

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第5750108号
(P5750108)

(45) 発行日 平成27年7月15日 (2015. 7. 15)

(24) 登録日 平成27年5月22日 (2015. 5. 22)

| | | |
|--------------------------|---------------|---|
| (51) Int. Cl. | F I | |
| C 1 2 Q 1/02 (2006. 01) | C 1 2 Q 1/02 | |
| G 0 6 F 19/24 (2011. 01) | G 0 6 F 19/24 | |
| C 1 2 M 1/34 (2006. 01) | C 1 2 M 1/34 | A |

請求項の数 25 (全 29 頁)

| | | | |
|---------------|-------------------------------|-----------|-----------------------|
| (21) 出願番号 | 特願2012-529340 (P2012-529340) | (73) 特許権者 | 508158676 |
| (86) (22) 出願日 | 平成22年9月21日 (2010. 9. 21) | | プラスティセル リミテッド |
| (65) 公表番号 | 特表2013-505012 (P2013-505012A) | | イギリス国 エヌダブリュ1 ○エヌエイ |
| (43) 公表日 | 平成25年2月14日 (2013. 2. 14) | | チ, ロンドン, ロイヤル カレッジ スト |
| (86) 国際出願番号 | PCT/GB2010/001768 | | リート 2, ザ ロンドン バイオサイエ |
| (87) 国際公開番号 | W02011/033274 | | ンス イノベーション センター |
| (87) 国際公開日 | 平成23年3月24日 (2011. 3. 24) | (74) 代理人 | 100103285 |
| 審査請求日 | 平成25年9月12日 (2013. 9. 12) | | 弁理士 森田 順之 |
| (31) 優先権主張番号 | 0916585.3 | (72) 発明者 | ジョンソン、クリストファー、ジェームス |
| (32) 優先日 | 平成21年9月21日 (2009. 9. 21) | | イギリス、ロンドン エスイー24 9エ |
| (33) 優先権主張国 | 英国 (GB) | | ヌエックス、ハーンヒル、フランクフルト |
| (31) 優先権主張番号 | 1011722.4 | | ロード 11ビー |
| (32) 優先日 | 平成22年7月12日 (2010. 7. 12) | | |
| (33) 優先権主張国 | 英国 (GB) | | |

最終頁に続く

(54) 【発明の名称】 細胞培養データを処理するための装置および方法

(57) 【特許請求の範囲】

【請求項 1】

細胞培養データを処理する方法であって、前記データは多数の試料の結果を含み、前記結果は前記各試料について細胞培養の複数の段階を順次実行することによって得られ、前記各段階は一組の特定条件をもつ細胞培養処理を表し、そのため前記各試料は前記細胞培養に適用された前記処理の固有性および順序によって定義される手順をたどり、

望ましい細胞培養成果を生み出した前記試料のサブセットを特定する工程と、

前記サブセット内の前記試料の前記結果をコンピュータで解析して前記結果を順序付けまたは分類する工程とを含み、前記結果を順序付けまたは分類する工程は前記望ましい細胞培養成果を得るために効果的な一つ以上の手順の特定に役立ち、前記順序付けまたは分類するための前記解析は異なる手順の間の類似性に関する情報を利用する、方法。

10

【請求項 2】

前記順序付けまたは分類するための前記解析は、前記各手順をたどる前記サブセットの試料の数をさらに利用する、請求項 1 に記載の方法。

【請求項 3】

前記異なる手順の間の前記類似性を評価する工程は、前記異なる手順の間の隔たりの程度を決定する工程を含む、請求項 1 または 2 に記載の方法。

【請求項 4】

前記異なる手順の間の前記隔たりの程度を決定する工程は、関心のある前記手順に共通の段階の数を数える工程を含む、請求項 3 に記載の方法。

20

【請求項 5】

前記隔たりの程度は、前記手順の個々の段階の処理の類似性の程度を与える、請求項 3 に記載の方法。

【請求項 6】

前記解析は前記サブセット内の前記各試料の記録を含む一組のデータに対して実行される、請求項 1 乃至 5 のいずれかに記載の方法。

【請求項 7】

前記各記録は、前記試料の識別子および前記試料に適用された前記手順に関する情報を含む、請求項 6 に記載の方法。

【請求項 8】

前記試料に適用された前記手順に関する前記情報は、前記試料に適用された前記処理の順序の一覧を含む、請求項 7 に記載の方法。

【請求項 9】

前記順序の一覧は二進列で表され、前記二進列の各桁は異なる段階の異なる処理に対応し、前記桁の値は前記処理がその特定の段階で前記試料に適用されたか否かを示す、請求項 8 に記載の方法。

【請求項 10】

前記結果を分類および順序付けする前記工程は、前記結果を集団化する工程を含む、請求項 1 乃至 9 のいずれかに記載の方法。

【請求項 11】

前記集団化する工程は階層的に集団化する工程を含む、請求項 10 に記載の方法。

【請求項 12】

前記階層的に集団化する工程はワード連鎖に基づく、請求項 11 に記載の方法。

【請求項 13】

前記集団化する工程は自己組織化マップを含む、請求項 10 に記載の方法。

【請求項 14】

前記分類および順序付けする工程を用いて追加検討のための細胞培養処理を特定する工程をさらに含む、請求項 1 乃至 13 のいずれかに記載の方法。

【請求項 15】

前記特定した細胞培養処理を前記追加検討する工程をさらに含む、請求項 14 に記載の方法。

【請求項 16】

前記方法は、望ましい細胞培養成果を生み出した試料の少なくとも前記サブセットの測定値を解析して、前記サブセット内の前記各試料用の前記手順を決定する工程をさらに含む、請求項 1 乃至 15 のいずれかに記載の方法。

【請求項 17】

前記測定値によってその試料用の前記手順の信頼性のある決定ができない場合は、試料の前記結果を廃棄する工程をさらに含む、請求項 16 に記載の方法。

【請求項 18】

前記測定値は流れ細胞測定または走査細胞測定によって得られて前記試料に関連する蛍光標識を識別し、異なる処理は前記試料に異なる蛍光標識を与える、請求項 16 または 17 に記載の方法。

【請求項 19】

前記望ましい細胞培養成果は一つ以上の試験に合格することによって決定され、結果の前記分類を解析する工程をさらに含んで前記一つ以上の試験に合格する試料の異なる分類を特定し、前記異なる分類は異なる細胞培養特性をもつ、請求項 1 乃至 18 のいずれかに記載の方法。

【請求項 20】

前記異なる分類は異なる細胞表現型に対応する、請求項 19 に記載の方法。

【請求項 21】

10

20

30

40

50

前記細胞培養の前記複数の段階を実行して処理のための結果を生成する工程をさらに含む、請求項 1 乃至 20 のいずれかに記載の方法。

【請求項 22】

請求項 1 乃至 21 のいずれかに記載の方法を実行するように構成された装置。

【請求項 23】

細胞培養データを処理する装置であって、前記データは多数の試料の結果を含み、前記結果は前記各試料について細胞培養の複数の段階を順次実行することによって得られ、前記各段階は一組の特定条件をもつ細胞培養処理を表し、そのため前記各試料は前記細胞培養に適用された前記処理の固有性および順序によって定義される手順をたどり、

望ましい細胞培養成果を生み出した前記試料のサブセットを特定するデータを含む記憶装置と、

10

前記サブセット内の前記試料の前記結果をコンピュータで解析して前記結果を順序付けまたは分類するよう構成された演算素子とを備え、前記結果の順序付けまたは分類する工程は前記望ましい細胞培養の成果を得るために効果的な一つ以上の手順の特定に役立ち、前記順序付けまたは分類するための前記解析は異なる手順の間の類似性に関する情報を利用する、装置。

【請求項 24】

前記装置は前記試料を受け取り、前記結果を決定する測定システムをさらに備える、請求項 23 に記載の装置。

【請求項 25】

20

前記測定システムは細胞測定システムを備える、請求項 24 に記載の装置。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は細胞培養に関し、具体的には多数の細胞培養試料に由来するデータを処理する装置および方法に関する。

【背景技術】

【0002】

近年、生命科学分野において細胞培養が中核技術になってきた。細胞培養の基礎となる科学が複雑であるため、処理および成長条件の違いが及ぼす影響はまだ十分に解明されていない。多くの細胞培養処理は、恐らく既存の処理からの類推によって試行錯誤的に開発されている。しかし、この方法は時間が掛かかり、不確実であり、かつ明らかに非効率である。

30

【0003】

多数の個別の段階が関与する細胞培養の手順を、考案し最適化することは特に難しい。ある段階の処理の変更が後続の段階の性能に影響する可能性があるため、処理の最適な組合せを考案するのは非常に困難であり、多数の実験が必要とされる。そのような実験は、自動細胞培養などの処理能力が高い方法が知られてはいるが、従来の細胞培養によって実施されている。これらの実験には、微小流体プラットフォーム（例えば非特許文献 1、Tourovskaia 等、2005 年）または細胞配列（例えば非特許文献 2）を使用するなど、細胞培養を小型化する方法を取り入れることもできる。当技術分野では、処理能力の高い細胞培養技術によって作られたデータを解析する方法が求められている。

40

【0004】

特許文献 1 および特許文献 2 には細胞が多数の異なる単位で培養される技術が説明されており、その内容を参照によって本明細書に援用する。各単位はビーズから形成することができ、その表面または細孔内で細胞が成長する。細胞単位（ビーズ）は別々の群に分割され、各群は特定の処理を受ける。この第 1 の段階（ラウンド）の後、細胞単位は必要に応じて再び一緒に貯留され、その後もう一度新しい群に分割される。新しい群はその後、第 2 のラウンドの処理に供される。さらに貯留、分割および処理のラウンドが続くこともある。細胞単位には、培養処理中に必要に応じて標識が付けられるため、実験の最後に所

50

与の細胞単位に適用された一連の処理を推定することができる。例えば、ふるい分け分析による判定時に特定の細胞種類が生育しているなど、望ましい終点に到達している細胞単位を識別することができ、それらが受けた一連の処理を識別することができる。

【0005】

これらの実験における細胞単位の数是非常に多く、数千あるいはそれ以上の場合がある。同様に、異なる細胞単位が受けてきた処理（すなわち手順）の組合せ候補の数もまた多数になる場合がある。例えば、3つの段階のそれぞれに10個の処理候補があるとする、1000（ 10^3 ）個の手順候補ができる。実験に50,000個の細胞単位が含まれるとすれば、それぞれの手順に平均50個のビーズが供されることになる。

【0006】

このような大規模なふるい分け実験の結果は検証を必要とする。というのは、一般にこれらの結果には、望ましい結果が得られているがその成果が当てにならない偽陽性、および潜在的に生産的な手順をたどる細胞単位が有効な成果をもたらさない偽陰性が含まれるからである。偽陰性の数から所与の手順の効率がある程度分かり、かなり低いことが多い（10%またはそれ以下）。

【0007】

一般に、これらの大規模な細胞培養実験の結果を解析する既存の方法は、陽性の結果を引き出した手順を最終検査の中で探すことである。その後、成功した手順は追加実験の対象となる。この追跡作業には、結果の統計値が良好な手順についてさらに多数の細胞単位を試験する、または（小さなビーズを使用するのではなく）従来のすなわち単分子層細胞培養を実施するなどの別の実験的方策を取り入れることもできる。

【0008】

この追跡作業は、追加検討が必要と思われる手順が多数ある場合には特に、その実施に比較的費用と時間が掛かる。追加検討のための一部の手順の選択を誘導することができ、さらにこれらの手順の効率を先験的に予測できれば、データ解析に役立つであろう。実験は多くの場合、例えば多数の細胞単位を各手順に供するなど並行して行われることから、ある手順の再現実験をN回以上実施して最終検査で陽性であれば、その手順が追跡作業に特に適している可能性がある（個別環境に応じて、Nを1、2、3・・・のように選ぶことができる）。

【0009】

そのような実験の重要な目標の一つは、細胞の分化を特定の表現型に向くように制御できるすなわち方向付けできることである。例えば、幹細胞から開始し、培養中に細胞の特定の型、例えば赤血球、心筋細胞、または脳細胞を作り出すように望むこともできる。結果として得られる分化細胞はその後、広範な潜在的用途に供され、その用途には、生体系のモデル化および研究、薬品用の毒性ふるい分け、再生薬開発用のふるい分け、および、例えば脳卒中または脊髄損傷時のヒトへの細胞移植によって死滅細胞または異常細胞を置換することなどがある。細胞培養実験はまた、その他の幅広い用途にも有用となり得る。

