(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2014 年 8 月 21 日(21.08.2014)

WIPO | PCT

(10) 国際公開番号
WO 2014/126213 A1

(51) 国際特許分類:
G06F 19/22 (2011.01) *C12N 15/115* (2010.01)

(21) 国際出願番号: PCT/JP2014/053516

(22) 国際出願日: 2014 年 2 月 14 日(14.02.2014)

(25) 国際出願の言語: 日本語

(26) 国際公開の言語: 日本語

(30) 優先権データ:
特願 2013-027851 2013 年 2 月 15 日(15.02.2013) JP

(71) 出願人: ＮＥＣソリューションイノベータ株式会社(NEC SOLUTION INNOVATORS, LTD.) [JP/JP]; 〒1368627 東京都江東区新木場一丁目１８番７号 Tokyo (JP).

(72) 発明者: 秋冨 穣(AKITOMI Jou); 〒1368627 東京都江東区新木場一丁目１８番７号 ＮＥＣソリューションイノベータ株式会社内 Tokyo (JP). 堀井 克紀(HORII Katsunori); 〒1368627 東京都江東区新木場一丁目１８番７号 ＮＥＣソリューションイノベータ株式会社内 Tokyo (JP).

(74) 代理人: 辻丸 光一郎, 外(TSUJIMARU Koichiro et al.); 〒6008813 京都府京都市下京区中堂寺南町１３４ 京都リサーチパーク１号館３０１号室 Kyoto (JP).

(81) 指定国 (表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) 指定国 (表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

(54) **Title**: SELECTION DEVICE FOR CANDIDATE SEQUENCE INFORMATION FOR SIMILARITY DETERMINATION, SELECTION METHOD, AND USE FOR SUCH DEVICE AND METHOD

(54) 発明の名称 : 類似判断の候補配列情報の選択装置、選択方法、およびそれらの用途

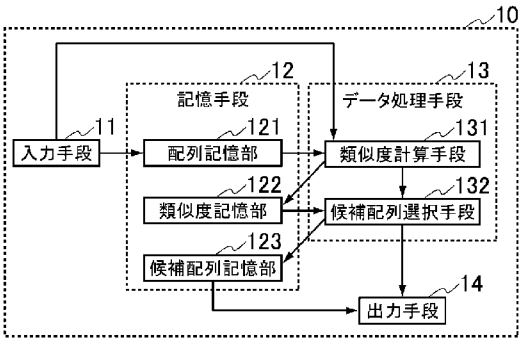


FIG. 1:
11 Input means
12 Storage means
13 Data processing means
14 Output means
121 Sequence storage section
122 Degree of similarity storage section
123 Candidate sequence storage means
131 Degree of similarity calculation means
132 Candidate sequence selection means

(57) **Abstract**: Provided is a device for easily determining the similarities between sequence information items. This candidate selection device (10) is provided with an input means (11), a sequence storage section (121), a degree of similarity storage section (122), a candidate sequence storage section (123), a degree of similarity calculation means (131), a candidate sequence selection means (132), and an output means (14). The input means (11) inputs information on sequence groups and virtual sequence groups. The degree of similarity calculation means (131) selects from the sequence groups a comparison source and a comparison destination, and calculates the difference in frequency of the virtual sequences with respect to a comparison source sequence and a comparison destination sequence, as a degree of similarity of the comparison destination sequence with respect to the comparison source sequence. In a case where the degree of similarity of the comparison destination sequence with respect to the comparison source sequence fulfills the allowable conditions of the degree of similarity set to the virtual sequence group, the candidate sequence selection means (132) selects the comparison source sequence and the comparison destination sequence as a candidate sequence group for determining the similarities between sequences. For the candidate sequence group, by determining the similarities between sequences, one sequence and a sequence similar thereto are selected as a similar sequence information group.

