

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第4704034号
(P4704034)

(45) 発行日 平成23年6月15日(2011.6.15)

(24) 登録日 平成23年3月18日(2011.3.18)

(51) Int. Cl. F I
 GO 1 N 30/86 (2006.01) GO 1 N 30/86 G
 GO 1 N 30/72 (2006.01) GO 1 N 30/72 C

請求項の数 22 (全 17 頁)

(21) 出願番号	特願2004-509381 (P2004-509381)	(73) 特許権者	509131764
(86) (22) 出願日	平成15年5月30日(2003.5.30)		ウオーターズ・テクノロジーズ・コーポレイション
(65) 公表番号	特表2005-528606 (P2005-528606A)		アメリカ合衆国、マサチューセッツ・01757、ミルフオード、メープル・ストリート・34
(43) 公表日	平成17年9月22日(2005.9.22)	(74) 代理人	100062007
(86) 国際出願番号	PCT/US2003/017190		弁理士 川口 義雄
(87) 国際公開番号	W02003/102543	(74) 代理人	100114188
(87) 国際公開日	平成15年12月11日(2003.12.11)		弁理士 小野 誠
審査請求日	平成18年5月11日(2006.5.11)	(74) 代理人	100140523
(31) 優先権主張番号	60/384,712		弁理士 渡邊 千尋
(32) 優先日	平成14年5月31日(2002.5.31)	(74) 復代理人	100119253
(33) 優先権主張国	米国 (US)		弁理士 金山 賢教

最終頁に続く

(54) 【発明の名称】 クロマトグラフィ／分光測定データの解析でデータピンニングを用いる方法

(57) 【特許請求の範囲】

【請求項1】

2次元解析法を使用して少なくとも2つのサンプルから得た3次元データを解析する方法であって、

(a) 当該サンプルについてデータを得るステップであって、当該データは少なくとも第1パラメーター、第2パラメーター及び第3パラメーターによって特徴づけられ、少なくとも第3パラメーターは時間と相関関係にあり、

(b) 得られた当該データをピンニングして当該サンプルの各々について複数のピンニングされたデータ集合を形成するステップであって、ピンサイズは時間と相関関係にある第3パラメーターに基づいて選択され、各々のピンニングされたデータ集合のデータ点は第1及び第2パラメーターを用いて特徴づけられ、

(c) 当該サンプルの各々について、第1パラメーターを2次元配列の一つの軸と関連付け、ピン番号を2次元配列の第2の軸と関連付けることにより、ピンニングされたデータ集合から2次元配列を形成するステップであって、ピンニングされたデータ集合のデータ点の第2パラメーターの値は2次元配列の要素と関連付けられ、

(d) 当該2次元配列の各々を2次元多変量統計解析にかけるステップ、及び

(e) 解析された2次元配列に関連するデータを比較することにより、少なくとも2つのサンプル間での少なくとも1つの成分のレベルの差異に関連する時間ピンを決定するステップ、

を含み、第1パラメーターがm/zパラメーター、第2パラメーターがイオン強度パラメ

10

20

ーター及び第3パラメーターが保持時間パラメーターである、前記方法。

【請求項2】

前記各ピンが共通の要素をもたない請求項1に記載の方法。

【請求項3】

前記各ピンが部分的に共通するところがある請求項1に記載の方法。

【請求項4】

前記ピンサイズが前記時間パラメータの全範囲を包含するピンから前記時間パラメータの2つ以下の同定し得る最小の増分を包含するピンにまでに及ぶ請求項1に記載の方法。

【請求項5】

前記2次元多変量統計解析が、古典的最小二乗分析(CSL)、主成分分析(PCA)、部分最小二乗分析(PLS)および部分最小二乗判別式分析(PLS-DA)から成る集合から選択される請求項1に記載の方法。

10

【請求項6】

前記2次元多変量統計解析が主成分分析(PCA)である請求項5に記載の方法。

【請求項7】

ステップcの変換が、各データ点が関連していることを前記第1のパラメータの各値が示す場合は前記データ点を併合することをさらに含み、前記併合は前記第2のパラメータの各値を加えることによって達成される請求項1に記載の方法。

【請求項8】

前記3次元データが分光測定法と組み合わせられたクロマトグラフィ法から得られる請求項1に記載の方法。

20

【請求項9】

前記クロマトグラフィ法が液体クロマトグラフィ法である請求項8に記載の方法。

【請求項10】

前記分光測定法が、紫外線分光法、赤外線分光法、質量分析法および核磁気共鳴から成る集合から選ばれる請求項8に記載の方法。

【請求項11】

前記分光測定法が質量分析法である請求項10に記載の方法。

【請求項12】

非時間ベースのパラメータが、前記サンプルの物理的/化学的性質に対応する物理的にアクセスできる変数および各ピンに含まれる各種の変数に対する前記変数の強度であり、前記時間パラメータが R_t の継続時間である請求項8に記載の方法。

30

【請求項13】

前記サンプルの物理的/化学的性質に対応する前記物理的にアクセスできる変数が m/z である請求項12に記載の方法。

【請求項14】

ステップcの変換が、各データ点が同じイオンの同位元素であること示す m/z パラメータを有する前記データ点を併合することをさらに含み、前記併合が各強度値を加えることによる請求項13に記載の方法。

【請求項15】

前記3次元データが液体クロマトグラフィ/質量分析法から得られる請求項8に記載の方法。

40

【請求項16】

(a) サンプルについてクロマトグラフィ/質量分析データを得るステップであって、前記データは、 R_t 、 m/z 及び強度を含む少なくとも3つのパラメーターによって特徴づけられ、

(b) 得られたデータをピンニングして複数のピンニングされたデータ集合を形成するステップであって、ピンサイズは R_t に基づいて選択され、各々のピンニングされたデータ集合のデータ点は、各ピンに関係付けられる1以上の各イオン種についての1つまたはそれ以上の m/z および強度パラメーターの値を用いて特徴づけられ、

50

(c) m/zパラメーターを2次元配列の一つの軸と関連付け、ピン番号を2次元配列の第2の軸と関連付けることにより、ピンニングされたデータ集合から2次元配列を形成するステップであって、ピンニングされたデータ集合のデータ点の強度パラメーターの加算した値は2次元配列の要素と関連付けられ、および