【発明の概要】

【課題を解決するための手段】

【0010】

本発明は添付の特許請求の範囲によって規定される。

本発明の一実施形態によれば、細胞培養データを処理する方法が提供される。データは多数の試料からの結果を含み、結果は細胞培養の複数の段階を各試料について連続して実行することによって得られ、各段階は一組の特定条件をもつ細胞培養処理を表し、そのため各試料は細胞培養に適用される処理の固有性および順序によって規定される手順に従う。本方法は望ましい細胞培養の成果を生み出した試料のサブセットを規定する工程と、その試料のサブセットからの結果をコンピュータによって解析して結果を順序付けする工程または分類する工程とを含む。順序付けする工程または分類する工程は、望ましい細胞培養成果を得るのに効果的な一つ以上の手順を特定するのに役立ち、順序付けまたは分類を生み出すための解析は、異なる手順の間の類似性に関する情報を利用する。

10

20

30

40

50

【 0 0 1 1 】

望ましい細胞培養の成果は一つ以上の条件によって定義されてもよい（潜在的に代替条件を表すものでよい）。結果を順序付けする工程または分類する工程は一般に、サブセットの試料を順序付けする工程または分類する工程、またはサブセットの試料に関連する手順を順序付けする工程または分類する工程を含む。どちらの場合にも解析では、望ましい細胞培養の成果の入手に効果的でありかつ偽陽性を示す可能性がある試料の区別に効果的である最良の機会が一般的に存在する手順の特定が試みられる。その結果、追跡実験のより良い目標を立てることができ、それによって実験時間および費用を削減することができる。

【 0 0 1 2 】

既存の方法（異なる手順を互いに独立して取り扱う）とは異なり、この解析は異なる手順間の類似性に関する情報を利用する。次にこの方法は、より堅牢かつ正確な順序付け／分類（したがって追加検討の手順のより良い特定）を実現するために、異なる手順のデータを結合する仕組みを提供する。

【 0 0 1 3 】

一実施形態では、順序付けするまたは分類するための解析は、（異なる手順間の類似性に関する情報に加えて）各手順に従うサブセットの試料の数をさらに利用する。例えば、少なくとも一つの陽性結果を出した I 個の手順があり、 $I(n)$ が n 番目の手順で陽性結果を出した試料の数であるとすれば、所与の手順（ $P(i)$ ）の点数（ S ）は次式で定義される。

【 数 1 】

$$S(P(i)) = \sum_{k=1}^{k=I} I(k) \times WT(P(i), P(k))$$

ここに $WT(P(i), P(k))$ は、手順 $P(i)$ と手順 $P(k)$ との間の類似性に基づく重み係数である（類似性が高いほど重み係数は大きい）。この方法では、ある手順が陽性結果を出した他の手順（とりわけ複数の陽性結果が得られた他の手順）に似ているほど、その手順の点数は高くなる。したがって、各手順の点数 S を基にして結果を分類することができる。この方法は点数付け公式化の単なる一候補に過ぎず、当然のことながら当業者は他に多数の候補を思い付くであろう。

【 0 0 1 4 】

上記公式化の重み係数は、異なる手順間の隔たり量の形態と考えることができる（隔たりが大きければ類似性は低く、故に重み付けが低くなる）。異なる手順間の重み係数（すなわち隔たり測定値）を決定する一つの方法は、関心のある手順に共通の段階の数を数えることである。例えば、重み係数は、関心のある手順に共通の段階の数に比例する（または他の何らかの相応しい依存性をもつ）可能性がある。

【 0 0 1 5 】

上記の方法では、個々の処理の比較に二値基準（0 または 1）が与えられる（同じか違うかのどちらか）。しかしながら、いくつかの実施形態では、もっと段階的な基準を使用して個々の処理を比較してもよい。例えば、処理 A が第 1 の条件の組の使用を含み、処理 B が第 2 の条件の組の使用を含む場合、類似性評価では、処理 A と処理 B との間で共通な条件の数を考慮に入れてもよい。同様に、処理 A、B および C 全てに同じ薬品が含まれるが濃度が異なる（ $A > B > C$ ）場合、A は C より B により類似していると考えられる（濃度の点で近いため）。別の例として、似た経路を有効にする処理がそうではない処理よりもより類似していると思ってもよい。

【 0 0 1 6 】

一実施形態では、解析はサブセットの各試料の記録を含む一組のデータに関して実施される。各記録は、試料の識別子および試料に適用された手順に関する情報を含んでもよい。試料に適用された手順に関する情報は、試料に適用された処理の順序の一覧を含んでもよい。別の実施形態では、手順に関する記録の情報は、単に手順の標識または他の識別子

10

20

30

40

50

を含むだけでもよく、これを用いて個々のデータの組にアクセスし、所与の手順で用いられる処理の情報（順序および固有性）を提供することができる。

【 0 0 1 7 】

一実施形態では、順序の一覧は二進列で表される。二進列の各ビットは異なる段階の異なる処理に対応する。そのため、例えば5ラウンドあって各ラウンドに8つの処理候補がある場合、二進列は40ビットを含む。二進列内の各ビットの値は、その特定の段階の試料に所与の処理が適用されるか否かを示す。二進列をこのように使うことによって、異なる手順間で共通な処理の数を直接数えることになり、分類および順序付けのための種々の手順への入力の手順の簡便な一形態が提供される。

【 0 0 1 8 】

一実施形態では、分類および順序付けする工程は試料（または手順）を集団化する工程が含まれる。この集団化する工程は、種々の技術を用いて実行することが可能であり、その技術には階層的集団化、自己組織化マップなどがある。最も高密度に集団化している（最も狭い分類）領域（複数領域も含む）が、追加検討で最も関心のある手順を示す傾向があることが分かる。というのは、その手順が陽性結果を生み出した手順に類似した手順を示すからである。一方、低密度の集団化（緩い分類）は、陽性結果を生み出した手順ではあるが他の類似手順はほとんど（または全く）陽性結果を生み出していない手順を示している。一般に、集団の密度が高いほど、手順が偽陽性を示す集団に含まれる可能性は低くなる。

【 0 0 1 9 】

結果を解析する他の技術としては、集団化よりもむしろ順序付けを行うものでもよい。例えば一つの方法は、上記にならって各試料（手順）に点数を付けることである。その後その試料（手順）をこの点数に従って順序付けまたは等級付けすることができる。この方法では、高得点の試料（手順）が、点数が低い試料（手順）よりも、他の成功した試料または手順により類似している傾向がある。一般に、所与の試料（手順）の点数が高いほど、偽陽性を示す可能性は低い。

【 0 0 2 0 】

したがって解析の結果（集団化、順序付け、または何らかの他の適切な形式で表される）は、望ましい細胞培養の成果を作り出す可能性が比較的高い最も関心のある手順の特定に役立つ。故に一実施形態では、本方法は、追加検討用の細胞培養処理を特定するために、結果の分類および順序付けを用いる工程をさらに含む。本方法は、特定された細胞培養処理を追加検討する工程を含んでもよい。

【 0 0 2 1 】

一実施形態では、本方法は、望ましい細胞培養成果を生み出した試料の少なくともサブセットの測定値を解析して、サブセットの各試料の手順を決定する工程をさらに含むことができる。例えば、異なる蛍光標識を試料に授与するように異なる処理を配列してしてもよく、流動細胞測定または走査細胞測定によって測定を実施して試料に関連する蛍光標識を識別してもよい。試料の手順を測定値によって確実に決定できなければ、その試料の結果は廃棄される場合もあるため、試料の結果は分類または順序付けの解析には含まれない。場合によっては、（全体よりもむしろ）部分的な手順が確実に決定されることがある。例えば測定値から、他のラウンドではなくあるラウンドの処理が明示される場合がある。そのような部分的な結果でも、その後続いて行われる解析によっては役に立つ場合がある。

【 0 0 2 2 】

一実施形態では、望ましい細胞培養成果は一つ以上の試験に合格することによって決定される。本方法は結果の分類を解析する工程をさらに含んで、前記一つ以上の試験に合格する試料の別の分類を特定する。これらの別の分類は細胞培養の別の特性を表してもよく、例えば異なる細胞表現型に対応してもよい。したがって分類は、単一の手順または関心のある経路の特定に限定されず、別の成果（全体に望ましい成果の範囲にある成果）に結びつき得る別の手順を特定するために用いられてもよい。例えば、望ましい成果はA型の

10

20

30

40

50

細胞の場合であっても、共に A 型に対応する A 1 および A 2 の亜型が存在する場合がある。結果の 2 つの分類がそれぞれ 2 つの異なる亜型に対応する場合があることが分かっており、それによって、（データの組の少なくとも一部の）分類の生物学的意義が実証される。

【0023】

一実施形態では、本方法は、細胞培養の複数の段階を実施する工程を含んで処理のための結果を生成する。故にいくつかの実施形態は、元の細胞培養実験を実施する工程から、結果を解析する工程および結果の解析に基づく追跡実験を実施する工程まで全手順を網羅して、どの手順が実際に望ましい細胞培養成果をもたらすかを裏付ける。

【0024】

本発明の一実施形態は、上記方法のいずれかを実行するためのコンピュータプログラムを備える。コンピュータプログラムは、不揮発性半導体記憶装置、光ディスク（CD、DVD など）、コンピュータハードディスクなど、コンピュータによって読み取り可能な何らかの適切な媒体に保存してもよい。コンピュータプログラムをインターネットなどのネットワークから供給可能にしてもよい。

【0025】

本発明の他の実施形態では、細胞培養データを処理する装置が提供される。データは多数の試料の結果を含み、結果は細胞培養の複数の段階を各試料について連続して実行することによって得られ、各段階は一組の特定条件をもつ細胞培養処理を表し、そのため各試料は細胞培養に適用される処理の固有性および順序によって規定される手順に従う。本装置は、望ましい細胞培養の成果を生み出した試料のサブセットを規定するデータを含む記憶装置と、サブセットの試料の結果をコンピュータによって解析して結果の順序付けまたは分類を行うよう構成された演算素子とを備え、順序付けまたは分類は、望ましい細胞培養成果を得るのに効果的な一つ以上の手順を特定するのに役立ち、順序付けまたは分類のための解析は、異なる手順の間の類似性に関する情報を利用する。

【0026】

本装置は、適切な符号でプログラムされたコンピュータシステム（複数のシステム含む）によって実行されてもよい。符号には、コンピュータシステムと共に一つ以上の演算素子による実行用のプログラム命令が含まれる。符号は光ディスク、磁気テープなどの持続性媒体に保存されてもよい。いくつかの実施形態では、処理の一部または全てを、専用のまたは特殊用途のハードウェアを使用して実行してもよく、または適切にプログラムされた汎用コンピュータワークステーションを用いて実行してもよい。本装置は、細胞培養実験で使用される機械の一部でもよく、またはそれに組み込まれてもよい。例えば本装置はフローサイトメトリー（流動細胞計測）システムを備え、これを用いて、細胞培養実験から結果を生成するとともにその結果を分類および順序付けの方法によって解析してもよい。本装置は、本方法の実施形態に関する上記の具体的特徴と同じ利益をもたらす。