(57) 要約:

[続葉有]

WO 2014/126213 A1

添付公開書類:

— 国際調査報告（条約第 21 条(3)）

配列情報間の類似を、容易に判断するための装置を提供する。本発明の候補選択装置１０は、入力手段１１、配列記憶部１２１、類似度記憶部１２２、候補配列記憶部１２３、類似度計算手段１３１、候補配列選択手段１３２、出力手段１４を備える。入力手段１１は、配列群および仮想配列群の情報を入力し、類似度計算手段１３１は、前記配列群から比較元と比較先とを選択し、比較元配列と前記比較先配列との前記各仮想配列の頻度の相違を、前記比較元配列に対する前記比較先配列の類似度として計算する。候補配列選択手段１３２は、前記比較元配列に対する前記比較先配列の類似度が、前記仮想配列群に設定した類似度の許容条件を満たす場合、前記比較元配列および前記比較先配列を、配列間の類似を判断する候補配列群として選択する。前記候補配列群について、配列間の類似を判断することにより、ある配列とこれに類似する配列とを類似配列情報群として選択する。

明　細　書

発明の名称 ：

類似判断の候補配列情報の選択装置、選択方法、およびそれらの用途

技術分野

[0001] 本発明は、配列情報群における配列情報間の類似の判断に関する発明であり、具体的には、配列情報から類似判断の候補配列情報を選択する候補選択方法、候補配列情報から類似配列情報群を選択する類似選択方法、目的の類似配列情報群の濃縮を判定する判定方法、およびこれらの方法を実行する各装置、プログラムならびに記録媒体に関する。

背景技術

[0002] 近年、抗体に代わるターゲットへの結合分子として、いわゆるアプタマーと呼ばれる核酸分子の開発が進められている。前記アプタマーは、一般に、ＳＥＬＥＸ（Systematic Evolution of Ligands by EXponential enrichiment）法により調製されている（特許文献１、非特許文献１）。ＳＥＬＥＸ法は、核酸ライブラリーと前記ターゲットとの接触、および、前記ターゲットに結合した核酸の増幅を、１セットの選択処理とし、複数ラウンドを繰り返し行う。これによって、初期のライブラリーから、ラウンド毎のライブラリーにおいて前記ターゲットに結合する核酸配列が濃縮される。そして、例えば、ライブラリー内で濃縮度合いが相対的に高い複数の核酸配列を、アプタマー候補群として選択し、さらに、前記ターゲットとの結合力等を評価することによって、最終的に前記ターゲットに結合するアプタマーを決定することができる。

[0003] このように、アプタマー候補群は、ライブラリー内における濃縮度合いによって選択できるため、ＳＥＬＥＸ法においては、濃縮度合いの評価が必要である。濃縮度合いの評価は、通常、以下のように行われている。まず、各ラウンドのライブラリーに含まれる核酸配列をシーケンスで解読する。そして、ライブラリー内における同じ核酸配列の出現数（以下、重複度ともいう

）をカウントする。このカウント数の増減により、各核酸配列の濃縮度合いを評価する。例えば、ｎ回目のラウンド（R_n）における核酸配列Ｘの重複度m_nと、次のラウンド、すなわちｎ＋１回目のラウンド（R_{n+1}）における核酸配列Ｘの重複度m_{n+1}とを比較して、重複度m_n＜重複度m_{n+1}であれば、核酸配列Ｘは、ラウンド（ｎ＋１）において、ラウンド（ｎ）よりも濃縮されていると判断できる。また、同じラウンドのライブラリー内において、核酸配列Ｘの重複度m_Xと核酸配列Ｙの重複度m_Yとを比較して、重複度の大きい方が、他方に比べて濃縮されていると判断できる。

先行技術文献

特許文献

[0004] 特許文献1：特許第２７６３９５８号

非特許文献

[0005] 非特許文献1：Ｓｃｉｅｎｃｅ．（１９９０）２４９，５０５－５１０．

発明の概要

発明が解決しようとする課題

[0006] しかしながら、濃縮度合いによってアプタマー候補群を選択しても、異なる全ての核酸配列について、前記ターゲットとの結合力を評価することは、非常に労力を有し、現実的ではない。

[0007] 一方、ライブラリー内には、ある核酸配列（以下、元配列ともいう）に対して完全に同じ塩基配列も含まれるが、前記元配列に対して数塩基程度のミスマッチを有する類似した核酸配列（以下、類似配列ともいう）が含まれる場合がある。そして、発明者らは、前記類似配列は、例えば、前記ターゲットとの結合の強さが前記元配列と異なることがあるが、前記ターゲットに対する特性等は、前記元配列と同一であることが多いとの知見を得ている。このため、核酸配列について、完全に同一か否かという分類ではなく、許容できる範囲で類似し合っている核酸配列を、同一の配列群とすることにより、アプタマーの評価を効率化できる。しかしながら、この場合、複数の核酸配