(d) 2次元配列をPCA分析にかけるステップを含む、主成分分析(PCA)を使用して3次元データを解析する方法。

【請求項17】

前記クロマトグラフィ/質量分析データが液体クロマトグラフィ/質量分析法から得られる請求項16に記載の方法。

【請求項18】

ステップcの変換が、各データ点が同じイオンの同位元素であることをm/zパラメータが示す場合は前記データ点を併合することをさらに含み、前記併合は各強度値を加えることによる請求項16に記載の方法。

【請求項19】

前記2次元配列がスプレッドシートに形成される請求項16に記載の方法。

【請求項20】

前記各ピンが共通の要素をもたない請求項16に記載の方法。

【請求項21】

前記各ピンが部分的に共通するところがある請求項16記載の方法。

【請求項22】

前記ピンサイズが前記時間パラメータの全範囲を包含するピンから前記時間パラメータの2つ以下の同定し得る最小の増分を包含するピンにまでに及ぶ請求項16に記載の方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、データピンニングのデータへの応用によってより少ない次元に適用できる解析法により多次元データを解析する方法に関する。特に、本発明は、データピンニングを用いるクロマトグラフィ/分光測定データの解析方法に関する。本発明は、主成分分析と組み合わせられたデータピンニングを用いる液体クロマトグラフィ/質量分析(LC/MS)データ解析を行う方法に特に使用される。

【背景技術】

【0002】

本出願は、参照によりその内容全体が本明細書に特に組み込まれている2002年5月31日に出版した米国仮出願番号60/384,712(整理番号WAA-306)の優先権を主張するものである。

【0003】

メタボノミクス(Metabonomics)は急速に発展している科学的研究の分野である。それは生体代謝特性を研究するためのシステムズアプローチであり、特定の疾患、毒性、および遺伝子機能に関する情報を提供することができる。メタボノミクスにおいて、動物または生物全体に対する薬剤候補の効果は、化合物投与に続くある時間の間に起こる代謝の変化を研究することにより調べられる。これらの研究で生じた分析データは、主成分分析(以後「PCA」と称す)などの多変量数学的手法によって解析される。この数学的手法は調べているサンプルの微妙な差異と全体の差異の両方を強調するために用いられる。

【0004】

今日まで、この分野の研究の大部分が、選択した分析方法として陽子核磁気共鳴(NMR)を利用してきた。NMRは非常に有効であるが、いくつかの欠点、すなわち、感度が劣る、分析に時間が掛かる、およびいくつかの化学物質類、たとえば、硫酸塩類の検出ができないなどの欠点を有している。NMRのさらなる欠点は、すべての信号が1つのスペ

10

20

30

40

50

クトルに含まれるので、1つの大きな化合物が簡単に低濃度分析物の微妙なしかし重要な変化を隠してしまうことである。さらに、生体異物に関連する化合物をNMRスペクトルから除去する必要があるが、この除去が関心のある内生化合物から得られる信号も除去し、そのため後に続くPCA分析に使用するデータ集合を低減してしまう。

【0005】

紫外線分光法、赤外線分光法、核磁気共鳴または質量分析法などの分光測定法と組み合わせられたガスおよび液体の両クロマトグラフィは、強力な手法に進化した。液体クロマトグラフィと結合されたエレクトロスプレー質量分析法(LC/MS)は、定量的および定性的両方のバイオ分析に対して選択される手法になっている。この手法は、確固たるものであり、感度が高く、かつ選択性があり、 $\mu\text{g}/\text{mL}$ までの範囲の感度が容易に達成される。短いカラムおよび急勾配の使用により、LC/MSが良好なクロマトグラフィの分解能と感度を備えた1分当たり1サンプル領域内の分析サイクル回数を達成することが可能になる。それゆえにLC/MSのメタボノミクスへの応用は論理的な手順である。質量分析計が高感度の定量化と構造上の情報の両方を与える一方で、クロマトグラフィのステップは、分析物を分離することにより時間分解能を与え、重なり合う信号の問題に対処する。LC/MSの使用は、クロマトグラフィのステップが任意のある時間における質量分析計に入る競争イオンの数を低減することから、イオン抑制を低減するので、質量分析法の導入以上に好ましい。NMRに対するLC/MSの別の利点はデータ解析において観測される。生体異物に関連する物質を除去する場合、1つまたは2つの特定質量におけるわずか一切れの時間だけがデータ集合から取り除かれることになる。したがって、残りのLC/MSデータは変更されないで残り、数理解析、たとえばPCAなどに利用できる。

【発明の開示】

【発明が解決しようとする課題】

【0006】

主成分分析はMSから得られるデータを解析する非常に有効な数学的デバイスである。しかしながら、PCAは2次元の手法であり、それに対してLC/MSから得られるデータは3次元である。現在、従来のPCA分析において通常失われるクロマトグラフィ分離情報をPCAが保持することを可能にする必要性が存在する。

【0007】

一般に、結果についての情報が、多数のパラメータにより特徴づけられるデータの中に含まれる実験条件が多数ある。説明のツールがこれらのパラメータを取り扱えない場合は情報が失われる。PCA法により分析されるLC/MSデータの解析の分解能を向上する方法は、他のデータ集合に対しても適用できる。

【課題を解決するための手段】

【0008】

本発明は、より少ない次元を取り扱うツールを使用して多次元データを解析する方法に関する。特定の実施形態において、本発明は、主成分分析(PCA)などの2次元多変量統計解析手法を用いて、液体クロマトグラフィ/質量分析(LC/MS)試験などのクロマトグラフィ/分光測定試験から得た3次元データを解析する方法に関する。一般には、PCA分析は2次元データだけを使用して行われる。本発明においては、3次元データを使用して解析が行われる。

【0009】

クロマトグラフィ操作から得られたデータは、いくつかのパラメータによって特徴づけられる。特に、クロマトグラフィは時間の関数としてサンプルを分離する。保持時間、すなわちサンプルの1成分が固相上に保持される時間の長さは特に情報を与えるものである。このサンプルの種々の成分は、条件により異なった時間に固相から溶離する。特性に依存するこの時間は、サンプルの特定の成分を区別する場合に有用である。