【図面の簡単な説明】

【0027】

以下の図面を参照して、実施例によって本発明の種々の実施形態をさらに詳細に説明する。

【図 1】本発明の一実施形態による細胞培養実験およびデータ処理の上位の流れ図である。

【図 1 A】本発明の一実施形態による、第 1 の実験用の、経路を決定するために用いられる解析における較正段階の画面図である。

【図 1 B】本発明の一実施形態による、第 1 の実験用の、一つの試料用の経路を決定する画面図である。

【図 1 C】本発明の一実施形態による、第 2 の実験用の、経路を決定するための解析における較正段階の画面図である。

【図 1 D】本発明の一実施形態による、第 2 の実験用の、一つの試料用の経路を決定する画面図である。

【図 2】本発明の一実施形態による、細胞培養実験のデータを処理する上位の流れ図である。

【図 3】本発明の一実施形態による、一組のデータ処理用の結果を生成する細胞培養実験の概略図である。

【図 4】図 3 の実験による順序付けされたデータを示す図である。

【図 4 A】図 4 の領域をより詳細に示す図である。

【図 5】階層的集団化を用いて分類された、図 3 の実験によるデータを示す図である。

【図 5 A】図 5 の線図の一部分を拡大して説明する図である。

【図 6 A】図 5 の線図のデータの一集団化配列を拡大して説明する図である。

【図 6 B】図 5 の線図のデータの他の集団化配列を拡大して説明する図である。

10

【図 7】(図 5 の集団化に関する変形形態として)階層的集団化を用いて分類された、図 3 の実験のデータを示す図である。

【図 8 A】経路 1 0 - 8 - 8 - 7 に対応し、顆粒球型細胞コロニーを示す図である。

【図 8 B】経路 1 0 - 8 - 8 - 7 に対応し、単球型細胞コロニーを示す図である。

【図 9 A】経路 3 - 8 - 9 - 7 に従うピーズに対応し、顆粒球型細胞コロニーを示す図である。

【図 9 B】経路 3 - 8 - 9 - 7 に従うピーズに対応し、顆粒球、赤血球、単球および巨核球 (G E M M) 混合型細胞コロニーを示す図である。

【図 1 0 A】経路 2 - 2 - 7 - 7 に従うピーズに対応し、図 1 0 の顆粒球型細胞コロニーを示す図である。

20

【図 1 0 B】経路 2 - 2 - 7 - 7 に従うピーズに対応し、図 1 0 の単球型細胞コロニーを示す図である。

【図 1 1 A】アイソタイプ試験および c d 1 1 b 染色細胞の流動細胞測定解析を表現する図である。

【図 1 1 B】経路 2 - 2 - 7 - 7 に対応し、c d 1 1 b 染色細胞の流動細胞測定解析を表現する図である。

【図 1 2 A】アイソタイプ試験および c d 4 5 染色細胞の流動細胞測定解析を表現する図である。

【図 1 2 B】経路 2 - 2 - 7 - 7 に対応し、c d 4 5 染色細胞の流動細胞測定解析を表現する図である。

30

【図 1 3 A】アイソタイプ試験および c d 1 1 b 染色細胞の流動細胞測定解析を表現する図である。

【図 1 3 B】経路 1 0 - 8 - 8 - 7 に対応し、c d 1 1 b 染色細胞の流動細胞測定解析を表現する図である。

【図 1 4 A】アイソタイプ試験および C D 4 5 染色細胞の流動細胞測定解析を表現する図である。

【図 1 4 B】経路 1 0 - 8 - 8 - 7 に対応し、C D 4 5 染色細胞の流動細胞測定解析を表現する図である。

【図 1 5 A】アイソタイプ試験および C D 4 5 染色細胞の流動細胞測定解析を表現する図である。

40

【図 1 5 B】経路 3 - 8 - 9 - 7 に対応し、C D 4 5 染色細胞の流動細胞測定解析を表現する図である。

【図 1 6】経路 1 0 - 1 - 8 - 5 に対応し、B - リンパ球型コロニーの 2 つの実施例を示す図である。

【図 1 7 A】アイソタイプ試験および C D 1 1 b 染色細胞の流動細胞測定解析を表現する図である。

【図 1 7 B】経路 1 0 - 1 - 8 - 5 に対応し、C D 1 1 b 染色細胞の流動細胞測定解析を表現する図である。

【図 1 8 A】アイソタイプ試験および C D 4 5 r / b 2 2 0 染色細胞の流動細胞測定解析を表現する図である。

50

【図18B】経路10-1-8-5に対応し、CD45r/b220染色細胞の流動細胞測定解析を表現する図である。

【図19A】アイソタイプ試験およびCD3e染色細胞の流動細胞測定解析を表現する図である。

【図19B】経路10-1-8-5に対応し、CD3e染色細胞の流動細胞測定解析を表現する図である。

【図20A】アイソタイプ試験およびCD49b染色細胞の流動細胞測定解析を表現する図である。

【図20B】経路10-1-8-5に対応し、CD49b染色細胞の流動細胞測定解析を表現する図である。

【図21】ウォード法を用いて階層的集団化によって分類された、図3の実験のデータを示す図である。

【発明を実施するための形態】

【0028】

図1は、本発明の一実施形態によるデータ処理を含む細胞培養実験の上位の流れ図である。本流れ図は、例えば数千個の試料などの多数の試料を含む体外細胞培養実験（操作100）を実行する工程に始まる。そのような細胞培養実験が、例えば特許文献1に説明されており、各試料は細胞培養の基質として作用するビーズを含む。そのビーズを一連の培地に浸漬することによって、各試料が一連の処理に供される。任意の所与の試料に行われるその一連の処理をその試料用の手順（経路）と考えることができる。

【0029】

一般に処理は複数の段階すなわちラウンドで行われ、全ての試料の処理のラウンドが同数になるように行われる。一部の試料が異なる数の処理を受ける場合でも、必要に応じてある試料のラウンドに無効な処理を「あてる」ことによって、処理のラウンドの数を設定された試料全体で均一化することができる。これによって、各試料は同じ決まった数の（所定数の）処理を受けるものと考えることができる。

【0030】

$N(i)$ が i 番目の段階すなわちラウンドにおける異なる処理候補の数であり、全体で I 個の段階があるとすると、実験の手順（ N ）の総数は $N = N(1) \times N(2) \cdots N(I)$ によって与えられる。一般に、複数の試料が（平均すると）それぞれの手順に供されるように、試料の数は N よりも大きくなるように選ばれる。

【0031】

特許文献1に説明されているように、各ラウンドの試料を分割する方法の候補が多数ある。一つの方法は、試料（例えばビーズ）の数を、ラウンド内の各処理に対して（ほぼ）均等に分けることである。その後各ラウンドの最後に、試料を次のラウンド用に再度分割するまで一緒に貯留する。試料の数が実験で試験される手順の総数より非常に多いと仮定すれば、統計的基礎に従って、この方法によって複数試料が確実に各手順に供される。

【0032】

貯留しその後各ラウンドの最後に分割するのではなく、もう一つの方法は分割した後に貯留することである。つまり、第1のラウンドの各処理後の試料を第2のラウンドの処理の数に分割する。その後、第2のラウンドの異なる処理のそれぞれに予定されている割り当てすなわち分割量を一緒にして処理を開始する。この方法は（統計的分配に頼るのではなく）手順全体にわたってさらに正確に試料を分配するが、分割および貯留の総量が非常に多いため、より実験全体を見通す能力に影響される。

【0033】

各処理は、細胞培養の物理的、化学的および/または生物学的な異なる条件を表す。例えば異なる処理は、異なる温度または照明条件、異なる成長培地の使用、特定のホルモンの有無等を含んでもよい。当業者は、多岐にわたる利用可能な種々の処理を思い付くだろう。さらなる考察には、特許文献1を参照されたい。

【0034】

10

20

30

40

50

各試料の手順は後の検出および解析のために記録される。この記録は、例えば所与の処理に供される全ての試料に、特定の蛍光標識を関連付けることによって物理的あるいは化学的に行ってもよい。その後、試料がたどったその手順を、その試料と関連した一組の標識から決定する。他の一方法は、例えば試料に R F I D 標識を含めることによって、各試料を識別できるように作ることである。それによって、特定の処理を受ける各試料の固有性を記録することができ、その結果、任意の所与の試料が受ける一連の処理の記録が得られる。試料の手順を監視し記録する種々の方法についてのさらなる情報は、特許文献 1 に見いだされる。

【 0 0 3 5 】

細胞培養実験の結果を精査して、陽性結果を生み出した試料を決定する（操作 1 1 0）。結果の精査には流動細胞測定または何らかの他の適切な技術を使用してもよい。本明細書で述べる「陽性」の結果とは望ましい成果の意味を含み、具体的な産物または効果があること（または無いこと）でもよい。また、この望ましい成果はさらに複雑な結果を表してもよい。例えばある成分が存在しかつ別成分が存在しないことでもよい。陽性試料を、細胞培養実験に供された試料の元の組のサブセットと考えることもできる。

【 0 0 3 6 】

ここで陽性試料の実験手順が決定される（操作 1 2 0）。この決定は種々の技術によって行ってもよく、例えば特許文献 1 を参照されたい。操作 1 1 0 および 1 2 0 の順序には自由度がある。例えば、一部の実験配列では、全ての試料の手順を決定してもよい。結果が陽性である試料のサブセットを一旦識別すれば、これが、対応する手順のサブセットに直接結び付く。他の実験では、結果が陽性である試料を最初に決定してもよい（図 1 に示す）。その後、陽性試料のこのサブセットについてのみ処理手順を特定する（これによって、判定が必要な手順の数を減らす）。他の実験では、陽性試料のサブセットの識別を処理手順と同時に実施してもよい。例えば、実験的成果と手順標識の両方を同時に検出することができる流動細胞測定または走査細胞測定の任意の適切な形態で識別してもよい。

【 0 0 3 7 】

出願人等は、第 1 のラウンドの処理に 1 A、1 B、1 C、・・・1 N、第 2 のラウンドの処理に 2 A、2 B、2 C、・・・2 N などの表示をすることができる。このとき、

（a）異なる処理の選択肢の数は、あるラウンドと他のラウンドでは異なる場合があり（すなわち「N」はラウンド間で異なる場合がある）、

（b）異なるラウンドの一組の処理の間になんらかの重複の程度（重複無し、部分的、完全）がある。例えば、第 1 のラウンドのある処理は、一つ以上後のラウンドの処理と同じである場合がある（例えば 1 B = 2 C = 3 C）。とりわけこれは、所与の処理に供する時間が重要か否かの検討（すなわち処理の順序付けの検討にも）に役立つ。

（c）各ラウンドは、一つ以上の試料がそのラウンド内の何らかの特定の処理に供されなかったことを示すために、「無効」処理を含む場合がある。

（d）手順候補（すなわち種々のラウンドの処理の有望な組合せ）が必ずしも全て実施されるわけではない。例えば、T 1 および T 2 が 2 つの処理であり、順序付けが重要であるか検討するのが望ましいとすれば、1 A = 2 A = T 1 および 1 B = 2 B = T 2 としてもよい。この場合、順序 1 A - 2 B および 1 B - 2 A は関心のある順序であるが、処理 T 1 または T 2 のどちらかの単純な繰返し（順序としては、1 A - 2 A および 1 B - 2 B）は実施されない場合がある。他の方法（例えば、貯留の後に分割する方法）の場合、陽性結果をもたらさない試料については追加解析（操作 1 2 0）を行わないことを考えれば全ての手順を実施するほうが簡単であるものの、実験手順によっては、後者の順序を全く実施しなくてもよい（特に、各ラウンドの最後において分割 - 貯留する方法では）。

【 0 0 3 8 】

したがって、実験段階の出力（すなわち操作 1 3 0 のデータ処理段階への入力）は一つ以上の成功した手順の組であり、各手順の成功は（例えば）標準分析によって測定され、各手順は手順を形成する一連の処理によって表される。例えば、4 ラウンドの処理がある場合、手順を上記の呼び方 1 C - 2 A - 3 C - 4 D を用いて表してもよい。複数の成功し

10

20

30

40

50

た試料が同じ手順をたどるとすれば、その手順はその一組のデータの中で複数回繰り返されることになる。

【 0 0 3 9 】

一般に大規模な細胞培養実験の目的、したがってデータ処理段階 1 3 0 の目的は、具体的な関心がある手順を特定することである。次に、これらの手順について追加の実験的検討（操作 1 4 0）を行う。その実施には比較的費用と時間が掛かる。したがって、特に実施する価値が最もあると思われるそれらの手順を順位付けする観点で、および偽陽性（すなわち偽陽性結果を生み出した試料／手順）を廃棄することができるという観点でも、できるだけ効果的であることが手順の特定には重要である。

【 0 0 4 0 】

図 1 に示した処理の中で、操作 1 0 0 および 1 4 0 は一般に実験室内で行われ、一般に操作 1 1 0 および 1 2 0 は実験測定値とコンピュータ解析とを一緒に含み、一般に操作 1 3 0 はコンピュータ解析を表す。実験測定値（および潜在的にコンピュータ解析）を細胞培養実験 1 0 0 自身に組み込んでもよく、または別の検討として後で実施してもよい。コンピュータ解析の実行には、専用ハードウェアを用いても良く、または記憶装置、演算素子などを備える従来のコンピュータを用いてもよい。演算素子が、コンピュータを制御するソフトウェアを実行して関連動作を実行する

【 0 0 4 1 】

一実施形態では、図 1 の操作 1 2 0 における経路の特定は、処理の各ラウンド（分割）の各貯留に蛍光標識を追加することによって実施される。標識付けは、（あるラウンドのみではなく）ラウンド全体の各処理に固有の標識を提供する。しかし処理の最終ラウンドには標識を使用しない。というのは、最終ラウンドの所与のビーズが供される特定の処理が、実験の最後のビーズを含む貯留（容器）から直接始まるためである。一般に蛍光標識は、試料例えば実験に使用されるビーズ（微小担体）よりも非常に小さく、処理のラウンドの貯留内のビーズに付着する。任意の所与の処理の間に任意の所与のビーズに付着する標識の数は、実験条件、ビーズの具体的な特性、不規則な変動等によって変動する。場合によっては、所与のビーズがわずかの標識しか（場合によっては標識を全く）拾わない場合がある一方、他の場合には、一回の処理で、所与のビーズが 1 0 0 以上の標識を拾う場合もある。

