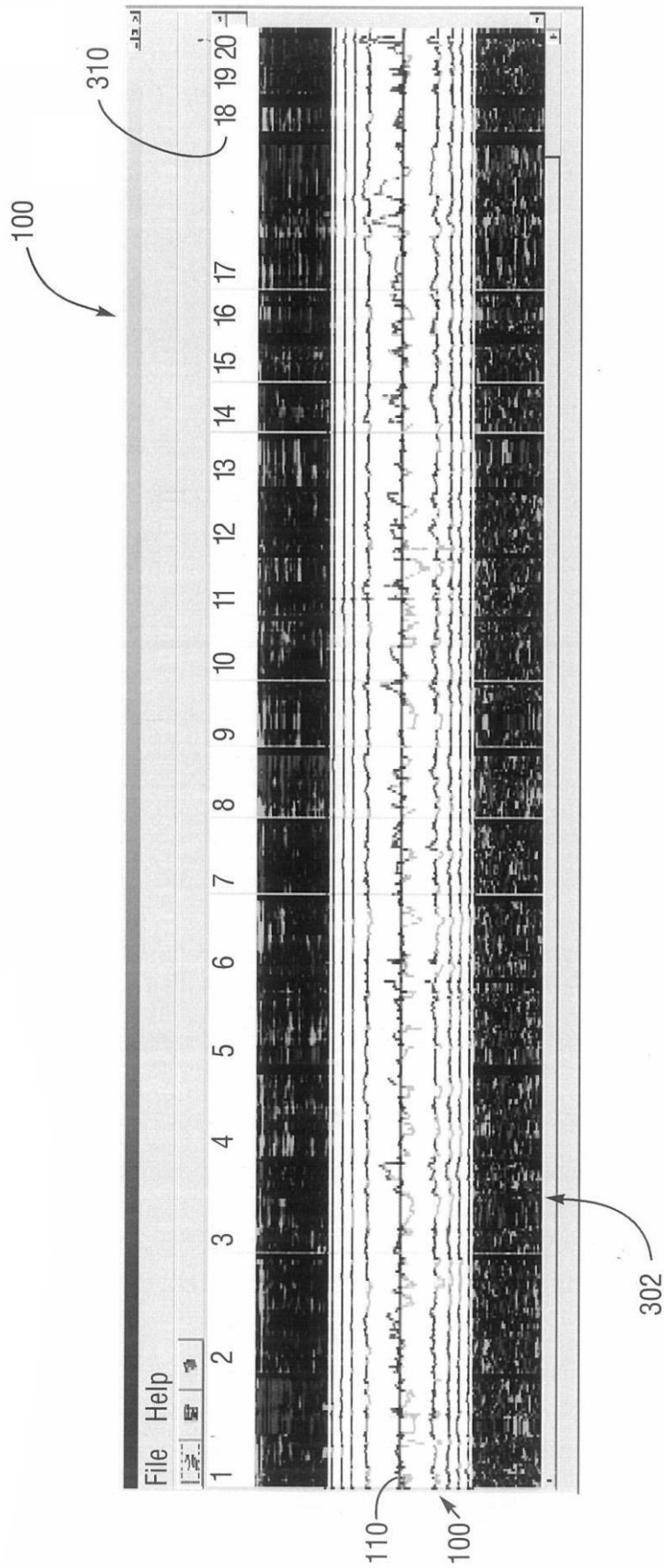
【図 1】



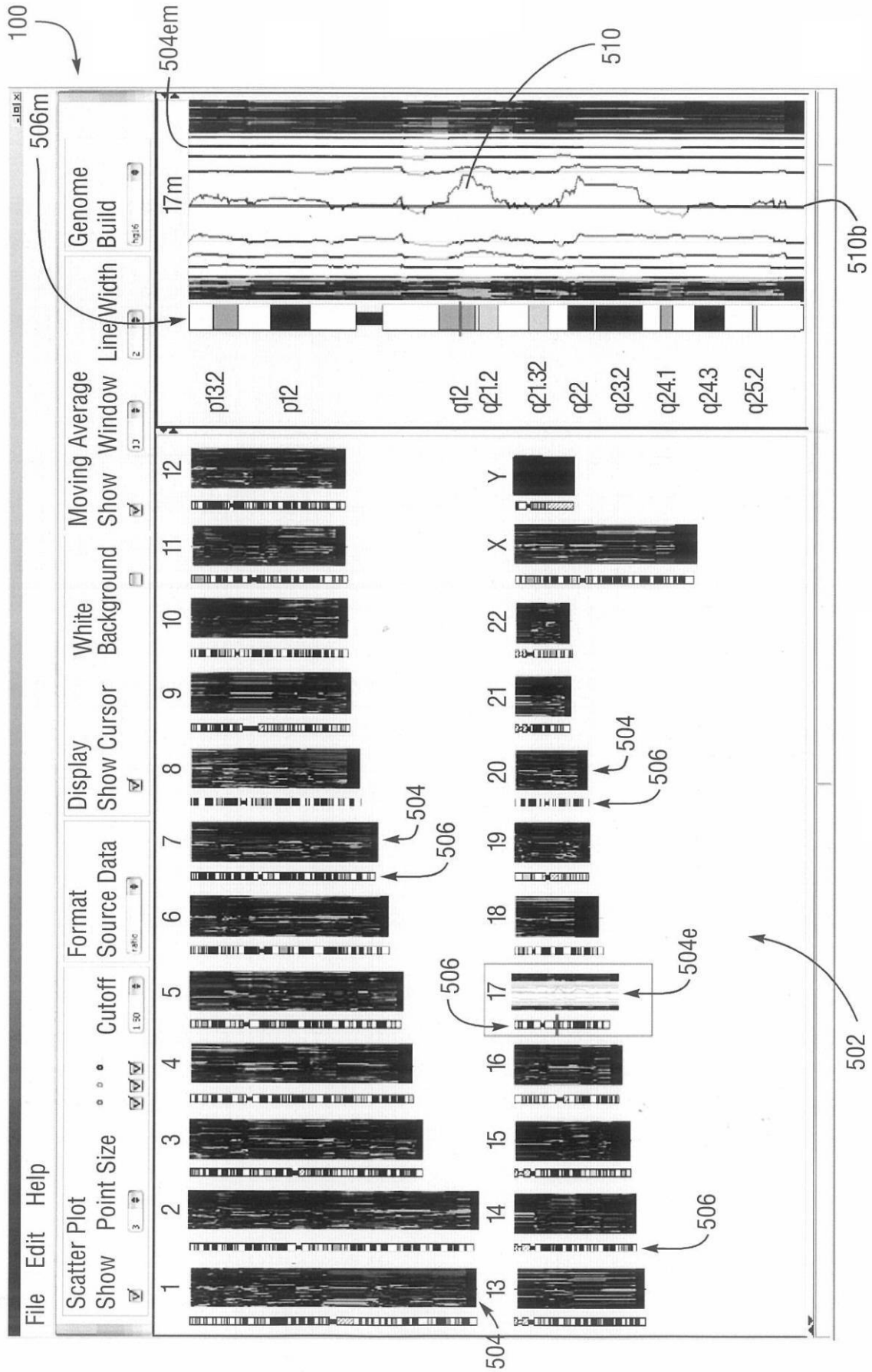
【図 2】