【0010】

さらなるパラメータは、クロマトグラフィ分離の後に使用される検出方法に特有のものでもよい。たとえば、蛍光の検出において、関心のあるパラメータは光の波長および強度

10

20

30

40

50

であり、一方、紫外線の検出においては、関心のあるパラメータは吸収された光の量である。多くの場合、解析方法で取り扱うことができるパラメータの数が多ければ多いほど、その結果の識別力が高くなる。1つのパラメータを完全に無視しなければならない場合、そのパラメータだけが異なる各データ点を識別する機会が失われる。

【0011】

LC/MS操作から得られるデータは3つのパラメータから成る。これらのパラメータの第一番目は、このシステムのLC部に固有のものである。液体クロマトグラフィは分析物を時間の関数として分離する。したがって、LC部から得られるデータは時間ベース、すなわち保持時間(R_t)である。種々の分析物が、溶離液などに応じて異なった時間にLCシステム(カラム)の固相エレメントから溶離することになる。この識別分離は、各分析物に対して R_t 特性を考慮している。この情報は、たとえば、特定の分析物の素性を決定する際に役立つことができる。

10

【0012】

第2および第3のパラメータは、LC/MSシステムのMS部に固有のものである。LCを經由してサンプルがいったん分離されると、個々の分析物はこのシステムのMS部に入る。MS部の条件により、分析物はある一定容積内でイオン化され、その後分子質量検出器に向けて加速される。質量電荷比(m/z)および強度は質量検出に従って算出することができる。この情報は2つ以上のサンプル間の差異を求めるのに使用される。

【0013】

一般に、PCA分析は、この m/z および強度パラメータだけを使用してこの定量試験を行うために使用される。それゆえ、 R_t データは全体的な分析には寄与しない。質量分析ではイオンピーク当たりおおよそ5から10原子質量単位(AMU)点がある豊富なデータを利用できるためこの犠牲は許容されてきた。本発明は、データビンニングを組み込むことにより、この過ちに対処するものである。データビンニングにより、PCA分析において R_t 情報が失われず、そのため、3次元データ解析が容易になる。サンプル間の比較は現時点で R_t 、 m/z および強度データを含む。この手法は、大量の複雑なデータの中に単純な疑問に対する答えを見いださなければならない複雑なサンプルの結果を調査するのに特に役に立つ。その結果、従来検出できなかったサンプル間の微妙な変化も現在では解明できる。

20

【0014】

1つの実施形態において、個々のLC/MSクロマトグラムは所定の部分(またはビン)に分割される。これらのビンは R_t パラメータを反映している。このビンのサイズは操作者が決定する。たとえば、ビンサイズは、クロマトグラムを通る時間が1分であることを反映して1分にすることができる。各ビン内で、そのビン時間の間に行われるすべてのスキャンについての強度情報は、対応する R_t 時に存在するそれぞれの種(m/z)について合計される。ビン、またはもっと正確に言えばその中に含まれるデータは、その後、変換され、市販のスプレッドシート、たとえば、Excelなどのデータ行列にエクスポートされる。したがって、元のグラフによるデータは表形式のデータに変換される。変換された各ビン内の情報はクロマトグラフィの時間、 m/z および強度である。2つ以上のビンをこの変換処理にかけることができる。たとえば、60分のLC手順の場合、1分につき1個のビンに相当する60個のビンとすることができる。別法として、重複するやり方で各ビンを構成することができるので、たとえば、40秒につき1回発生する91分用ビンが使用されて60分のLC手順を分析する。変換がいったん完了すると、各データ集合(表内のデータ)はゼロを用いて整列される。この各データ集合の整列は単に隙間を埋める手順を指す。たとえば、 $m/z = 100$ におけるビン3内に強度値102があり、 $m/z = 100$ におけるビン4内に強度値が無い場合は、ビン4内の $m/z = 100$ に対応する強度の位置にゼロが置かれる。言い換えると、ゼロはプレースホルダの役割を果たし、その結果適切なデータ整列が可能になる。したがって、各データ集合は、各 m/z につき1つの強度値(ゼロを含む)を有する。さらに、同位体存在量は一緒に、たとえば、 C^{12} およびその同位体 C^{13} を合わせた値にすることができる。最後に、スプレッドシー

30

40

50

トのファイルが市販のPCAシステム、たとえばPCA分析用MatLabに移される。

【発明を実施するための最良の形態】

【0015】

(詳細な説明)

本発明は、より少ない次元を取り扱うツールを使用して多次元データを解析する方法に関する。本発明は、その後に分光測定方法が続くクロマトグラフィ方法から得たデータの解析に特に応用できる。これらのケースにおけるデータ集合は時間成分、ならびに一般にサンプルの物理的または化学的性質に対する強度の値を有している。特定の実施形態において、本発明は、主成分分析(PCA)を用いて、液体クロマトグラフィ(LC)/質量分析(MS)方法から得たデータを解析する方法に関する。一般には、PCA分析は2次元データだけを使用して行われる。本発明においては、3次元データを使用して解析が行われる。LC/MSおよびPCA分析の詳細説明では、一般的ケースならびにこの特定の実施形態の方法論について示す。

10

【0016】

このシステムのクロマトグラフィ部はサンプルマトリックス中の分析物の分離を可能にする。このような方法から得られた情報は一般に保持時間(R_t)として報告される。この分析物がシステムを通過して進行するとともに、LC/MSシステムのMS領域に入る。それらはそこでイオン化され、その後質量検出器が、 m/z および強度の情報をもたらすこれらイオン化種を検出する。したがって、LC/MSシステムは少なくとも3個の情報を供給することができる。多くの多変量統計解析手法が使用できると同時に、主成分分析(PCA)はデータを解析する確固たる方法であり、異なったサンプルを比較する場合、かなり有効である。しかしながら、PCA分析は一般に2次元データ集合だけを使用して行われる。したがって、LC/MS操作から得られるデータの少なくとも1つのパラメータが失われることが予期される。本発明においては、PCA分析に加えて3次元データ集合を使用して解析が行われる。

20

【0017】

ガスクロマトグラフィ(GC)はガス状のサンプルに対して同じ様なデータを与えるが、サンプルマトリックス中の分析物を分離するのに一般に用いられる方法はクロマトグラフィ、特に、液体クロマトグラフィ(LC)である。高性能液体クロマトグラフィ(以下において「HPLC」と称す)により、あるサンプル内の分析物の迅速で効率的な分離および特性記述が可能になる。HPLCシステムの構成部品には、システムを通る水相および/または有機相の移動を容易にする高圧ポンプが含まれる。この水相(すなわち移動相)は、HPLCシステムを最初に平衡させるのに使用される溶剤を含む。この溶剤は、HPLCシステム全体を通過して移動する分析物に水性環境も与える。最後に、移動相はHPLCカラムから分析物を溶離する溶剤を含む。