【 0 0 4 2 】

標識は、（蛍光の）色、標識の大きさおよび標識の蛍光強度など一つ以上の特性の組合せによって識別されてもよい。具体的な一実施形態では、（ 3、4、5 ）で表された 3 つの大きさ、（ R（赤）、B（青））で表された 2 つの色および（ 0 1、0 2、0 3、0 4、0 5、0 6、0 7、0 8、0 9、1 0、1 1、1 2 ）で表された 1 2 の蛍光強度が利用できる。したがって、所与の標識の表示は例えば 4 R 1 1 のように表され、大きさが 4、赤色および強度が 1 1 を示し、あるいは 3 B 0 3 のように表され、大きさが 3、青色および強度 0 3 を示す。これによって総数 $3 \times 2 \times 12 = 72$ 個の固有の識別子が得られる。

【表 1】

| | T1 | T2 | T3 | T4 | T5 | T6 | T7 | T8 | T9 | T10 |
|-----|------|------|------|------|------|------|------|------|------|------|
| 分割1 | 5R03 | 5R04 | 5R05 | 5R06 | 5R07 | 5R08 | 5R09 | 5R10 | 5R11 | 5R12 |
| 分割2 | 3R02 | 3R03 | 3R04 | 3R05 | 3R06 | 3R07 | 3R08 | 3R09 | 3R10 | 3R11 |
| 分割3 | 4R02 | 4R03 | 4R04 | 4R06 | 4R07 | 4R08 | 4R09 | 4R10 | 4R11 | 4R12 |

【表 2】

| | T1 | T2 | T3 | T4 | T5 | T6 | T7 | T8 | T9 | T10 | T11 | T12 | T13 | T14 | T15 |
|-----|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| 分割1 | 4R02 | 4R03 | 4R04 | 4R06 | 4R07 | 4R08 | 4R09 | 4R10 | 4R11 | 4R12 | 3R07 | 3R08 | 3R09 | 3R10 | 3R11 |
| 分割2 | 5R03 | 5R04 | 5R05 | 5R06 | 5R07 | 5R08 | 5R09 | 5R10 | 5R11 | 5R12 | 3R02 | 3R03 | 3R04 | 3R05 | 3R06 |

【0043】

表1および2は、2つの異なる実験に採用された標識付け配列を説明している。第1の表には4つのラウンドまたは分割があり、それぞれ10個の異なる処理を含み（表1）、第2の表には3つのラウンドまたは分割があり、それぞれ15個の異なる処理を含む（表2）。各実験に対し、最終ラウンドでは標識を付けていない（上記理由による）。各ラウンドのT1、T2・・・T15のような処理の表示は、全てのラウンドで同じ組の処理が使用されることを示すものではない。実際、これは一般には当てはまらない。言い換えれば、分割1の処理T5は分割2の処理T5と同じ表記であるものの、極一般的には異なる処理を意味する。同様に、分割3の処理T5は、分割1および分割2のT5とは異なる別の処理を表すことがある。

10

【0044】

一実施形態では、流動（または走査）細胞測定による測定値の代わりに、所与の試料に関連する標識がその試料から分離される。流動細胞測定により各標識について4つの測定値が得られる。一つが蛍光波長を、一つが蛍光強度を、2つがそれぞれ横方向および前方向の散乱強度を表す。これらの4つの測定値により、上記規定の3つの変数（大きさ、色および強度）を、したがって所与のビーズまたは試料の標識の固有性を決定することができる。

20

【0045】

図1Aは、上記表1の実験の標識のコンピュータで行う較正を説明する画面図である。この解析システムでは、特定の実験の結果を選択するのに実験列を用いる。流動細胞測定の所与のセッションを選択するのにセッション列を用いる。そのような測定値の異なるセッションは別々に較正される。その理由は、あるセッションから他のセッションへの出力時に、装置が若干変更される場合があるからである。

30

【0046】

各個々の試料から得られた標識の結果は一組のファイルに含まれ（各試料に一ファイル）、表題「データの組」の下に別々に羅列される。また、対照ファイルがあり（図1Aで選択されている）、セッションの開始時に測定される標識から得られたデータを含む。一実施形態では、対照ファイル用のこれらの標識は細胞培養実験で使用される標識と同じであるが、試料に付けられて実際に実験に供された標識ではない。他の実施形態では、実験に供された試料から取り外された標識を直接用いて較正することもできる。後者の場合、試料が陽性結果を生み出したか否かに関係なく、試料の標識を較正に使用できる（この方法がより多くのデータ点を較正に与えるためである）。

【0047】

図1Aの画面の右側部分には区画が3箇所ある。左下の区画には横方向散乱を表すX軸および前方散乱を表すY軸がある。3つの異なる大きさの標識がY軸に沿って分けられた集団ごとにはっきり見え、3つの独立した四角によって描かれている。Y軸の最小値の集団が大きさ3に、Y軸の中央値の集団が大きさ4に、Y軸の最大値の集団が大きさ5に対応する。

40

【0048】

図1Aの画面の右側部分の右下の区画には、データが3つの四角のそれぞれの中に左側部分からZ軸に沿って延び、赤色蛍光波長の強度値を表す。表1の実験は全て同じ色の標識を使用したため、4番目の軸（青色蛍光波長の強度値）に沿った区画は表示されていない。左下の区画内の3つの集団（四角）のそれぞれは右下の区画の複数の集団を含み、各

50

集団は異なる強度に対応することがわかる。これらの異なる強度も独立した四角によって描かれている。各四角に付属する数値はその四角に該当する標識の数を示す。

【 0 0 4 9 】

右上の区画には3列の四角があり、これらは右下の3列の四角に対応している（各列は左下の区画の四角の一つに対応する）。右上の区画の四角は、右下の区画の各四角内のZ値に対応する標識の数の分布を示すヒストグラムである。

【 0 0 5 0 】

一実施形態では、流動細胞測定セッションの一組のデータの解析には、対照ファイルのデータの分散図（左下に示す）を描画する第一の工程を含む。その後、データの集団（すなわち図1Aの左下に示した3つの集団）は四角によって区分される（使用者の手作業またはコンピュータ解析のどちらかによって）。左下に示した各四角内で対応する集団用の対応するはしご型が決定され、これが右下に示されている（同じく使用者の手作業またはコンピュータ解析のどちらかによって）。はしご型の各区分は異なる標識強度（すなわちその標識の大きさ）に対応する。

【 0 0 5 1 】

図1Bは図1Aに類似した構成の標識を示すが、全てのデータが単一の試料に由来している点が異なる。この場合、四角ノはしご形の位置は図2Aに設定された全データの（再計算ではなく）解析によって管理される。平均すると試料は各処理あたり（非常に大まかに）100個の標識を拾い、明らかに各四角内の標識測定値を集団化する工程があることがわかる。このようにして、大きさ3に関しては、標識の大部分は、表1の標識3R06に対応するはしご型の5番目の区分に集団化される。大きさ4に関しては、同様に標識の大部分は、表1の標識4R07に対応するはしご型の5番目の区分に集団化される。大きさ5に関しては、標識の大部分は、表1の標識5R11に対応するはしご型の9番目の区分に集団化される。

【 0 0 5 2 】

図1Cおよび1Dは概ね図1Aおよび1Bにそれぞれ類似しているが、表1ではなく表2の実験に対応する。すなわち図1Cは対照ファイル内の一揃いの試料の結果を表す一方、図1Dは単一の試料の結果を表す。図1Dの試料が、使用された一番小さい強度の標識（4R02）および二番目に大きな強度5の標識（5R11）に対応する処理をたどったことがはっきりわかる。

【 0 0 5 3 】

図1Bおよび1Dの区画は関心がある試料の関連性がある経路をはっきり特定するものの、場合によっては、必ずしも標識の測定値が明確な特定を裏付けるわけではない。これは、試料が所与の処理の標識を十分に拾っていない、または次の（またはもっと後の）処理の中で標識を失った可能性があるなど、いろいろな原因によって生じる。また、第1の試料および第2の試料が処理の後ラウンドの同じ分割に入った場合、第1の試料は第2の試料から伝達された標識によって汚染された可能性がある（このとき、伝達標識は第1の試料ではなく第2の試料が供されたもっと前のラウンドの処理に対応する）。さらなる可能性としては、標識を光学測定する際の問題の存在である。

【 0 0 5 4 】

さらに、特定の試料の標識を一つ以上の処理に正確に割り付けることができるとしても、この割り付けは利用可能な経路に対応しなければならない。具体的には、標識識別は必ず各分割の一つの処理につながなければならない。試料の特定の分割の中で処理が特定されなければ、結果的に試料の経路特定は不完全となる。一方、試料の単一の分割内で複数の処理が特定されたら、それは何らかの誤りを示しており（例えば、2つのピーズが特定の処理の間にくっついてしまっている）、試料の完全な経路を決定できない。この場合、最終または最終の1つ前の分割における処理の情報によって部分的な経路を決定してもよい。

【 0 0 5 5 】

従って、図1の操作120（経路を決定する）は、経路を正しく決定できなかった試料

10

20

30

40

50

を廃棄または棄却する工程を含んでもよい（そのような試料が陽性結果を生み出しているとしても）。所与の経路の確実性の決定に、種々の統計的試験を用いることができることがわかる。それ故、所与の実験によって作られた（確実な経路の）陽性試料の総数は、個々の統計的試験および経路の決定の許容／棄却（および上記の部分的な経路を用いるか否か）に採用される関連閾値によって若干変動することがある。

【0056】

上記の様に、大規模細胞培養実験のデータを解析するひとつの方法は、特定の手順をたどった成功した試料（細胞単位）の数を数えることである。実際に、その特定の手順をたどった試料の数に従って手順の順位が付けられる。この方法では、各手順を他の手順とは独立に扱って統計値（その手順に関連する試料の数）を決定し、次にこれを用いて追加検討の手順を順位付け／選択する。その一方、本明細書に記載の細胞培養実験データを処理する方法（操作130）は、種々の手順の間の隔たりの程度に基づいて手順を分類および順序付けするなど、手順間の依存性または関連性を探す。この方法は、関心のある手順の潜在的価値を見分ける能力が高いことが確認されている。

【0057】

図2は、本発明の一実施形態による図1のデータ解析操作130の処理を説明する上位の流れ図である。この処理は、特定された手順（経路）の表示を二値表現に変換することから始まる（操作300）。具体的には、（例えば）第1のラウンドに6つの処理候補があるとすると、このラウンドは、ある桁は1、残りの桁はゼロである6桁の二進列で表される。二進列内の「1」の位置は対応する処理を示す。例えば、1A = 100000、1B = 010000、1C = 001000、1D = 000100、1E = 000010および1F = 000001とすることができる。後続のラウンドの処理も同様に二進列で表すことができる。その結果、異なるラウンドのこれら二進列を連結することによって、またはベクトルを形成して、ベクトルの要素の数がラウンドの数に対応し各要素が関連ラウンドの二進列を含むようにすることによって処理全体を示すことができる。

【0058】

この処理では、ここで手順の間の類似性（隔たり）が決定される（操作310）。この決定は種々の方法で行うことができる。（例えば）任意の2つの手順に対する一つの方法は、（a）処理の各ラウンドに対し、そのラウンドに対応する2つの二値符号にAND操作を行い、（b）処理の全てのラウンドについて（a）のゼロではない結果の数を合計することである。この処理の結果は2つの手順の間で処理が共通の（重複する）ラウンドの数を表す。この重複は、ゼロ（共通ラウンドは無い）、部分的（一部のラウンドが共通だが全てではない）、または完全（全てが共通のラウンド）となる可能性があり、2つの手順の間の隔たりの程度の一形態と考えることができる。

【0059】

この別の見方は、所与のラウンドの二値符号を、N次元空間への種々の手順の配置と考えることである（Nはラウンド内の処理選択枝の数）。その結果、この空間における所与のラウンドの処理の間の隔たりの値がゼロ（一致の場合）および1（不一致の場合）に数値化される。

【0060】

この処理ではここで、決定された隔たりすなわち類似性の程度に基づいて分類および順序付けが行われる（操作320）。各手順の単一桁の数（その手順をたどった試料の数）を他の全ての手順と独立に決定し、その数に基づいて順序付け／順位付けを行う既存の方法とは異なり、図2の方法は手順の間の依存関係を含む。具体的には、隔たりの程度は手順の別対ごとに計算され、この隔たりの程度を用いて分類および順序付けを決定する。