列を一個ずつ照らし合わせて類似か否かを判断することも、労力、コストおよび時間がかかる。特に、次世代シーケンサー等を用いて大量の核酸配列の情報が得られた場合等、非常に計算コストがかかる。また、このような問題は、核酸配列に特化した問題ではなく、要素が並んだ配列情報について、共通する問題である。

[0008] そこで、本発明は、容易に、配列情報間の類似を判断するための装置、方法、プログラムおよび記録媒体を提供することを目的とする。

課題を解決するための手段

[0009] 前記目的を達成するために、本発明の候補選択装置は、下記（ａ）、（ｂ）、（ｃ）および（ｄ）手段を備えることを特徴とする、配列情報群から、配列情報間の類似の判断候補となる候補配列情報群を選択する候補選択装置である。

（ａ）配列情報群の各配列情報について、仮想配列情報群の各仮想配列情報の頻度をカウントする工程を実行する手段

（ｂ）前記配列情報群から、比較元となる配列情報と比較先となる配列情報とを選択する工程を実行する手段

（ｃ）前記比較元配列情報の前記各仮想配列情報の頻度と、前記比較先配列情報の前記各仮想配列情報の頻度との相違を、前記比較元配列情報に対する前記比較先配列情報の類似度として計算する工程を実行する手段

（ｄ）前記比較元配列情報に対する前記比較先配列情報の類似度が、前記仮想配列情報群に設定した類似度の許容条件を満たす場合、前記比較元配列情報および前記比較先配列情報を、配列情報間の類似を判断する候補配列情報群として選択する工程を実行する手段

[0010] 本発明の類似選択装置は、下記（Ａ）および（Ｂ）手段を備え、前記（Ａ）手段が、前記本発明の候補選択装置であることを特徴とする、配列情報群から、相互に類似する類似配列情報群を選択する類似選択装置である。

（Ａ）配列情報群から、配列情報間の類似の判断候補となる候補配列情報群を選択する工程を実行する手段

（Ｂ）前記候補配列情報群の各候補配列情報を相互に対比し、同一および類似する配列情報を類似配列情報群（Ｇ３）として選択する工程を実行する手段

[0011] 本発明の判定装置は、下記（Ｘ）および（Ｙ）手段を備え、前記（Ｘ）手段が、前記本発明の類似選択装置であることを特徴とする、目的の類似配列情報群の濃縮の判定装置である。

（Ｘ）配列情報群から、目的配列情報とそれに類似する配列情報とを目的の類似配列情報群として選択する工程を実行する手段

（Ｙ）前記類似配列情報群における前記目的配列情報と前記類似する配列情報との重複度の合計から、前記類似配列情報群の濃縮を判定する工程を実行する手段

[0012] 本発明の候補選択方法は、下記（ａ）、（ｂ）、（ｃ）および（ｄ）工程を含むことを特徴とする、配列情報群から、配列情報間の類似の判断候補となる候補配列情報群を選択する候補選択方法である。

（ａ）配列情報群の各配列情報について、仮想配列情報群の各仮想配列情報の頻度をカウントする工程

（ｂ）前記配列情報群から、比較元となる配列情報と比較先となる配列情報とを選択する工程

（ｃ）前記比較元配列情報の前記各仮想配列情報の頻度と、前記比較先配列情報の前記各仮想配列情報の頻度との相違を、前記比較元配列情報に対する前記比較先配列情報の類似度として計算する工程

（ｄ）前記比較元配列情報に対する前記比較先配列情報の類似度が、前記仮想配列情報群に設定した類似度の許容条件を満たす場合、前記比較元配列情報および前記比較先配列情報を、配列情報間の類似を判断する候補配列情報群として選択する工程

[0013] 本発明の類似選択方法は、下記（Ａ）および（Ｂ）工程を含み、

前記（Ａ）工程が、前記本発明の候補選択方法を含むことを特徴とする、配列情報群から、相互に類似する類似配列情報群を選択する類似選択方法であ

る。

（Ａ）配列情報群から、配列情報間の類似の判断候補となる候補配列情報群を選択する工程

（Ｂ）前記候補配列情報群の各候補配列情報を相互に対比し、同一および類似する配列情報を類似配列情報群（Ｇ３）として選択する工程

[0014] 本発明の判定方法は、下記（Ｘ）および（Ｙ）工程を含み、前記（Ｘ）工程が、前記本発明の類似選択方法を含むことを特徴とする、目的の類似配列情報群の濃縮の判定方法である。