30

【0018】

HPLCシステムの別の構成部品はクロマトグラフィ分離カラムである。このカラムは固相を含む。この固相は移動相と組み合わせ、サンプルマトリックスに含まれる分析物の時間差のある分離を行う。この固相は一般に、ある種の分析物と相互に作用する化学重合体から成る。たとえば、逆相HPLCカラムは、疎水性の力によって分析物と相互に作用する固相の化学的性質(たとえば、シリカのビーズに付いている18個の炭素から成る炭化水素鎖)を有している。

40

【0019】

分析物がいったんカラムから溶離されると、検出器内に移動しそこを通りぬける。HPLCシステムで用いることができるものとして多様な検出器システムがある。たとえば、紫外線(「UV」)域内の分析物を検出するUV検出器、赤外線(IR)域内の分析物を検出するIR検出器、電解電量法を使用して分析物を検出する電気化学検出器、磁場に対する反応に基づいて分析物を検出する核磁気共鳴検出器および質量電荷比に基づいて分析物を検出する質量分析検出器がある。

【0020】

50

LC/MSシステムにおいて、サンプルマトリックスは最初LC分離にかけられ、これによってサンプル内の分析物の分離を行う。多くの場合分析物の R_t によって伝えられるこの情報は、分析物の特性を記述し、分析物を特定するのに使用することができる。しかしながら、特に異なるサンプル間で比較を行うために、分析物をMS検出にかけることによってさらなる情報を得ることができる。一般に質量分析計は、ある容積内の分子をイオン化し、その後イオン化された分子を分子質量検出器に向けて加速する。分子のイオン化は、電子イオン化、化学イオン化、エレクトロスプレーイオン化、または光イオン化により生じ得る。このイオン化プロセスは、大気圧を含む種々の圧力条件下で発生することができる。MSから得られる情報には特定の m/z 値に対する応答の強度が含まれる。2つ以上のサンプルを比較する場合に重要なのはこの強度（ならびに m/z 値）情報である。

【0021】

MS情報は比較分析に利用することができる。たとえば、人間の細胞系について実験しているとする。この研究の目標は、もしあるとすれば、化合物Xは特定の代謝産物に対してどのような作用を及ぼすか判定することである。話を続けて、操作者は、対照サンプルすなわち化合物Xがななら加えられていない細胞培養からの細胞抽出であるサンプル1と、化合物Xが加えられた細胞培養系から取った細胞抽出であるサンプル2の2つのサンプルを得たとする。これら2つのサンプルはLC/MSにかけることができる。さらに、関心のある代謝産物が、逆相カラムを用いるクロマトグラフィ単離に基づいて分析できるとする。LCを通過していったん分離されると、サンプル1とサンプル2のサンプルは別々にMSにかけることができる。2つのサンプルから発生したMS信号は、その後、特に関心のある代謝産物について比較を行って、化合物Xの添加によってその強度が増加するか、減少するか、または影響を受けないか決定することができる。

【0022】

図1は、LC/MS実験に適用した場合の本発明の様々なステップを記した流れ図を示す。図1aは、仮想サンプルに関してLC/MS手順から得られたデータをグラフ形式で示す。図示されるように、このグラフは3つのパラメータを規定している。X軸は m/z パラメータを規定しており、たとえば、X軸は100から800 m/z まで0.1 m/z 刻みで延びることができる。図1aのY軸は強度または応答パラメータを表す。強度は、あるイオンのパーセンテージ、たとえば基底のピークに設定することができる。最後に、時間パラメータを表すZ軸がある。図1aに示されているのは、3つのパラメータすべてではないにしても少なくとも1つが異なるイオン種である。それでも依然として各イオン種は図1aのパラメータによって一意的に定められる。

【0023】

現存のPCA分析では、図1aに示されたデータを分析前に2次元データ集合に変換することが要求される。たとえば、従来のPCA分析は m/z および強度の各パラメータを使用して実行することができる。しかしながら、このことにより時間パラメータが削除される。より完璧な分析は、時間次元を含め入手できるすべての関連情報を含む。図1bは、3つのパラメータすべてを保存し分析にける方法を表している。データビンニングは、多数の測定点に相当する1つのデータ集合を、測定点を集めた多数のデータ集合に分離するプロセスである。ピンは1つのパラメータの範囲、すなわちピンパラメータに対応する。各ピン内のピンパラメータの変動は、ピンの値におけるピン内のすべてのデータ点を累積することなどにより除かれる。データ集合全体を包含するピンはピンパラメータを無視するのと同じ効果をもつことに注目すること。共通の要素をもたないピンを用いることは最も一般的であるが、ある分析に対しては一部共通するところがあるピンを使用してもよい。図1aから図1bへの移行において、所定のピンサイズが選択されるとデータビンニングのプロセスが生じる。ピンサイズは時間次元の関数である。たとえば、ピンサイズは2分にすることができる。多数のピンが図1bに示されている。各ピンは、スキャン全体で見いだされ、そのピンの範囲内に入るイオン種についての m/z および強度の両方の情報を含んでいる。ピン（または領域）サイズは、経験および試験されるサンプルに基づいて従業者によって選択される。選択方法は当業者によく知られている。たとえば、ソフ

10

20

30

40

50

トウェアの *Mass Lynx* を用いることができる。このソフトウェアを使用して、ピンが以下のように集められる。(1) 全体のイオンクロマトグラムを開く、(2) マウス右クリックでピン領域を選択しピンエリア上にドラッグすれば、すべてを含めた質量スペクトル(ピン)が現れる、(3) すべてを含めた質量スペクトル(ピン)のウィンドウ内に「リストスペクトル(List Spectrum)」のオプションが現れ、始動する、(4) リストスペクトル(List Spectrum)が選ばれると、ウィンドウが対応する存在量を備えた m/z を 2 行形式で表示する、(5) 2 つのカラム内に含まれるデータを複写する、および市販の(6) *Excel* のようなスプレッドシートに貼り付ける。