【0061】

隔たりの程度を用いて手順を分類および順序付けするには、公知の種々のアルゴリズムがある。これには、階層的集団化、自己組織化マッピングおよび指紋解析法がある。これらのアルゴリズムは、手順の間にある関連性、例えば何らかの形態の隔たりまたは類似性に着目して、試料／手順の分類、順序付け、またはその他の組織化を行う。手順の間のこ

10

20

30

40

50

の関連性を使用すると、別次元の情報が細胞培養実験の結果から抽出され、結果をさらに強力に解析することになる。そしてこれによって、追加検討すべきこれらの手順を、感度良かつ効果的に区別（操作 3 3 0）できるようになる（図 1 の操作 1 4 0 のように）。

【 0 0 6 2 】

当業者には、図 2 に示した種々の操作は、具体的な実装形態によって変更可能であることは自明である。例えば、操作 3 0 0 における手順の表示の二値符号への変換によって、多次元空間内マッピングが提供され、操作 3 1 0 でこれを用いて手順の間の隔たりが決定される。しかし、分類 / 順序付けに用いられる類似性または隔たりを、他の何らかのしくみによって得てもよい（多次元空間内の位置を参照せずに）。一つの可能性は、1 B と 1 D など、一対の処理表示を用いて参照表の隔たり数値を閲覧し、この隔たりを分類 / 順序付けに投入することである。この環境下では操作 3 0 0 の二値への変換は不要となる。別の可能性は、各経路を、その経路に関連する処理の情報を含む表に関連付けられた固有の表示で表わすことである。この経路表示を用いてその表から対応する経路情報を抽出し、必要に応じて類似性 / 隔たりを計算する。

【 0 0 6 3 】

いくつかの実施形態では、手順の種々の処理の二値符号は分類 / 順序付けアルゴリズムに直接投入され、何ら隔たりを明確に計算しない（すなわち操作 3 1 0 を無関連の工程として除外する）。この方法では、隔たりまたは何らかの対応する関連性の程度は、分類 / 順序付けアルゴリズムの一部として明確に決定される。当業者なら図 2 に示した処理をさらに変更できることは分かっている。

【 0 0 6 4 】

ここで、図 1 および 2 のデータ解析操作 1 3 0 の実施例を、図 3 に説明した実験のデータと関連付けて説明する。この実験は 3 0 0, 0 0 0 個の試料（生体高分子ビーズ）を含み、多能性幹細胞から開始した。各試料を、それぞれ 1 0 個の異なる処理候補を含む 4 つのラウンドの処理に供した。試料を処理の第 1、第 2 および第 3 の各ラウンドの後に貯留した後（無作為かつ均等に）再分割した。したがって、実験には 1 0, 0 0 0 (1 0⁴) 個の手順候補があり、そのため各手順を 3 0 個（平均）のビーズがたどったはずである。

【 0 0 6 5 】

処理の 4 番目（最終）のラウンドの後、ビーズを走査して 2 つの異なる陽性の成果を探した。すなわち（ a ）（標的大腸菌細胞を摂取する）食細胞の存在および（ b ）緑色蛍光神経系細胞の存在である。実験は（ a ）について陽性の結果がある 1 0 1 個のビーズ（試料）を算出し、（ b ）について陽性結果がある 8 4 個のビーズを算出した（これらに対し、各ビーズ（試料）の全手順を正確に決定した）。次に食細胞の結果のデータ処理をさらに詳しく説明する。

【 0 0 6 6 】

細胞培養実験の一組のデータを、ビーズごと別々の記録にして A S C I I ファイル書式にした。各記録は、ビーズ識別子を標識で分けた一覧および関連する 4 0 個の二値記述子を含む（上記二値符号化に対応する）。次にこの一組のデータを集団解析にかけて、類似性に基づいてビーズを分類した。上記のように、試料は N 次元空間内の位置またはベクトルと考えられ、ここで N は試料ごとの記述子の数に等しい（4 つのラウンドのそれぞれに 1 0 個の異なる処理があるため、ここでは 4 0）。次に類似性が、ユークリッド距離または何らかの他の適切な量に基づいて計算される（個々の応用例によって対称または非対称の場合がある）。

【 0 0 6 7 】

図 4 は上記の実験の結果を説明する区画である（図 4 の左上部をさらに詳細に説明した図 4 A が付属されている）。x および y 軸は、食細胞の存在について陽性の結果を出した各試料 / ビーズの順序の一覧を表す（各軸上に同じ順序付けがある）。区画の各四角は関心のあるビーズの対に対し共通な段階の数を表す。具体的には、黒い四角は 4 つの段階が共通、暗色の四角は 3 つの段階が共通、中間の濃さの四角は 2 つの段階が共通、明るい四角は 1 つの段階が共通、白い四角は共通の段階が無いことを表す。ビーズの固有性が図 4

10

20

30

40

50

Aにx軸およびy軸に沿って表示されている。これらの表示によって各ビーズを一意に特定することができる。(図4Aで用いられている正確な表示方法の論理的根拠は本発明の理解とは無関係)。左上から右下への対角線は黒いが、これは各ビーズの自分自身との交点であるからである(したがって必然的に共通の4つの段階がある)。また、共通の段階の数の決定は可換演算であるため、区画はこの対角線に対して対称である。

【0068】

ビーズは次のように軸上に並べられる。各ビーズはベクトルをもつと考えられ(N4、N3、N2、N1、N0)、ここにN4はビーズが4つの段階全てを共有する他のビーズの数、N3はビーズが3つの段階を共有する他のビーズの数、N2はビーズが2つの段階を共有する他のビーズの数、以下同じである。合計T個の(成功した)ビーズがあると仮定すると、 $N4 + N3 + N2 + N1 + N0 = T$ である(各ビーズの自分自身との一致をN4の値に含む)。次に任意の2つのビーズを、そのN4の値に従って互いに順序付ける。その2つのビーズが同じN4の値をもつ場合は、そのN3の値に従って互いに順序付ける。同じN3およびN4の値をもつ場合は、N2の値に従って互いに順序付ける、というように続ける。

【0069】

この方法では、他のビーズと共通性(関連性または類似性)が高いビーズが一般に区画の左上に集められる。上記の通り、ビーズの間の共通の段階の数はビーズの間の(より正確には対応する実験手順の間の)類似性(隔たり)の程度と考えることができる。これは、共通の段階の数値が大きい(類似性が高い)ことがビーズの間の隔たりが小さいことを表す一方、共通の段階の数が小さい(類似性が低い)ことがビーズの間の隔たりが大きいことを表すという反比例の関係であることが分かる。図4Aは成功したビーズのサブセットだけを表示している(他のビーズとの共通性が低い一部のビーズは省略されている)。

【0070】

上記の様に、関心のある手順を特定する一般的な従来の方法は、複数の成功したビーズが同じ手順をたどった状態を探すことである。図4のデータによれば、成功した手順が4つあり、これを2つ以上のビーズがたどっている。これらが図4Aに具体的に示されており、群A(ビーズ26、25および13)、群B(ビーズ28および24)、群C(ビーズ20および12)および群D(ビーズ98および101)である。

【0071】

従来の計数方法では、群Aはほとんどのビーズ(3個)を含むことから最も有望な群と考えられる。しかし、残り3つの群はそれぞれ2個のビーズを含み、これらを区別するのは難しい。しかし図4Aの区画から、一方の群A、群B、群Cと、他方の群Dと間の大きな違いがわかる。図4Aの左上隅のやや暗色の影の部分は、群A、群B、群Cが全てまとめて集団化され、互に大きく重なっていることを説明している。これによって、望ましい最終結果を生み出す良い手順をこれらの群に含まれる手順の中でまたはその近くで見付ける可能性を非常に高くする。

【0072】

その一方、群Dの黒い四角は比較的離れており、他の成功したビーズ具体的には群A、群B、群Cとの重なり(関連性)が非常にわずかである。このことが図4Aから(実質的に)白い行列よってはっきりわかり、そこには群Dの小さな黒い四角が配置されている。これは、群Dの手順が特定条件の影響を非常に受けやすく、本質的にあまり魅力がない手順となり得ることを示している(信頼性のある基盤で実施するのが難しい場合があるため)。別の可能性は、群Dが単なる偶発事象あって実際に有効な手順を表さないということである。そのような偶発事象の確率を、100個の成功したビーズが10000個の手順候補全体に無作為に配置されることを仮定して、所与の手順がそのような環境下で2つ以上のビーズによって共有される可能性を計算することによって推定できる。これは、 $1 - (9999! / 9900! * 10000^{99}) = 0.39$ で得られ、そのため実際には、陽性の結果を誤って分類する可能性は決して低くない。

【0073】

(理論的な、またはシミュレーションによる)確率計算を用いて任意の所与の結果の統計的有意性を評価できることは自明である。例えば、純粋に無作為に少なくとも3重項(3つの試料がそれぞれ同じ手順を共有する)を得る確率は次式で与えられる。

【数2】

$$1 - \frac{\sum_{d=0}^{I/2} \binom{N}{d} \binom{N-d}{I-2d} \frac{I!}{(2!)^d}}{N^I}$$

ここに、Nは経路候補の総数、Iは陽性結果の数、項数dは二重項(2つの試料が共に同じ手順を共有する)の数を示す(そのため、Iが奇数の場合は、項数は(I-1)/2で終わる)。したがってこの情報を用いて(i)潜在的偽陽性の識別を助け、(ii)統計的信頼性を高めるためにピーズの数などの最初の実験変数を設定することができる。

【0074】

例えばN=10,000かつI=101(図4の一組のデータの場合)に対し、三重項を偶発事象として得る可能性は上記の式から約0.0016となることがわかる。したがって、二重項(2つの陽性試料が同じ手順を共有する)は偶発する可能性もあるが(上記決定された0.39の確率によってわかるように)三重項との関連では可能性は低い。

【0075】

図4および4Aの試料の順序付けは、群Dと群A~Cの区画が互いに無関係であることを強く示唆するとしても、群Dと群A~Cを明確に分離しない。この分離を、軸上の試料を順序付けする他の方法でもっと明確にしてもよい。例えばその方法の一候補は、各ピーズに、他のピーズとの共通の段階の総数を加算することによって得られる点数を与えることである。例えば、101個のピーズがあり、特定のピーズが2つのピーズ(自身を含む)と4つの段階を共有し、6つのピーズと3つの段階を共有し、11個のピーズと2つの段階を共有し、23個のピーズと1つの段階を共有し、残りのピーズとは共有する段階は無いとすると、(2×4)+(6×3)+(11×2)+(23×1)=71の点数が与えられる。したがってこの点数は、所与のピーズの他の全てのピーズからの全隔たり(平均または総計)の何らかの形態を表す。次にこの点数に従ってピーズを軸に沿って順序付けることができる。この方法では、群A、B、Cの点数が、群Dの点数よりも大きくなり、したがって群A、B、Cは同じ集団に留まる一方、群Dは分離されることになる。これは、隔たりに基づいて順序付けする技術の単なる一説明であって、当業者はそのような順序付けを実行する他の何らかの可能な方法を思い付くことができる。

【0076】

図5は、(図4の区画ではなく)図3の細胞培養実験の結果を処理する別の仕組み、すなわち階層的集団化を説明している。図5の各列は、先頭にピーズIDが付けられた行に示された成功したピーズに対応する。図5のピーズに適用された表示は、図4に示された表示と同じである(これによって、これら2つの異なる分類の仕組みによる結果を、互いに直接比較することができる)。

【0077】

また図5には行ブロックが4つ描かれており、各行ブロックは、行ブロック見出しによって示す様に処理の一つのラウンドを表す。各行ブロックはさらに10行に分割され、各行はその特定のラウンドの一つの処理を表す。各列(ピーズ)には、各行ブロックの一つの暗色の四角があり、その特定のラウンドのそのピーズが受けた処理を示す。

【0078】

図5の左端は系統樹を示し、そのピーズの階層的集団を説明している(これは、ピーズが掲載される順序を決定する)。系統樹は一種の家系図と考えることができ、どのピーズがそれらの各手順の間の隔たり(関連性)の観点で他のピーズに近いかを示す。2つのピーズをつなぐ系統樹内をさらに左へ進むほど、すなわち上位階層へ進むほど、それらのピーズはさらに離れると考えられる。