（Ｘ）配列情報群から、目的配列情報とそれに類似する配列情報とを目的の類似配列情報群として選択する工程

（Ｙ）前記類似配列情報群における前記目的配列情報と前記類似する配列情報との重複度の合計から、前記類似配列情報群の濃縮を判定する工程

[0015] 本発明のプログラムは、前記本発明の候補選択方法、前記本発明の類似選択方法および前記本発明の判定方法からなる群から選択された少なくとも一つを、コンピュータ上で実行可能なことを特徴とするプログラムである。

[0016] 本発明の記録媒体は、前記本発明のプログラムを記録していることを特徴とする。

発明の効果

[0017] 本発明によれば、配列情報間の類似を判断するにあたって、まず、類似を判断するための候補配列群が選択される。このため、例えば、全ての配列情報間の類似を確認する従来の方法とは異なり、簡便に効率よく類似の判断を行うことができる。このため、例えば、アプタマーの濃縮の判定等についても、労力、時間およびコストの軽減が可能となる。

図面の簡単な説明

[0018] [図1]図１は、本発明の候補選択装置の実施形態を示すブロック図である。

[図2]図２は、本発明の候補選択方法および候補選択プログラムの実施形態を示すフローチャートである。

[図3]図３は、本発明の候補選択方法および候補選択プログラムの実施形態を

示すフローチャートである。

[図4]図４は、本発明の類似選択装置の実施形態を示すブロック図である。

[図5]図５は、本発明の類似選択方法および類似選択プログラムの実施形態を説明するためのフローチャートである。

[図6]図６は、本発明の類似選択方法および類似選択プログラムの実施形態を説明するためのフローチャートである。

[図7]図７は、本発明の類似選択装置のその他の実施形態を示すブロック図である。

[図8]図８は、本発明の類似選択方法および類似選択プログラムのその他の実施形態を説明するためのフローチャートである。

[図9]図９は、本発明の類似選択方法および類似選択プログラムのその他の実施形態を説明するためのフローチャートである。

発明を実施するための形態

[0019] 本発明において、「配列情報群」は、複数の配列情報から構成される群を意味し、前記複数の配列情報は、例えば、全て、異なる配列情報でもよいし、同じ配列情報と異なる配列情報とを含んでもよい。本発明は、異なる配列情報間における類似を判断するにあたって、類似判断の候補となる候補配列情報の選択を目的とする。このため、前記複数の配列情報は、例えば、全て、異なる配列情報が好ましい。前記配列情報群に含まれる前記配列情報の個数は、特に制限されない。

[0020] 本発明において、「配列情報」は、特に制限されず、要素の並びに関する情報である。前記要素は、例えば、文字および記号の少なくとも一方があげられ、具体例として、核酸の種類を示す文字または記号、アミノ酸の種類を示す文字または記号等があげられる。核酸の種類を示す文字または記号としては、例えば、Ａ、Ｇ、Ｃ、ＴおよびＵ等の塩基の種類を示す文字または記号があげられる。アミノ酸の種類を示す文字または記号としては、例えば、Ｍｅｔ等の３文字表記、Ｍ等の１文字表記の文字または記号があげられる。前記配列情報は、具体例として、核酸配列の配列情報、アミノ酸配列の配列

情報等があげられる。前記配列情報の長さは、前記配列情報を構成する要素の数ともいうことができる。前記配列情報の長さは、特に制限されず、要素が、例えば、５～２００個であり、好ましくは、１０～１５０個であり、さらに好ましくは２０～１２０個である。

[0021] 本発明において、「仮想配列情報群」は、複数の仮想配列情報から構成される群を意味する。前記仮想配列情報は、前記配列情報を構成する要素（構成単位ともいう）から構築された仮想の配列情報である。前記要素は、前記配列情報群の配列情報の種類に応じて決定でき、具体的には、前記配列情報群における配列情報と同じ要素である。前記仮想配列情報は、例えば、前記要素を任意に並べた情報ということができ、前記仮想配列情報群は、複数の、任意の異なる並びの情報から構成される群ということができる。前記仮想配列情報の長さは、前記仮想配列情報を構成する要素の数ともいうことができる。前記仮想配列情報の長さは、特に制限されず、要素が、例えば、１～１０個であり、好ましくは、１～７個であり、さらに好ましくは１～４個である。前記仮想配列情報群の各仮想配列情報は、例えば、全て同じ長さであることが好ましい。