【0024】

データピンニングプロセスの後に、変換ステップがある。(図 1 b から図 1 c 参照)。この変換は、得られた分離リストをグラフによるデータ形式から行列データ表現に変換する。従業者は、単に「切り取りおよび貼り付け」手順を用い、2 行のデータ集合情報を *Excel* のような市販スプレッドシートに与えることでこの変換を始めることができる。変換のこの部分がいったん完了すると、データ集合はゼロを用いて整列させる必要がある。各ピンは、分析されるすべてのイオンの値を有する必要がある。分析の忠実度を確保するために、そのピンについてイオン信号が観測されないところではゼロが強度値として挿入され、各データ集合が各 m/z パラメータに対して 1 つの強度値を有する結果になる。その後、 m/z パラメータの各行は、行列の 1 つの軸として作用する 1 つの行に融合させることができる。別法として、この変換を自動化するためにソフトウェアプログラムを使用することができる。表の形式で表したものが図 1 c に示されている。図 1 c には m/z の行とピンの列がある。 m/z の行は分析される m/z イオンを含んでいる。ピンの列は分析される望ましいピンを含んでいる。行列の交差部分は、適切な m/z およびピンパラメータに対応する強度値を含んでいる。さらに、データを解析にかける前に、試験によって引き起こされた m/z および強度における知られた変換に基づいて、操作者がデータ操作を行ってもよい。たとえば、同位体存在量は一緒にすることができ、その結果、たとえば C^{12} およびその同位体 C^{13} に対する各値が一緒にされる。

【0025】

この行列またはスプレッドシートは次に、各種サンプル間の比較を行うことができる市販の *PCA* プログラムに移され、解析される。このような市販の *PCA* プログラムの 1 つが、マサチューセッツ州ネーティックの *MathWorks* 社が販売している *MatLab* である。別の *PCA* プログラムパッケージは、ワシントン州ウディンビル(*Woodinville*)にある *InfoMatrix* 社から入手でき、*Pirouette* と呼ばれている。

【0026】

複数のスキヤンの結果として生じたデータをピンニングを通じて合同することにより質量分析法特有の高分解能のいくつかをばかすことによって、2 次元解析は得られた 3 次元データを取り扱うことができる。これにより、試料間の変化が予期されるが種々のスキヤンにおける変化の位置は分からない大量のデータを解析することが可能になる。変動の位置が分かった後で、後半の分析は、生データ本来の高分解能を使用してさらに差異の特徴を調べることができる。このピンニング手法は、データ集合を多くのタイプの多変量統計解析に適合させるように使用することができる。

【0027】

本発明の特徴およびその他の詳細について以下の実施例でより詳細に説明し指摘する。本発明の特定の実施形態を図を用いて示すが、本発明を限定するものとして示すものではないことが理解されるだろう。本発明の主な特徴は発明の精神と範囲から逸脱することなく様々な実施形態で使用することができる。

【実施例】

【0028】

ラットの尿サンプル

合計20匹のラットのサンプルがこの実験に使用された。これらのラットは4つのグループに分けられた。グループIは投与マトリックスだけを受け入れ、グループII、IIIおよびIVは化合物A、BおよびCをそれぞれ経口投与された。ラットの尿のサンプルは2つの時点、投与後0～8および8～24時間に集められた。ラットの尿は分析前に-20℃に冷凍保存された。

【0029】

クロマトグラフィ

クロマトグラフィは、カラムオープンを備えたWaters社のAlliance 2795HTおよびWaters社の2996PDA検出器を使用して実施した。ラットの尿(蒸留水で1:4に希釈)を10cm×2.1mmのWaters社Symmetry C₁₈ 3.5µmカラムに20µL/分で注入した。カラムは40℃に保持し、流量600µL/分において勾配条件下、溶離液「A」は0.1%蟻酸水溶液、溶離液「B」はアセトニトリルとして溶離した。カラムは、最初の1分間100%Aに保持し、その後、移動相を次の7分間にわたって30%Bまで急激に増加させた。その後、移動相を0.1分間にわたって95%Bまで急激に増加させた。この移動相は、カラムを洗浄するために0.5分間保持され、その後0.1分間にわたって最初の状態に戻された。最初の注入の10分後に次の注入が行われた。

【0030】

質量分析

質量分析の解析は、エレクトロスプレーインターフェース(「ESI」)を備えたMicroMass社Quattro Microで行われた。この計測器は陰イオンモードでキャピラリー電圧3kVの状態で作動され、コーン電圧は25Vに設定、ネブライザガスは脱溶媒温度150℃およびガス供給温度70℃の状態で作動され600L/hrに設定された。カラム溶離液は5:1の比率に分割され、その結果100L/hrが質量分析供給源に入った。この計測器はスキャン時間200ミリ秒およびスキャン間遅れ時間50ミリ秒で100から800m/zまでスキャンするフルスキャンモードで操作された。データは0分から10分まで集められた。クロマトグラムの最初の部分は導入されず廃棄した。

【0031】

データ解析

段階Iおよび段階IIの代謝産物などの任意の薬物関連生成物の存在について、LC/MSクロマトグラムを検査した。個々のLC/MSクロマトグラムはその後それぞれ10画分(1画分100スキャン)に分割された。100スキャンの個別クロマトグラム画分のそれぞれに対してMS信号が組み合わされた。ピークリストおよびイオン存在量はその後Excelファイルにエクスポートされた。次に、イオン信号が観測されない箇所にゼロが強度値として挿入されてデータ集合が整列され、その結果各データ集合が、100から800の各m/z整数に対して1つの強度値を有した。その後、社内のプログラムを使用してExcelファイルがPCA分析用Matlabに移された。

【0032】

検討

ラットの尿は数千の化合物を含む可能性があり、その大部分は不明のままである。これら化合物の大多数は中程度の極性かまたは高い極性であり、したがって、尿サンプルの全成分の完全な溶離を行うには有機含有率が0から30%の勾配を用いるだけでよい。陰イオンエレクトロスプレーイオン化(ESI)MSは、陽イオンESIよりも情報が多いデータ集合をもたらすので、MS検出モードとして陰イオンを選んだ。この研究で用いたクロマトグラフィのカラム長は、勾配時間、および溶離液流量は、サンプル処理量とクロマトグラフィ分解能の間で最良のバランスを得るように選んだ。その結果、サンプル処理量は1時間当たり6サンプル、ピークキャパシティは78となった。このピークキャパシティは非常に大きくはないが、100から800m/zをスキャンする質量分析法の使用ははるかに大きな有効ピークキャパシティをもたらす。本研究においては、単純な逆相クロマトグラフィシステムを用いた。アミノ酸や糖などの極性の高い多くの化合物は保持され