【0079】

図5の階層的集団化は集団化3.0 (Cluster 3.0) プログラムを用いて実行され、系統樹は、ジャバツリービュー (Java Treeview) プログラムを用いて描写された。集団化プログラムは非特許文献3 (Eisen等、1998年) に記載されており、非特許文献4または非特許文献5からダウンロードできる (ジャバツリービュープログラムも後者のサイトから利用できる)。

【0080】

階層的集団化は凝集型 (上昇型) でもまたは対立型 (下降型) でもよい。前者の方法によれば、全ての対象すなわち試料は最初にそれらの個々の集団で表され、次にそれらが一つに統合される。ある凝集型アルゴリズムでは、最短の点間距離によって隔てられた対が第1の集団を形成する。次の集団を、点間距離が最短の2つの対象の間で再び形成し、対象を個々の試料または先に生成された集団のどちらかによって表すことができる。この手順を続けて、系統樹を全データに拡げる。

【0081】

階層的集団化アルゴリズムには、一つ以上の対象が集団である場合 (個々の試料ではなく集団) の点間距離を決定する仕組みが含まれる。これを行う仕組みは一般に連鎖と呼ばれ、集団の要素の間の平均差、集団要素の間の最大 (または最小) 距離などの種々の基準を基にすることができる。連鎖方法の選択、ならびに隔たりまたは類似性測定法の選択、および入力データの最初の順序付けが、集団化解析の出力に影響する可能性がある。

【0082】

図5は、1個以上のピースが所与の手順を共用する4つの分類を示している。これらの分類は図4に示したものと同一である (同じ表記で表示されている)。群Dが群A、B、Cと離れていることは系統樹から容易にわかり、図4と同じ結論である。上記のように、群Dは偽陽性を示すことがこの構成から強く示唆される (2つのピースが同じ手順を共有してはいるが)。

【0083】

図5の集団化の程度を説明する一つの方法は、「a C b」命名法の使用である。ここに「a」はピースの数、「b」は共通の段階の手順の数である。この命名法を用いると、群Aを3 C 4と表記することができ (3つピースが4つの段階全てを共有する)、一方群B、C、Dを2 C 4と表記することができる (2つのピースが4つの段階全てを共有する)。

【0084】

図5Aは図5の部分拡大図であり、群A、BおよびCを含む。4つの行ブロック内の処理を見ると、この部分のピースの間の処理に関して、変動の多くが、3番目および4番目のラウンドで発生していることが容易にわかる。その一方、この集団の大部分のピースが第1のラウンドの処理1~10および第2のラウンドの処理2~1を使用しており、これらの処理が陽性結果を得ることに重要な貢献をしていることを強く示唆している。

【0085】

図6Aは図5Aのデータの3つの集団を示し、集団の間に存在するいくつかのピースは省略されている。図5の群A、BおよびCの位置は図6Aにも示されている。図6Aの集団は5 C 3、4 C 3および6 C 3と表記され、それらが5個、4個および6個のピースをそれぞれ含み、3つの処理を共有していることを示している。3つの場合全てにおいて、処理の最終ラウンドで、処理の唯一の変動が発生する。

【0086】

図6Bは図5Aのデータの別の集団を示す。この集団は群A、BおよびCの7個のピースを含み、図5および5Aに示した順序付けからいくつかの中間ピースを省略することによって形成されている。図6Bの集団は7 C 3と記載することができる。すなわち7個のピースが3つの処理を共有している。これらの7個のピースでは、変動は処理の3番目のラウンドで発生する (図6Aの処理の4番目のラウンドではない)。

【0087】

図 6 A および 6 B は、例えば図 3 の一組の実験データの集団化を、集団階層の異なる深さで見るときの種々の潜在的な方法があることを説明している。この中で、図 6 A および図 6 B の構成は、同一の基礎となる一組のデータの補助的な、ビーズ / 手順の間の関係の別の特徴を強調する表現と考えることができる。この見方を組み合わせることによって、どの手順をさらに検討するかを選択するために、一組のデータ全体の理解を深めることができる。

【 0 0 8 8 】

同様に、図 4 の順序付けされた区画および図 5 および 5 A の系統樹も、同一データの補助的な表現と考えることができる。というのは、両者は共に特性の異なる特徴をわずかに測定するものの、同じ基礎となる特性（関連性）を示すからである。このことは、図 4 の最上位の試料が図 5 A の系統樹の最下位の試料（すなわち最も狭く集団化されている）にも対応するという事実から、さらに実証される。これによって、試料のこの小さなサブセットの重要性を再確認することになる。

【 0 0 8 9 】

図 4 ~ 6 の結果は、その処理がそのラウンドで実施されたか否かを示すために、ラウンド内の各処理を二値の桁で表すことによって生成された。例えば、A、B、C、D の 4 つの処理候補があるとすると、それぞれ (1 0 0 0)、(0 1 0 0)、(0 0 1 0) および (0 0 0 1) で表わすことができる。しかし、別の実施形態では、種々の処理の表記に別のさらに詳細な表記を使用してもよい。例えば、各処理が条件 X 1、X 2、X 3、Y 1、Y 2、Z 1 および Z 2 の組全体から選択したある条件を含んでもよい。A が (X 1、X 2、Y 2) を含み、B が (X 1、X 2、Y 1、Z 1) を含む場合、処理の各条件候補に二値の桁を割り当てることによって、これら 2 つの処理をそれぞれ (1 1 0 0 1 0 0) および (1 1 0 1 0 1 0) と表示することができる。これらの表示は処理の条件を規定し、これを集団化アルゴリズムへの入力として（または試料の順序付けまたは分類の他の形態として）用いることができる。

【 0 0 9 0 】

図 7 は、集団化 3.0 ソフトウェアを図 4 ~ 6 と同じ一組のデータに用いて実行した階層的集団化の結果を説明している。図 7 の構成は図 5 に類似しており、各列は特定のラウンド / 処理（右端に示した）に対応する一方、系統樹（集団）内の各終点は陽性ビーズに対応する。（ビーズ識別子が図 7 の下部全体に記載されているが、用いられた識別子は図 5 で用いられたものとはやや異なる）。図 7 の集団化の結果は図 5 に示した集団化と類似している（が同じではない）。この変動は、集団化が行われた方法の違い、例えば入力データの順序付けの違い、集団に対する隔たりを決定する規定の違いなどにあると考えられる。

【 0 0 9 1 】

図 7 は 2 つの特徴ある集団を表している。図 7 で集団 A と表記された第 1 の集団は、経路 A X - B X - C 8 / C 9 - D 7 に対応する。図 7 で集団 B と表記された第 2 の集団は、経路 A 1 0 - B 1 - C X - D 1 / D 5 / D 6 に対応する。この命名方では、集団 B は分割 A の処理 1 0、分割 B の処理 1、分割 C の任意の処理および分割 D の処理 1、5 または 6 に対応する。また、集団 B は一般に図 6 B に示した集団に対応するが、分割 D の他の処理用に一部追加の試料を含む。

【 0 0 9 2 】

本実験で陽性結果がある試料を特定するために用いられる生物学的ふるい（図 1 の操作 1 1 0 に対応する）は、食作用を目的とする大腸菌吸収法（internalization assay）であった。多くの細胞表現型が食作用特性を示すことから、このふるいは単一の表現型に特化したものではない。上記集団 A および B に対応する経路をさらに調査したときに得られた分化細胞の表現型検査から、2 つの集団が細胞を異なる表現型に分化させることが実証された。

【 0 0 9 3 】

集団 A では、この集団の経路は造血前駆細胞を 9 日で作ることがわかった。これらの前

10

20

30

40

50

駆細胞は、図 8、9 および 10 に説明したように、半個体培地内のコロニー形成法で単球、顆粒球および赤血球コロニーを生じさせた。具体的には、図 8 は経路 10 - 8 - 8 - 7 をたどったビーズに対応し、顆粒球型細胞コロニー（図 8 A）および単球型細胞コロニー（図 8 B）を示す。図 9 は経路 3 - 8 - 9 - 7 をたどったビーズに対応し、顆粒球型細胞コロニー（図 9 A）、および顆粒球、赤血球、単球および巨核球（GEMM）混合型細胞コロニー（図 9 B）を示す。図 10 は経路 2 - 2 - 7 - 7 をたどったビーズに対応し、顆粒球型細胞コロニー（図 10 A）および単球型細胞コロニー（図 10 B）を示す。

【0094】

集団 A に含まれる経路で生成された細胞が半個体培地から分離され、図 11 ~ 15 に示すように、汎白血球マーカー c d 4 5 および骨髓細胞系列 c d 1 1 b で有意に染色された。具体的には、これらの図は次のことを示す。

10

【0095】

図 11 は c d 1 1 b 染色細胞の流動細胞測定解析であり、図 11 A がアイソタイプ試験を表し（母集団の 0 . 3 % が陽性）、図 11 B が経路 2 - 2 - 7 - 7 に対応する（母集団の 9 . 7 % が陽性）。

【0096】

図 12 は C D 4 5 染色細胞の流動細胞測定解析であり、図 12 A がアイソタイプ試験を表し（母集団の 0 . 9 % が陽性）、図 12 B が経路 2 - 2 - 7 - 7 に対応する（母集団の 1 5 . 6 % が陽性）。

20

【0097】

図 13 は c d 1 1 b 染色細胞の流動細胞測定解析であり、図 13 A がアイソタイプ試験を表し（母集団の 1 . 8 % が陽性）、図 13 B が経路 10 - 8 - 8 - 7 に対応する（母集団の 5 . 5 % が陽性）。

【0098】

図 14 は C D 4 5 染色細胞の流動細胞測定解析であり、図 14 A がアイソタイプ試験を表し（母集団の 1 . 7 % が陽性）、図 12 B が経路 10 - 8 - 8 - 7 に対応する（母集団の 8 . 7 % が陽性）。

【0099】

図 15 は C D 4 5 染色細胞の流動細胞測定解析であり、図 15 A がアイソタイプ試験を表し（母集団の 2 . 6 % が陽性）、図 15 B が経路 3 - 8 - 9 - 7 に対応する（母集団の 2 8 . 8 % が陽性）。

30

【0100】

集団 B では、この集団の経路は単球、顆粒球および赤血球コロニー半個体培地は成長しなかったが、その代わりに、I I - 7 サイトカインを含み特別に調合された半個体培地の中で、B リンパ球型コロニーが成長した。これを図 16 で経路 10 - 1 - 8 - 5 について説明しており、B リンパ球型コロニーの 2 つの実施例を示している。

【0101】

この経路（10 - 1 - 8 - 5）で作られた細胞は、図 17 ~ 20 に示したように、骨髓マーカー C D 1 1 b に陰性、リンパマーカー C D 4 5 R / B 2 2 0、C D 3 e および C D 4 9 b に陽性であった。具体的には、これらの図は次のことを示す。

40

【0102】

図 17 は C D 1 1 b 染色細胞の流動細胞測定解析であり、図 17 A がアイソタイプ試験を表わし（母集団の 0 . 6 % が陽性）、図 17 B が経路 10 - 1 - 8 - 5 に対応する（母集団の 1 . 6 % が陽性）。

【0103】

図 18 は C D 4 5 r / B 2 2 0 染色細胞の流動細胞測定解析であり、図 18 A がアイソタイプ試験を表わし（母集団の 0 . 3 % が陽性）、図 18 B が経路 10 - 1 - 8 - 5 に対応する（母集団の 2 1 . 7 % が陽性）。

【0104】

図 19 は C D 3 e 染色細胞の流動細胞測定解析であり、図 19 A がアイソタイプ試験を

50

表わし（母集団の 0.2% が陽性）、図 19B が経路 10 - 1 - 8 - 5 に対応する（母集団の 41.6% が陽性）。

【0105】

図 20 は CD49b 染色細胞の流動細胞測定解析であり、図 20A がアイソタイプ試験を表わし（母集団の 1.0% が陽性）、図 20B が経路 10 - 1 - 8 - 5 に対応する（母集団の 9.9% が陽性）。

【0106】

したがって、集団 A は造血幹細胞すなわち骨髓前駆細胞からの一つの生物学的経路に対応し、一方集団 B は別の生物学的経路すなわちリンパ前駆細胞に対応することがわかる。したがって、本明細書に記載したように結果を分類することは、実験から陽性結果を識別するのを改良する助けとなるだけでなく、細胞培養実験において陽性結果の種々の型を識別し区別する助けにもなる。具体的には、結果をこのように分類または集団化することは、任意の所与の経路について陽性の成果の数を単に数えるよりも一層意味のあるかつ役に立つ方法で、生物学的に重要な種々の陽性成果を示す可能性がある。