[0022] 本発明において、前記配列情報群から選択した比較または対比し合う配列情報を、それぞれ、比較元配列情報および比較先配列情報という。ある配列情報に対して、他の配列情報を対比する場合、前者の配列情報を「比較元」ともいい、後者の他の配列情報を「比較先」ともいう。

[0023] 本発明において、「仮想配列情報の頻度」とは、対象となる配列情報において、前記仮想配列情報が出現する頻度を意味し、例えば、頻度ベクトルの要素、出現数ということもできる。また、「頻度の相違」とは、二つ以上の配列情報間の頻度の相違を意味し、例えば、比較先の配列情報の頻度と比較元の配列情報の頻度との相違である。

[0024] 本発明において、「類似度」は、比較元配列情報に対する比較先配列情報の類似の程度を示す。また、本発明において、「類似度の許容条件」は、前記比較元配列情報に対して、前記比較先配列情報が類似判断の候補となり得

ることを示す、類似度の条件である。前記類似度の許容条件は、任意に設定でき、例えば、２つの配列情報を対比した場合に許容できる要素のミスマッチの個数に基づいて設定できる。２つの配列情報の対比とは、例えば、２つの配列情報の要素の並びの対比である。前記類似度の許容条件は、例えば、２つの配列情報を対比した場合に許容できるミスマッチの個数（Ｍ）に、前記仮想配列情報の長さ（要素の個数Ｎ）を乗じた値を設定できる。

[0025] 本発明において、「重複度」とは、複数の配列情報から構成される配列情報群において、完全に同一である配列情報の個数を意味し、例えば、出現数ということもできる。また、本発明において、「類似重複度」とは、複数の配列情報から構成される配列情報群において、完全に同一である配列情報の重複度と、前記配列情報に類似する他の配列情報の重複度との合計を意味する。前記配列情報に対して、類似する他の配列情報が２つ以上存在する場合、例えば、前記配列情報と、類似する各他の配列情報との間の重複度の合計を、それぞれの類似重複度とする。

[0026] ＜本発明の候補選択装置および候補選択方法＞

本発明の候補選択装置は、前述のように、下記（ａ）、（ｂ）、（ｃ）および（ｄ）手段を備えることを特徴とする、配列情報群から、配列情報間の類似の判断候補となる候補配列情報群を選択する候補選択装置である。

（ａ）配列情報群の各配列情報について、仮想配列情報群の各仮想配列情報の頻度をカウントする工程を実行する手段

（ｂ）前記配列情報群から、比較元となる配列情報と比較先となる配列情報とを選択する工程を実行する手段

（ｃ）前記比較元配列情報の前記各仮想配列情報の頻度と、前記比較先配列情報の前記各仮想配列情報の頻度との相違を、前記比較元配列情報に対する前記比較先配列情報の類似度として計算する工程を実行する手段

（ｄ）前記比較元配列情報に対する前記比較先配列情報の類似度が、前記仮想配列情報群に設定した類似度の許容条件を満たす場合、前記比較元配列情報および前記比較先配列情報を、配列情報間の類似を判断する候補配列情報

群として選択する工程を実行する手段

[0027] 本発明の候補選択装置において、前記仮想配列情報群が、配列情報を構成する要素から構築された仮想配列情報の群であることが好ましい。

[0028] 本発明の候補選択装置において、前記（ｃ）手段が、下記（ｃ１）および（ｃ２）工程を実行する手段であることが好ましい。

（ｃ１）前記仮想配列情報ごとに、前記比較元配列情報における頻度と前記比較先配列情報における頻度との差を求める工程

（ｃ２）前記各仮想配列情報の頻度の差のうち、正数の差のみの総和の絶対値または負数の差のみの総和の絶対値を求め、前記絶対値を、前記比較元配列情報に対する前記比較先配列情報の類似度として計算する工程

[0029] 本発明の候補選択装置において、前記類似度の許容条件が、２つの配列情報を対比した場合に許容できるミスマッチの個数に基づき設定された条件であることが好ましい。２つの配列情報の対比とは、２つの配列情報のアラインメントということもできる。

[0030] 本発明の候補選択装置において、例えば、前記配列情報が、塩基配列であり、前記配列情報を構成する要素が、Ａ、Ｇ、Ｃ、ＴおよびＵの塩基であることが好ましい。

[0031] 本発明の候補選択装置において、前記仮想配列情報の塩基長が、例えば、１～１０塩基長であることが好ましい。

[0032] 本発明の候補選択装置において、前記仮想配列情報群の各仮想配列情報が、すべて同じ塩基長であることが好ましい。

[0033] 本発明の候補選択装置において、前記類似度の許容条件が、２つの配列情報を対比した場合に許容できるミスマッチの塩基数に基づき設定された条件であることが好ましい。