10

20

30

40

50

ないことが十分に了解されるであろう。このシステムを変更することによってこれら化合物の解明が可能になる。

【0033】

対照ラットの尿と時点1で投与された動物の尿サンプルのベースピーク強度（「BPI」）クロマトグラムの予備的な比較によって定性的な差異が明らかになっている（図2参照）。これらのクロマトグラムから確認できるように、投与されたサンプルグループのクロマトグラムには、対照グループのそれと比較した場合に高いレベルを示すいくつかのピークがある。これらのピークは、クロマトグラムの5～9分の時間域で溶離する。これらピークの徹底した調査によって、ピークはいずれも投与した化合物に起因する代謝産物ではないことが確認された。したがってこれらのピークは、動物の代謝状態の変化によるものでなければならない。

10

【0034】

化合物「A」の場合の時点2のサンプルから得られたBPIクロマトグラムは、時点1のサンプルと比較した場合、いくつかのピーク強度におけるいくつかの微妙な変化を示している（図3参照）。時点2のサンプルにおいて、5から7分の間に溶離する分析物は、時点1の同一分析物と比較した相対濃度を低下させているようである。化合物「B」および「C」のサンプルは、時点1と時点2の間の変動が少ないことを示している（図4および5参照）。これらの時点1と時点2のサンプルは昼と夜のサンプルなので、生体異物の投与が無くても2つのサンプルタイプの間には著しい変化が見られることが予期される。このことは、昼と夜の間の動物代謝の変化に関連した化合物は保持されなかったかまたは検出されなかったことを示唆している。

20

【0035】

対照サンプルと比較して投与サンプルで変化すると識別されるピークを表1に挙げる。（ $m/z = 338$ におけるこれらイオンの1つの抽出されたイオンクロマトグラムおよびこのピークのMSスペクトルが図6に示されている。）

【0036】

【表1】

表1

投与化合物	分析物 m/z 値	相対変化
A	283	10倍の増加
A	461	5倍の増加
A	187	10倍の増加
B	338	2倍の増加
B	283	10倍の増加
B	461	10倍の増加
B	187	10倍の増加
C	283	20倍の増加
C	187	30倍の増加

30

40

【0037】

表1のデータは、対照サンプルおよび投与サンプルのこれらのピークについて観察されたピーク強度の変化を示している。このデータから観察できるように、ピーク濃度の変化は非常に著しい。これらの化合物は対照サンプル中にも観測されるという事実は、これらのピークが投与化合物の代謝生成物ではないことを立証している。

50

【0038】

データ集合全体のPCA分析は図7に呈示されている。この図で、対照サンプルに関するデータは丸で囲んだ部分内に含まれる。この対照サンプルは投与動物のサンプルから明らかに分けられている。このPCA分析は、保持時間および m/z 値のデータ集合全体を使用して実施される。このクロマトグラムを10個の時間域に細分化することによって、同様にクロマトグラムの特定時間域でPCA分析を行うことが可能になった。

【0039】

5分から6分のデータのPCA分析が図8に呈示されている。この図では、丸で囲んだ部分内部に含まれている対照サンプルと投与サンプルとを分離することが可能である。1つに合わせたMSスペクトルは、投与サンプルおよび対照サンプルをPCAプロットの異なる領域に位置づけることができるMSピークは $m/z = 283$ および $m/z = 261$ であったことを示している。 $m/z = 283$ のイオンは投与サンプルの強度が増加したのに対して、 $m/z = 261$ のイオンは投与サンプルの強度が低下した。この結果は、LC/MSデータを解析することによって得られた観測値を確認している。

【0040】

図9は3分から4分のLC/MSデータのPCA分析を示している。このデータでは対照サンプルと投与サンプルは分離されていない。このデータは、クロマトグラムの領域内では、投与サンプルと対照サンプルの間に有意な差が無かったことを示している。

【0041】

PCAデータ分析に結合されたLC/MSの応用は、3つの候補薬剤の投与に続くラットの尿の選別を使用することに成功した。この方法を使用して、対照サンプルと投与サンプルを区別することができた。PCA分析を使用して、分離を担うMSスペクトルの各成分を特定することもできる。このデータは、薬の発見および開発におけるメタボノミックな応用のために、LC/MSが陽子NMRに取って代わる実行可能な方法であるかまたはこれを補完させるものであることを明確に実証している。

【0042】

本発明をその実施形態に関連して特に示し、説明したが、本明細書に添付した請求項で定義したように本発明の精神と範囲から逸脱することなく形態と詳細の数多の変更が可能であることを当業者は了解するであろう。