【0107】

図 21 は、図 3 の実験データの階層的集団化を用いた別の表現である（図 5 および 7 と同じ）。図 21 の集団化は、結果を集団に組み込むためにワード法（ワード連鎖とも呼ばれる）を用いて行われた。この技術は、対象を集団化して変動を最小化することに基づいており、変動を、集団化のための隔たりの程度と考えることができる。詳細については、（例えば）非特許文献 6 を参照されたい。

【0108】

図 21 の系統樹の下帯には 4 つの細片があり、1 つの細片が処理の各ラウンドに対応し、一番上（系統樹の傍）に第 1 の処理が、一番下（系統樹から一番遠い部分）に 4 番目の処理がある。各細片は色分けされて、そのラウンドでそのビーズすなわち試料に用いられた個々の処理を示す。すなわち、系統樹の所与の分岐に対応する試料の処理が、分岐の直下に各細片に 1 つの 4 つの色によって表わされている。図の最下部の横棒には、対応する処理数に各色を関連付けるキーが設けられている。

【0109】

また図 21 上に示されているのは集団 A および集団 B であり、図 7 に示した集団上に同様に表記されたものと同じ 2 つの集団である。具体的には、集団 A は経路 AX - BX - C8 / C9 - D7 に対応し、一方集団 B は A10 - B1 - CX - D1 / D5 / D6 に対応する（または、少し一般的には A10 - B1 - CX - DX）。上記のように、これらの 2 つの集団は細胞を異なる表現型に分化させる。階層的集団化へのワード法の使用が、細胞培養データの検討および解析に効果的な道具を提供することがわかる。

【0110】

要約すれば、上記実施形態を実施例だけで説明してきたが、当業者は、添付された特許請求の範囲によって定義される本発明の範囲にある多くの潜在的な修正あるいは変更を想起することができる。

【先行技術文献】

【特許文献】

【0111】

【特許文献 1】欧州特許出願公開第 1551954 (A) 号明細書

【特許文献 2】国際公開第 2007/023297 号パンフレット

【非特許文献】

【0112】

【非特許文献 1】“Differentiation-on-a-chip: A microfluidic platform for long-term cell culture studies (半導体素子による差別化: 細胞培養の長期研究のための微小流体プラットフォーム)”, Anna Tourovskaia, Xavier Figueroa-Masot and Albert Folch; Lab Chip, 2005, 5, 14 - 19

【非特許文献 2】<http://biopoets.berkeley.edu/publications/2006uTAS-Dino-Culture>

10

20

30

40

50

.pdf

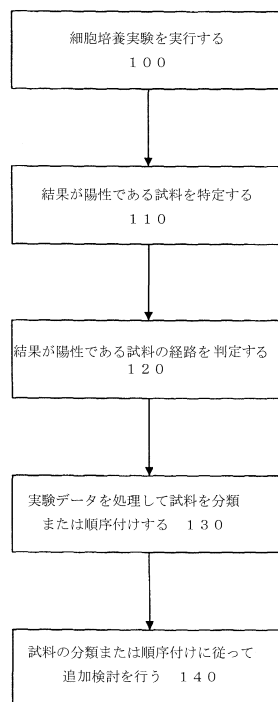
【非特許文献 3】“Cluster Analysis and Display of Genome-Wide Expression Patterns (集団化解析と遺伝子の表示 - 広範な表現型)” by M Eisen et al, Proceedings of the National Academy of Science, USA, 1998, 95 14863

【非特許文献 4】<http://www.geo.vu.nl/~huik/cluster.htm>

【非特許文献 5】<http://bonsai.ims.u-tokyo.ac.jp/~mdehoon/software/cluster/software.htm>.

【非特許文献 6】<http://www.statistics.com/resources/glossary/w/wardslnkg.php>

【図 1】



【図 1 A】

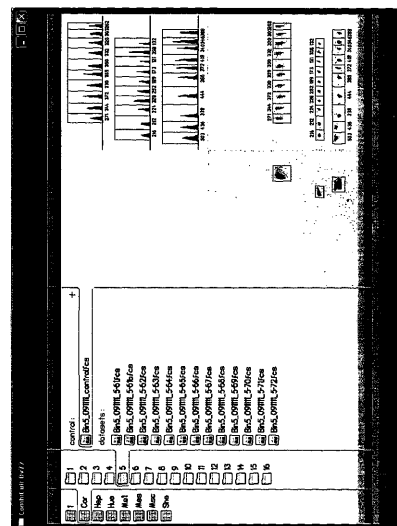


Figure 1A

【図 1 B】

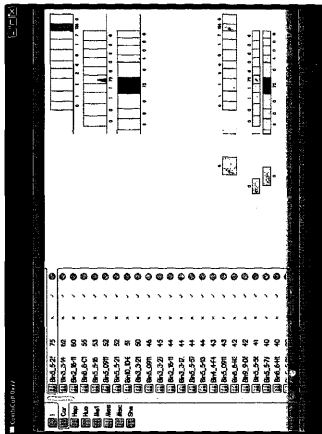


Figure 1B

【図 1 C】

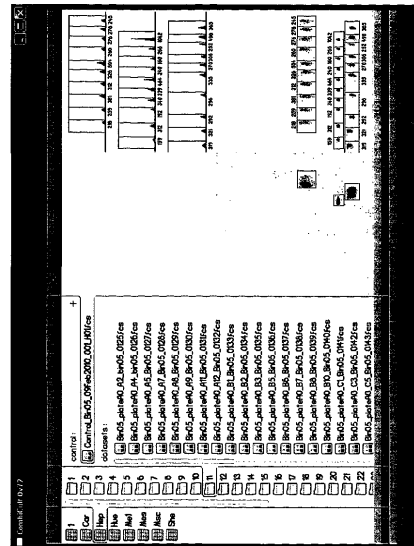


Figure 1C

【図 1 D】

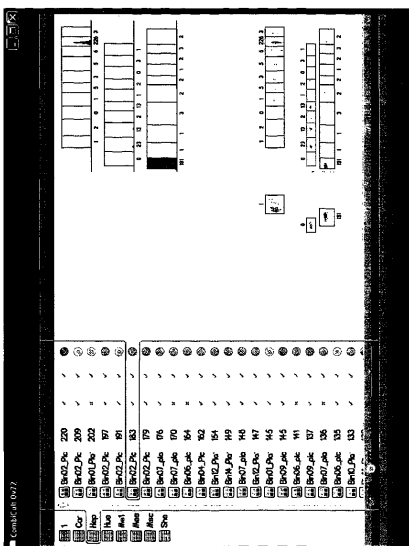
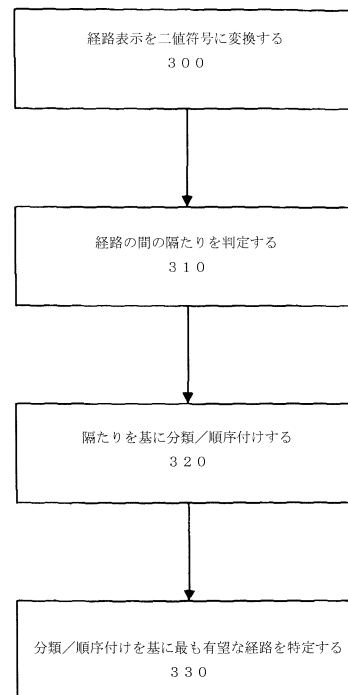