[0034] 本発明の候補選択装置において、前記類似度の許容条件が、２つの配列情報を対比した場合に許容できるミスマッチの塩基数（Ｍ）に前記仮想配列情報の塩基長（Ｎ）を乗じた値であることが好ましい。

[0035] 本発明の候補選択装置は、さらに、下記（ｅ）手段を有することが好まし

い。

（ｅ）前記（ｂ）、（ｃ）および（ｄ）手段による各工程の反復を実行する手段

この場合、前記（ｂ）手段は、例えば、前記工程の実行ごとに、前記配列情報群から、異なる配列情報を前記比較元配列情報として選択することが好ましい。

[0036] 本発明の候補選択方法は、前述のように、下記（ａ）、（ｂ）、（ｃ）および（ｄ）工程を含むことを特徴とする、配列情報群から、配列情報間の類似の判断候補となる候補配列情報群を選択する候補選択方法である。本発明の候補選択方法は、特に示さない限り、前記本発明の候補選択装置における説明を援用できる。

（ａ）配列情報群の各配列情報について、仮想配列情報群の各仮想配列情報の頻度をカウントする工程

（ｂ）前記配列情報群から、比較元となる配列情報と比較先となる配列情報とを選択する工程

（ｃ）前記比較元配列情報の前記各仮想配列情報の頻度と、前記比較先配列情報の前記各仮想配列情報の頻度との相違を、前記比較元配列情報に対する前記比較先配列情報の類似度として計算する工程

（ｄ）前記比較元配列情報に対する前記比較先配列情報の類似度が、前記仮想配列情報群に設定した類似度の許容条件を満たす場合、前記比較元配列情報および前記比較先配列情報を、配列情報間の類似を判断する候補配列情報群として選択する工程

[0037] 本発明の候補選択方法は、前記仮想配列情報群が、配列情報を構成する要素から構築された仮想配列情報の群であることが好ましい。

[0038] 本発明の候補選択方法は、前記（ｃ）工程が、下記（ｃ１）および（ｃ２）工程を含むことが好ましい。

（ｃ１）前記仮想配列情報ごとに、前記比較元配列情報における頻度と前記比較先配列情報における頻度との差を求める工程

（ｃ２）前記各仮想配列情報の頻度の差のうち、正数の差のみの総和の絶対値または負数の差のみの総和の絶対値を求め、前記絶対値を、前記比較元配列情報に対する前記比較先配列情報の類似度として計算する工程

[0039] 本発明の候補選択方法は、前記類似度の許容条件が、２つの配列情報を対比した場合に許容できるミスマッチの個数に基づき設定された条件であることが好ましい。

[0040] 本発明の候補選択方法は、前記配列情報が、塩基配列であり、前記配列情報を構成する要素が、Ａ、Ｇ、Ｃ、ＴおよびＵの塩基であることが好ましい。

[0041] 本発明の候補選択方法は、前記仮想配列情報の塩基長が、１～１０塩基長であることが好ましい。

[0042] 本発明の候補選択方法は、前記仮想配列情報群の各仮想配列情報が、すべて同じ塩基長であることが好ましい。

[0043] 本発明の候補選択方法は、前記類似度の許容条件が、２つの配列情報を対比した場合に許容できるミスマッチの塩基数に基づき設定された条件であることが好ましい。

[0044] 本発明の候補選択方法は、前記類似度の許容条件が、２つの配列情報を対比した場合に許容できるミスマッチの塩基数（Ｍ）に前記仮想配列情報の塩基長（Ｎ）を乗じた値であることが好ましい。

[0045] 本発明の候補選択方法は、さらに、下記（ｅ）工程を含むことが好ましい。この場合、前記（ｂ）工程において、前記工程の実行ごとに、前記配列情報群から、異なる配列情報を前記比較元配列情報として選択することが好ましい。

（ｅ）前記（ｂ）、（ｃ）および（ｄ）工程を反復する工程

[0046] 本発明の候補選択方法は、前記各工程が、全て、コンピュータ上で実行されることが好ましい。本発明の候補選択方法は、例えば、前記各工程が、全て、前記本発明の候補選択装置により実行されてもよい。