【図面の簡単な説明】

【0043】

【図1A】本発明の実施形態に伴う各ステップを示す概略図である。

【図1B】本発明の実施形態に伴う各ステップを示す概略図である。

【図1C】本発明の実施形態に伴う各ステップを示す概略図である。

【図2A】対照ラットおよび投与ラットに対する陰イオンBPILC/MSクロマトグラムの比較図である。

【図2B】対照ラットおよび投与ラットに対する陰イオンBPILC/MSクロマトグラムの比較図である。

【図2C】対照ラットおよび投与ラットに対する陰イオンBPILC/MSクロマトグラムの比較図である。

【図2D】対照ラットおよび投与ラットに対する陰イオンBPILC/MSクロマトグラムの比較図である。

【図3A】化合物Aを投与したラットの時点1および時点2における陰イオンBPILC/MS尿クロマトグラムの比較図である。

【図3B】化合物Aを投与したラットの時点1および時点2における陰イオンBPILC/MS尿クロマトグラムの比較図である。

【図4A】化合物Bを投与したラットの時点1および時点2における陰イオンBPILC/MS尿クロマトグラムの比較図である。

【図4B】化合物Bを投与したラットの時点1および時点2における陰イオンBPILC/MS尿クロマトグラムの比較図である。

【図 5 A】化合物 C を投与したラットの時点 1 および時点 2 における陰イオン B P I L C / M S 尿クロマトグラムの比較図である。

【図 5 B】化合物 C を投与したラットの時点 1 および時点 2 における陰イオン B P I L C / M S 尿クロマトグラムの比較図である。

【図 6 A】抽出したイオンクロマトグラムの図である。

【図 6 B】 m/z 338 のピークについての M S スペクトル図である。

【図 7】ラットの尿サンプル 1 ~ 24 から成る全データ集合 P C A プロット図である。

【図 8 A】統合して重ね合わせた質量スペクトルの図である。

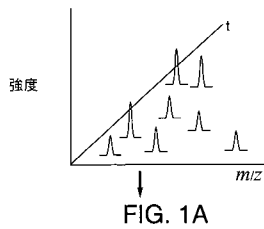
【図 8 B】5分から6分の部分の P C A プロット図である。

【図 9 A】統合して重ね合わせた質量スペクトルの図である。

【図 9 B】3分から4分の部分の P C A プロット図である。

10

【図 1 A】



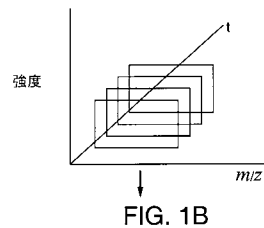
【図 1 C】

m/z	1	2	3	4	ピン
100	20	30	50	0	
120	50	0	100	10	
140	30	10	0	115	
160	100	15	20	60	
180	10	80	120	85	
200	45	30	130	20	
220	20	35	45	10	