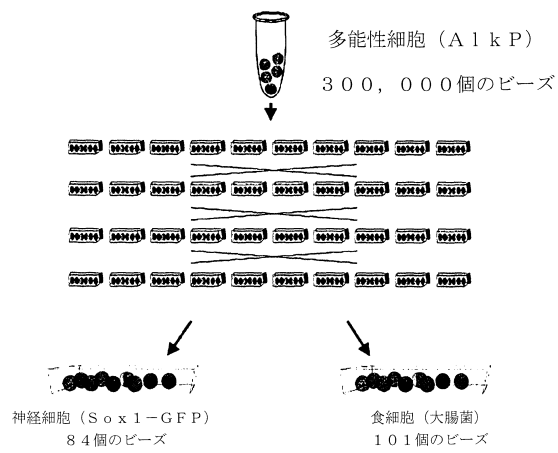
Figure 1D

【図 2】

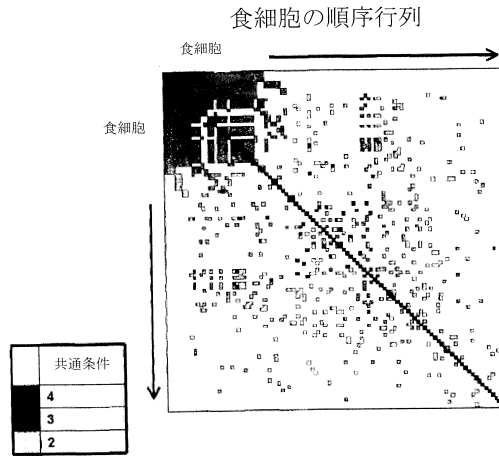


【図 3】

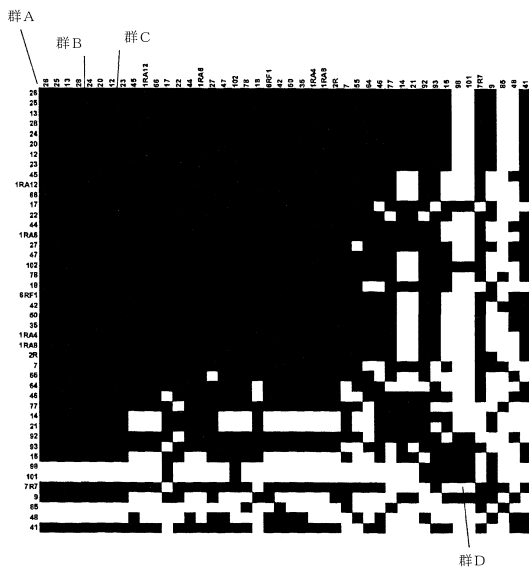
検査と結果



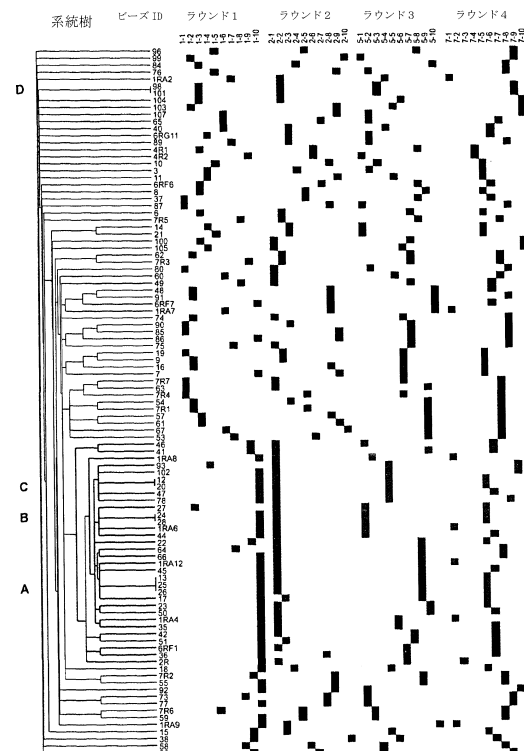
【図 4】



【図 4 A】



【図 5】



【図 5 A】

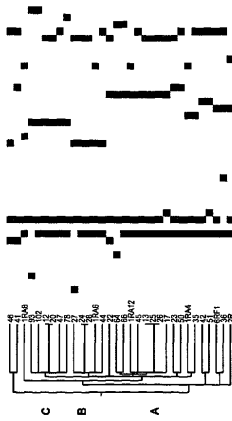


Figure 5A

【図 6 A】

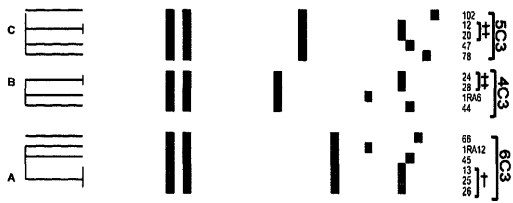


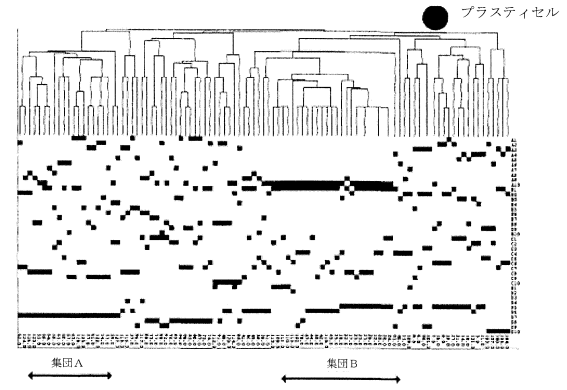
Figure 6A

【図 6 B】



Figure 6B

【図 7】



【図 8 A】

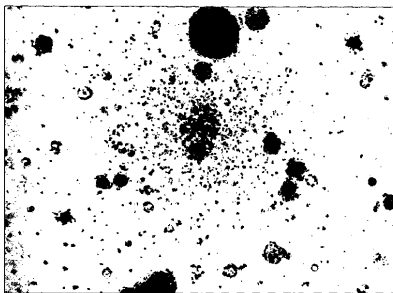


Figure 8A

【図 8 B】

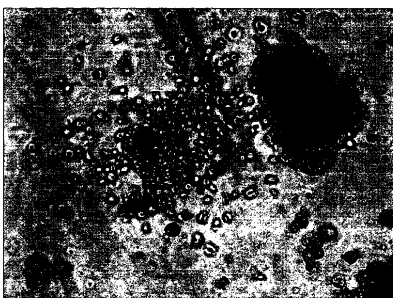


Figure 8B

【図 9 A】

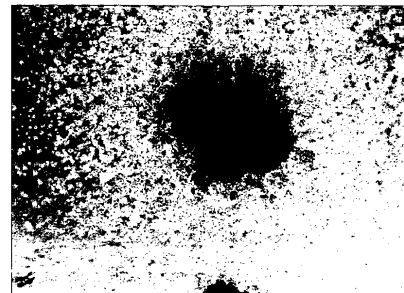


Figure 9A

【図 9 B】

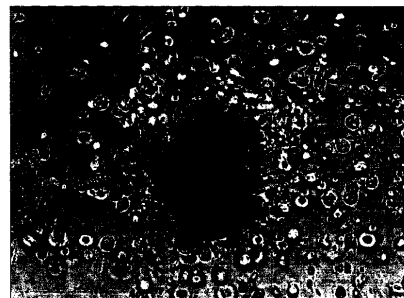


Figure 9B

【図10A】

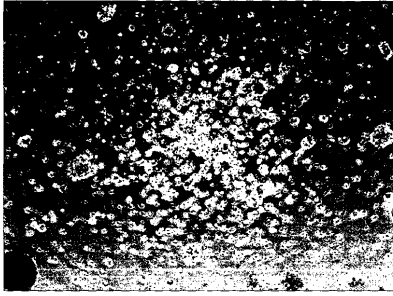


Figure 10A

【図10B】

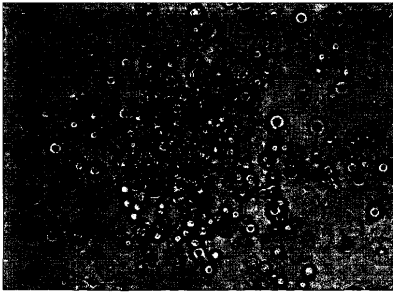
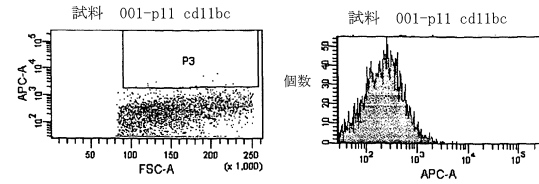
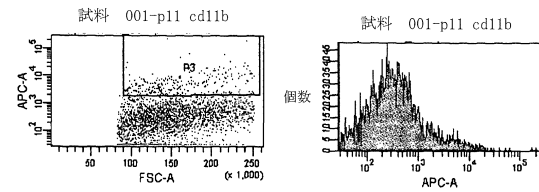


Figure 10B

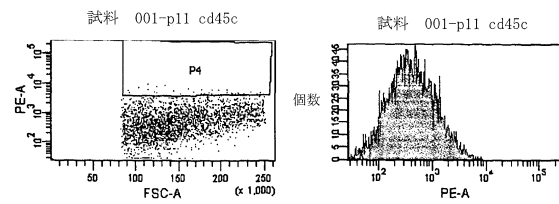
【図11A】



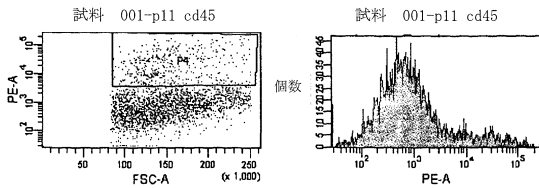
【図11B】



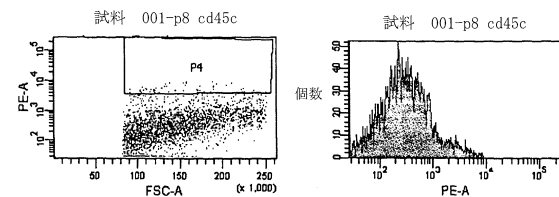
【図12A】



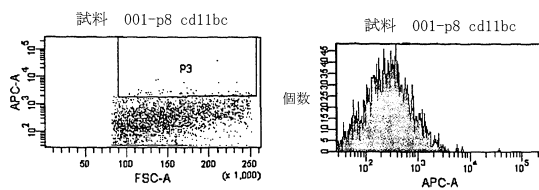
【図12B】



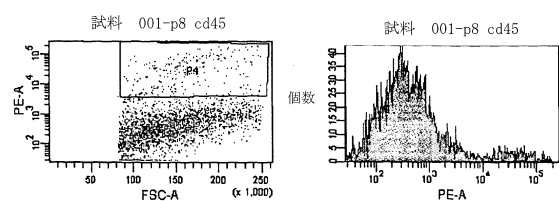
【図14A】



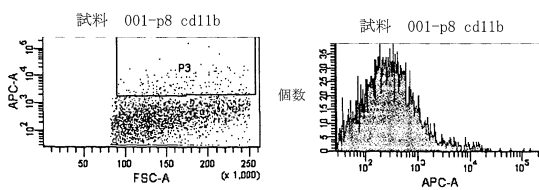
【図13A】



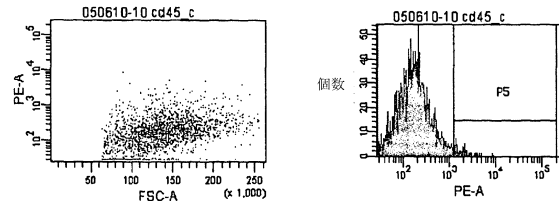
【図14B】



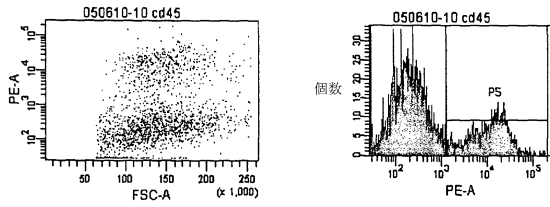
【図13B】



【図15A】



【図 15 B】



【図 16 A】

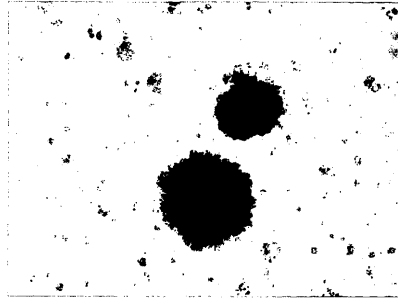


Figure 16A

【図 16 B】

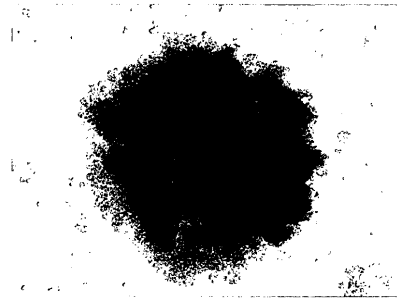
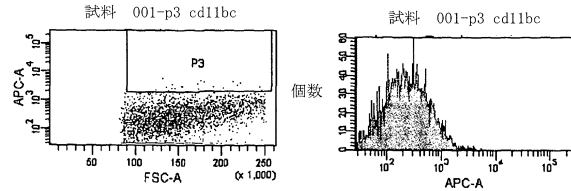
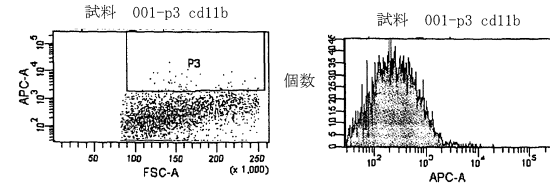


Figure 16B

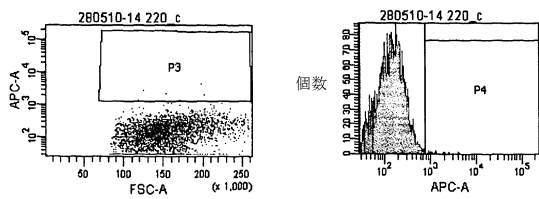
【図 17 A】



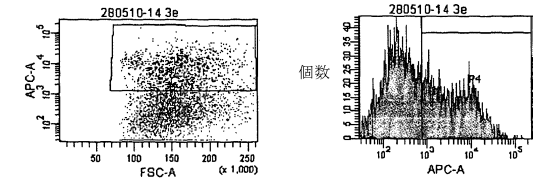
【図 17 B】



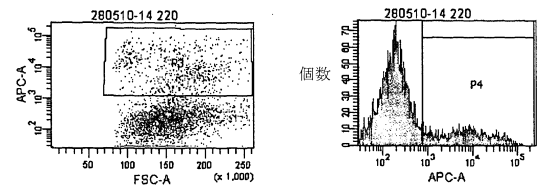
【図 18 A】



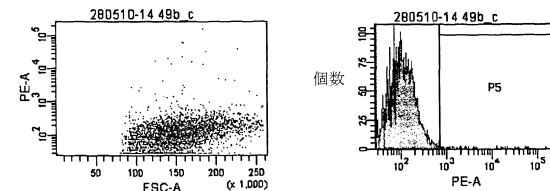
【図 19 B】



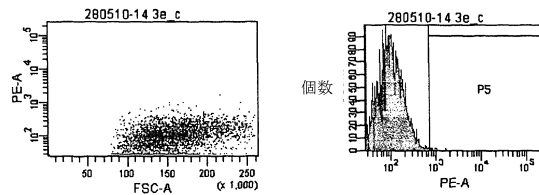
【図 18 B】



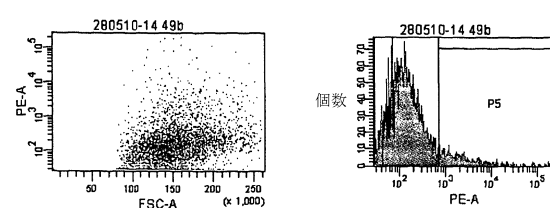
【図 20 A】



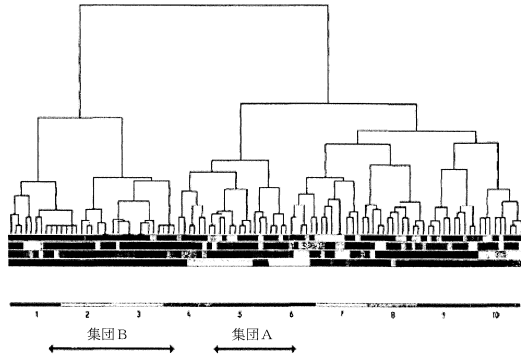
【図 19 A】



【図 20 B】



【図 21】



 フロントページの続き

- (72)発明者 タルニナ、マリナ
イギリス、ケント ピーアール5 1ディーエックス、オーピントン、ペッツ ウッド、トランス
ミア クローズ 49エー
- (72)発明者 チュー、イェン
イギリス、ロンドン エスダブリュー3 6ピーピー、シドニー ストリート 10
- (72)発明者 リスター、サイモン
イギリス、ハートフォードシャー エスジ-3 6ピーアール、ネブワース、オールド ネブワー
ス、パーク レーン 174
- (72)発明者 メリフィールド、ロバート
イギリス、エンフィールド イーエヌ2 0キューイー、カレッジ ロード 1

審査官 松原 寛子

- (56)参考文献 米国特許出願公開第2006/0035211(US, A1)
特表2009-506314(JP, A)
特表2006-500941(JP, A)
特表2007-517524(JP, A)
特開平03-259078(JP, A)

(58)調査した分野(Int.Cl., DB名)

C12Q 1/02
C12M 1/34
G06F 19/24
CAplus/MEDLINE/EMBASE/BIOSIS(STN)
JSTPlus/JMEDPlus/JST7580(JDreamIII)