[0047] 以下、図面を参照しながら本発明のさらに具体的な実施形態について説明

する。ただし、本発明は、以下の実施形態に限定されない。以下、配列情報を配列、配列情報群を配列群と示す。

[0048] ［実施形態１］

実施形態１は、本発明の候補選択装置および候補選択方法に関する。本実施形態は、前記配列として、核酸の塩基配列を使用する一例である。

[0049] 本実施形態によれば、複数の塩基配列からなる塩基配列群から、塩基配列間の類似の判断候補となる候補配列群を選択できる。

[0050] 図１に、本実施形態の候補選択装置の一例の構成を示す。図１に示すように、候補選択装置１０は、入力手段１１、配列記憶部１２１、類似度記憶部１２２および候補配列記憶部１２３、類似度計算手段１３１および候補配列選択手段１３２、ならびに出力手段１４を備える。類似度計算手段１３１および候補配列選択手段１３２は、例えば、図１に示すように、ハードウェアであるデータ処理手段（データ処理装置）１３に組み込まれてもよく、ソフトウェアまたは前記ソフトウェアが組み込まれたハードウェアでもよい。各記憶部１２１、１２２、１２３は、例えば、図１に示すように、ハードウェアである記憶手段１２に組み込まれてもよい。データ処理手段１３は、ＣＰＵ等を備えてもよい。

[0051] 配列記憶部１２１は、入力手段１１および類似度計算手段１３１と、類似度記憶部１２２は、類似度計算手段１３１および候補配列選択手段１３２と、候補配列記憶部１２３は、候補配列選択手段１３２および出力手段１４と、それぞれ電気的に接続されている。また、入力手段１１は、類似度計算手段１３１と、類似度計算手段１３１は、候補配列選択手段１３２と、候補配列選択手段１３２は、出力手段１４と、それぞれ電気的に接続されてよい。候補選択装置１０は、例えば、情報を記憶手段１２に記憶させ、記憶させた情報をデータ処理手段１３に出力してデータ処理を行ってもよいし、前記情報をデータ処理手段１３に入力してデータ処理を行ってもよい。

[0052] 入力手段１１は、配列群および仮想配列群の情報を入力する手段（入力装置）である。入力手段１１は、特に制限されず、例えば、キーボード、マウ

ス等のコンピュータに備わる通常の入力手段、入力ファイルおよび他のコンピュータ等を用いることができる。入力手段１１は、例えば、データベースに格納された、前記配列群および仮想配列群の情報を読み出す手段でもよい。この場合、例えば、予めサーバに格納された配列情報が、回線網を通じて、入力手段１１に呼び出される。また、入力手段１１は、例えば、通信インターフェースを含んでもよい。

[0053] 前記配列群における入力する配列の数は、特に制限されず、下限は、例えば、５個、好ましくは１０個であり、上限は、例えば、１０００万個、好ましくは１００万個である。入力する配列の情報項目は、例えば、配列を構成する要素の順序、すなわち塩基の並びである。前記配列の長さは、特に制限されず、例えば、５～２００塩基長であり、好ましくは、１０～１５０塩基長であり、さらに好ましくは２０～１２０塩基長である。

[0054] 前記仮想配列群の仮想配列の数は、特に制限されず、前記仮想配列の塩基長に応じて適宜決定できる。前記塩基長は、その下限が、例えば、１塩基長であり、好ましくは２塩基長であり、より好ましくは３塩基長であり、その上限が、例えば、１０塩基長であり、好ましくは９塩基長であり、より好ましくは８塩基長であり、さらに好ましくは７塩基長である。前記仮想配列群において、前記各仮想配列の長さは、全て同じ長さが好ましい。

[0055] 前記仮想配列を構成する要素が４つの塩基（Ａ、Ｃ、Ｇ、およびＴまたはＵ）であり、前記仮想配列の塩基長がｎ（正数）の場合、前記仮想配列群における前記仮想配列の数は、例えば、４のｎ乗個（4^n個）である。具体例として、前記要素が４つの塩基Ａ、Ｃ、ＧおよびＴの場合、前記１塩基長の仮想配列の数は、４の１乗、つまり、Ａ、Ｃ、ＧおよびＴの４個であり、前記２塩基長の仮想配列の数は、４の２乗、つまり、ＡＡ、ＡＣ、ＡＧ、ＡＴ、ＣＣ、ＣＡ、ＣＧ、ＣＴ、ＧＧ、ＧＡ、ＧＣ、ＧＴ、ＴＴ、ＴＡ、ＴＣ、ＴＧの１６個である。