↓
PCA

FIG. 1C

【図 1 B】



【 図 2 A 】

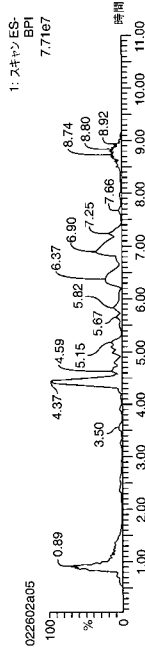


FIG. 2A

【 図 2 B 】

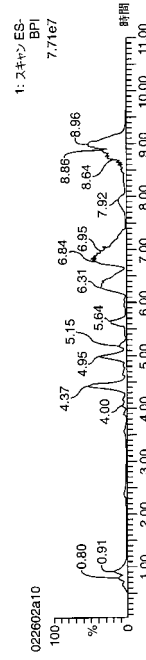


FIG. 2B

【 図 2 C 】

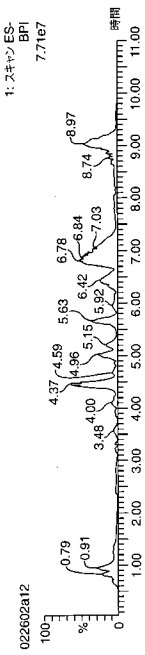


FIG. 2C

【 図 2 D 】

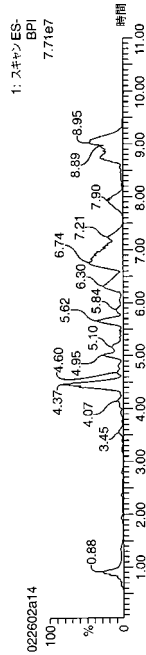


FIG. 2D

【 図 3 A 】

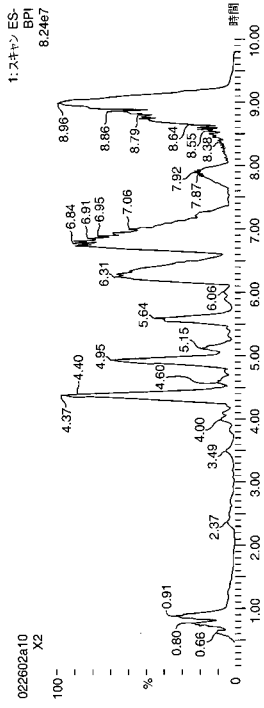


FIG. 3A

【 図 3 B 】

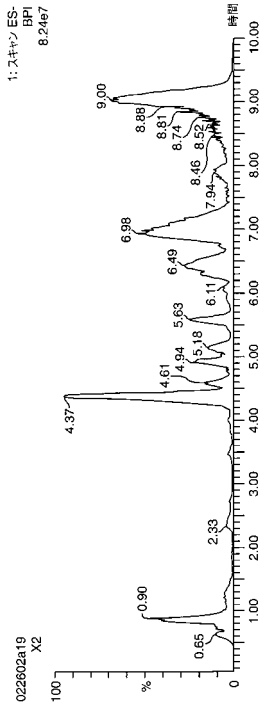


FIG. 3B

【 図 4 A 】

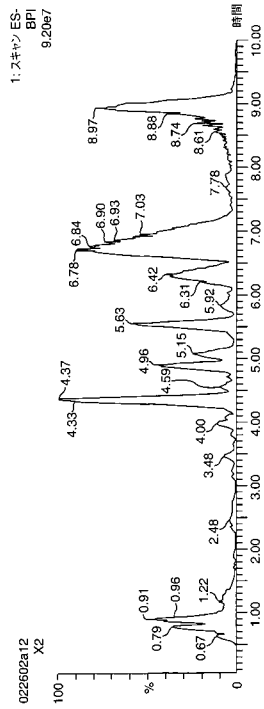


FIG. 4A

【 図 4 B 】

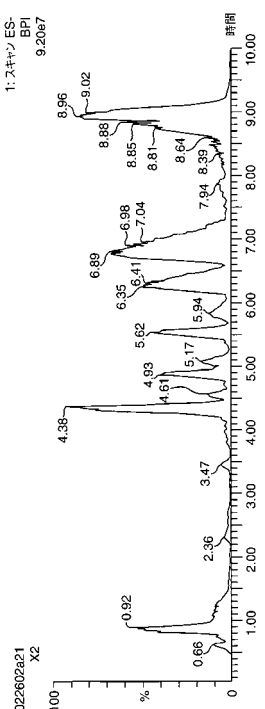


FIG. 4B

【 図 5 A 】

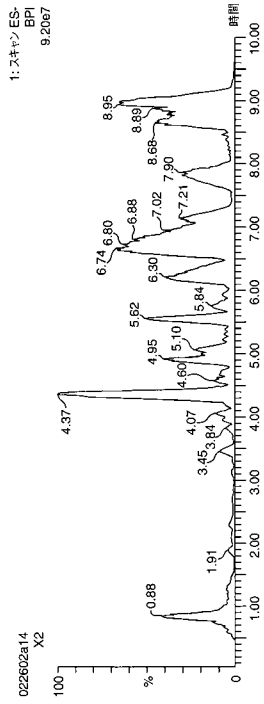


FIG. 5A

【 図 5 B 】

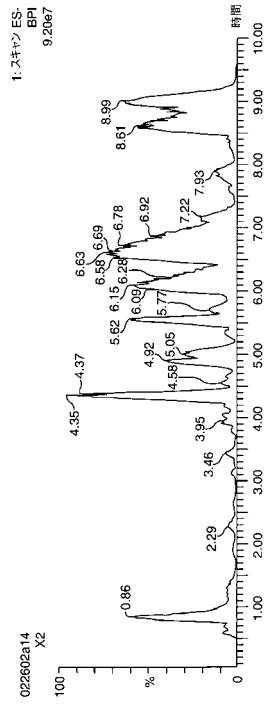


FIG. 5B

【 図 6 A 】

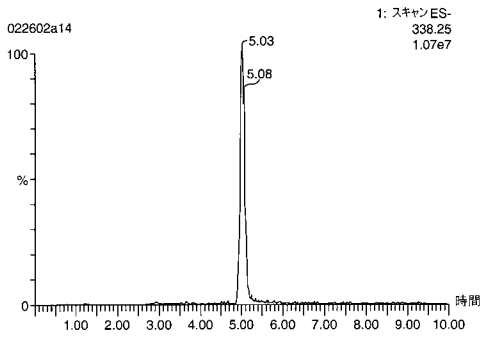


FIG. 6A

【 図 7 】

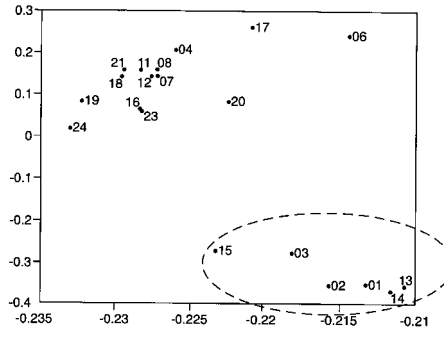


FIG. 7

【 図 6 B 】

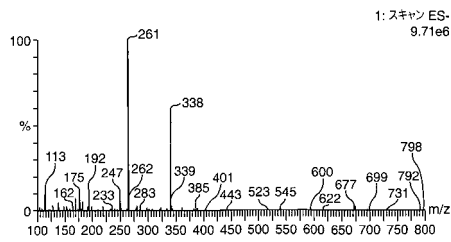
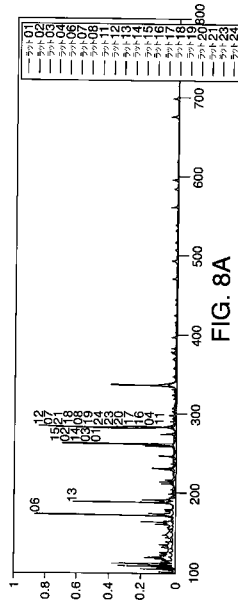
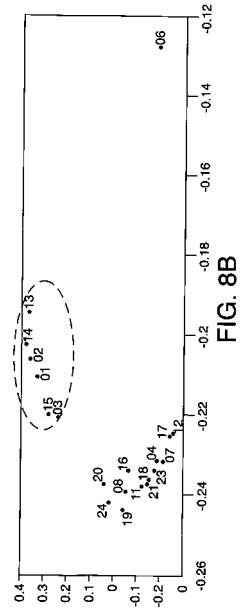


FIG. 6B

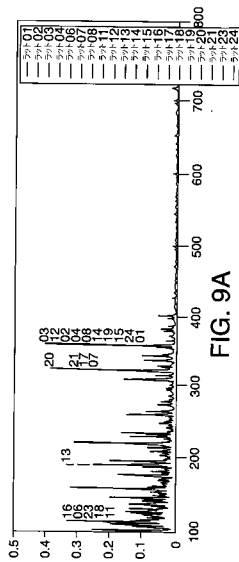
【 8 A 】



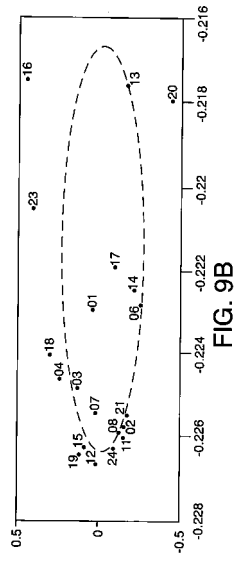
【 8 B 】



【 9 A 】



【 9 B 】



フロントページの続き

- (74)代理人 100103920
弁理士 大崎 勝真
- (74)代理人 100124855
弁理士 坪倉 道明
- (74)代理人 100113332
弁理士 一入 章夫
- (74)代理人 100119253
弁理士 金山 賢教
- (72)発明者 プラム, ロバート・ステイブン
アメリカ合衆国、マサチューセッツ・01757、ミルフオード、マツギル・レーン・6
- (72)発明者 スタンフ, クリス・リー
アメリカ合衆国、マサチューセッツ・01757、ミルフオード、リンドン・ロード・4、アパートメント・9
- (72)発明者 ゴーレンシユタイン, マーク・ブイ
アメリカ合衆国、マサチューセッツ・02492、ニーダム、ブルツクライン・ストリート・313

審査官 河野 隆一朗

- (56)参考文献 特開2001-028252(JP, A)
特開2002-005890(JP, A)
特開平08-334493(JP, A)
特開平11-326304(JP, A)
特開昭63-308560(JP, A)
国際公開第02/003056(WO, A1)
国際公開第03/077552(WO, A1)
米国特許出願公開第2003/0058239(US, A1)
Frøde Brakstad, The feasibility of latent variables applied to GC-MS data, Chemometrics and Intelligent Laboratory Systems, 1995年10月, Vol.29, No.2, P.157-176

(58)調査した分野(Int.Cl., DB名)

G01N 30/86
G01N 30/72
JSTPlus(JDreamII)
JST7580(JDreamII)