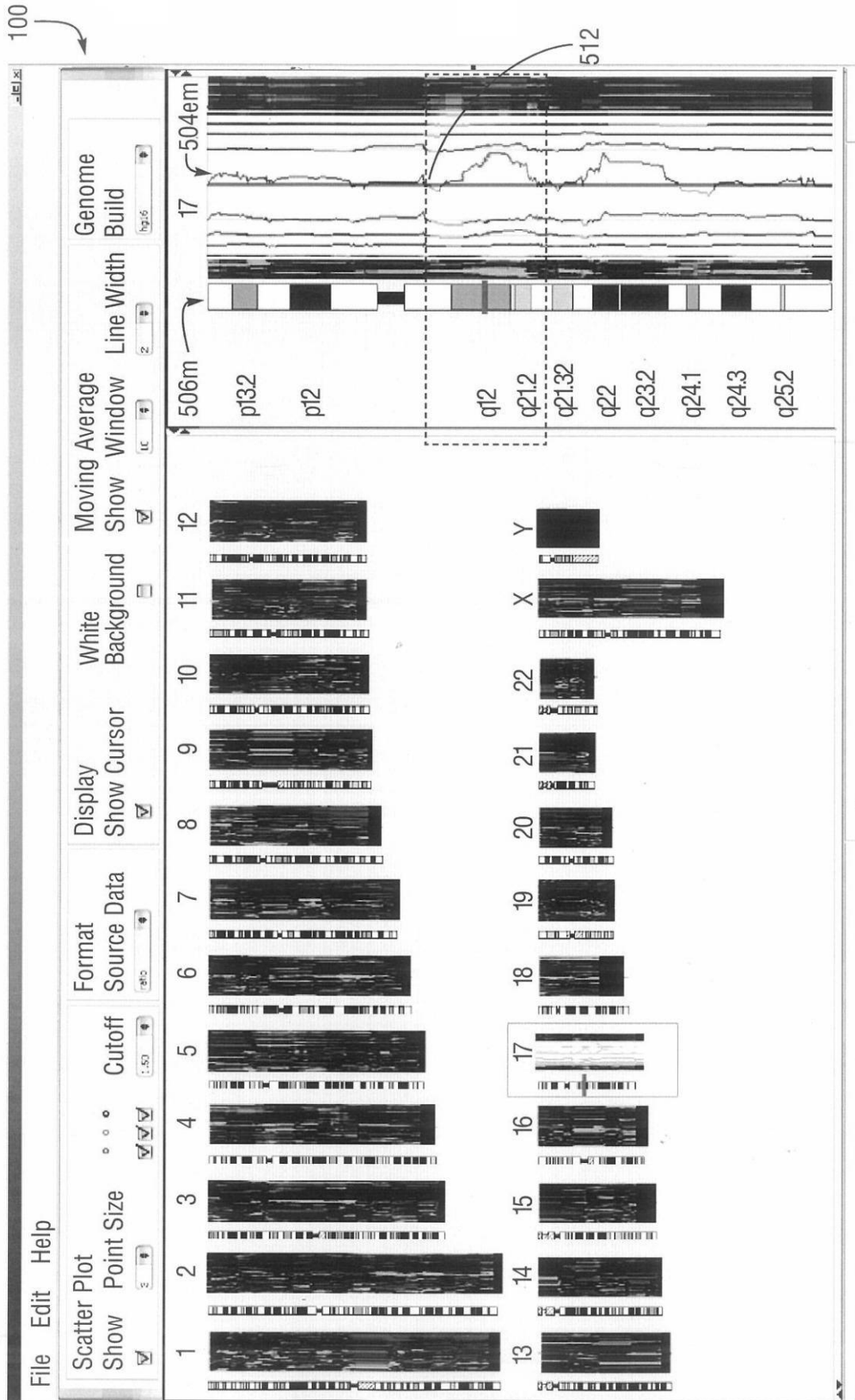
【図 3】



【図 5 A】



【図 5 B】

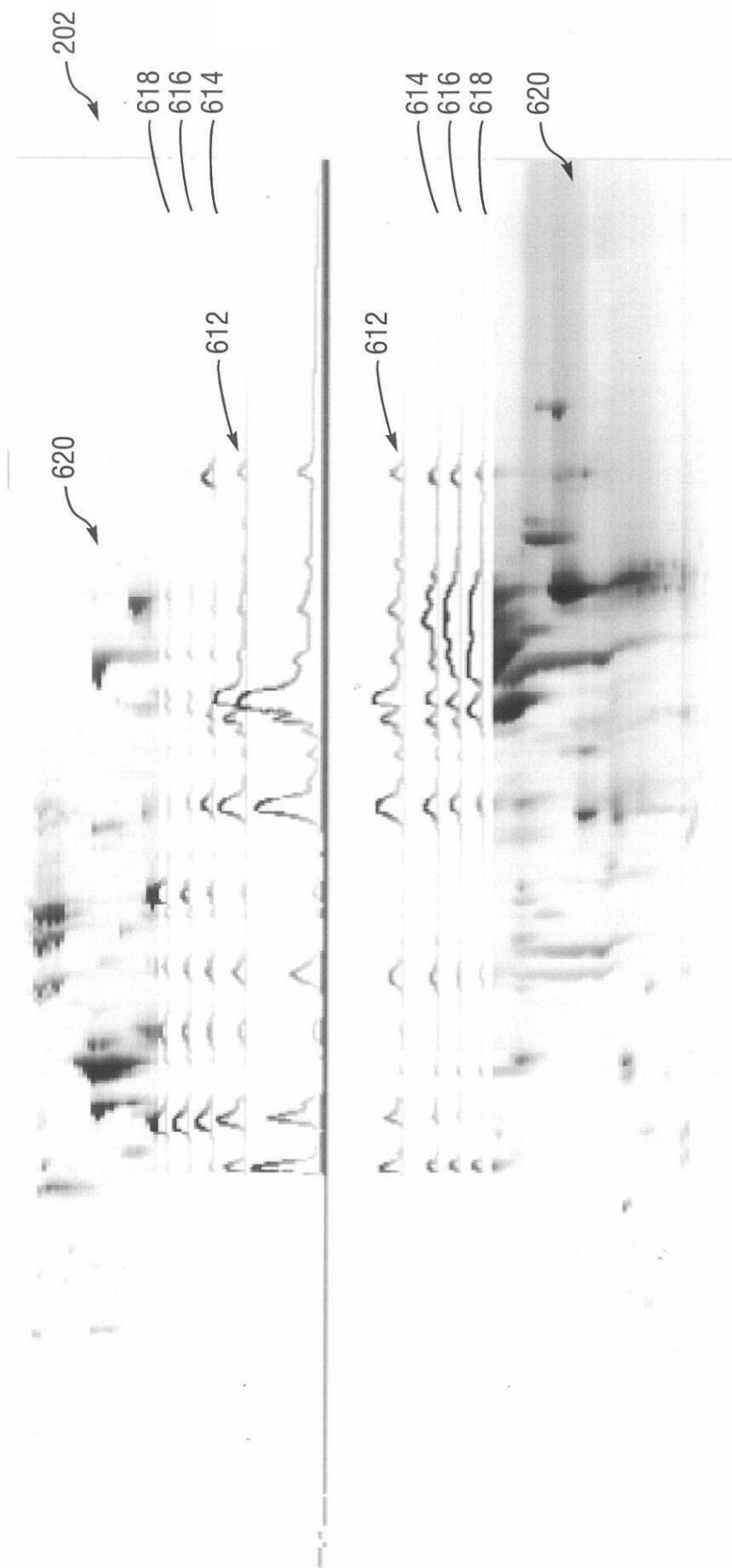


【図 6 A】

102



【図 6 B】



---

フロントページの続き

(72)発明者 ロバート・エイチ・キンケイド

アメリカ合衆国カリフォルニア州 9 4 0 1 9 , ハーフムーンベイ , スピンドリフト・ウェイ・ 5 1  
9

F ターム(参考) 5B080 BA00 FA02



【外国語明細書】

2006139762000001.pdf