[0056] 類似度計算手段１３１は、前記（ａ）工程として、前記配列群の各配列について各仮想配列群の頻度のカウント、前記（ｂ）工程として、前記配列群

からの比較元配列と比較先配列との選択、前記（ｃ）工程として、前記比較元配列に対する前記比較先配列の類似度の計算を行う。前記（ａ）、（ｂ）および（ｃ）工程の順序は、特に制限されず、順不同である。

[0057] 前記（ｃ）工程における前記類似度の計算は、前述のように、前記（ｃ１）として、前記仮想配列ごとに、前記比較元配列における頻度（S_n）と前記比較先配列における頻度（T_n）との差（$S_n - T_n$）を求め、前記（ｃ２）工程として、前記頻度の差（$S_n - T_n$）のうち、正数の差のみの総和の絶対値または負数の差のみの総和の絶対値を求めることで行える。すなわち、前記総和の絶対値を、前記類似度とする。

[0058] 候補配列選択手段１３２は、前記比較元配列に対する前記比較先配列の類似度と、前記仮想配列群に設定した類似度の許容条件とに基づいて、配列情報間の類似を判断する候補配列の選択を行う。ここで選択された複数の候補配列が、候補配列群となる。

[0059] 前記類似度の許容条件は、２つの配列を対比した場合に許容できるミスマッチの塩基数に基づき設定でき、具体例として、前記許容できるミスマッチの塩基数（Ｍ）に前記仮想配列の塩基長（Ｎ）を乗じた値（Ｎ×Ｍ）があげられる。例えば、塩基長Ｎ＝１の前記仮想配列（Ａ、Ｃ、ＧおよびＴ）であって、前記許容できるミスマッチの塩基数Ｍ＝２に設定した場合、許容条件（Ｎ×Ｍ）は、１×２＝２となる。そして、前記類似度が２以下の場合、許容条件の数値以下となり許容条件を満たすため、前記比較元配列および前記比較先配列は、配列情報間の類似を判断する候補配列として選択する。他方、前記類似度が２を超える場合、許容条件の数値を超え許容条件を満たさないため、前記比較先配列は、前記比較元配列との類似を判断する候補配列として選択しない。

[0060] 前記許容条件の一例として、前記許容できるミスマッチの塩基数（Ｍ）に前記仮想配列の塩基長（Ｎ）を乗じた値（Ｎ×Ｍ）を設定するのは、以下の理由による。例えば、以下の２つの配列をアラインメントした場合、大文字の１塩基がミスマッチである。これらの配列について、塩基長Ｎ＝２の仮想

配列の頻度をカウントした場合、対象元配列Ｓｅｑ１において、下線部がｃｇおよびｇｇとカウントされるのに対し、対象先配列Ｓｅｑ２において、下線部がｃＡおよびＡｇとカウントされる。つまり、許容できるミスマッチの塩基数が１であっても、１つミスマッチの存在によって、カウントされる仮想配列は、最大２つが変動することになる。このため、前記許容できるミスマッチの塩基数（Ｍ）に、前記仮想配列の塩基長（Ｎ）を乗じることで、カウントへの影響を補正できる。

対象元配列Ｓｅｑ１：ａａｃｃｇｇｔｔ

対象先配列Ｓｅｑ２：ａａｃｃＡｇｔｔ

[0061] 出力手段（出力装置）１４は、候補配列選択手段１３２の結果を出力する手段であればよい。また、前記出力手段１４は、候補配列記憶部１２３に記憶された情報を出力する手段でもよい。前記出力手段１４は、特に制限されず、例えば、ディスプレイ装置、印刷装置等のコンピュータに備わる通常の出力装置、出力ファイル、および、他のコンピュータ等を使用できる。

[0062] つぎに、図２および図３のフローチャートを参照し、本実施形態の候補選択方法を説明する。本実施形態の候補選択方法は、Ａ１ステップ（配列入力）、Ａ２ステップ（類似度計算）およびＡ３ステップ（候補配列選択）を含む。

[0063] （Ａ１）配列入力

配列群の各配列および仮想配列群の各仮想配列を、それぞれ入力し、配列記憶部１２１に記憶させる。前記配列群および前記仮想配列群の情報項目は、例えば、配列における塩基の順序があげられる。

[0064] （Ａ２）類似度計算

前記配列群から、新しい比較元配列のセット（Ａ２１）および新しい比較先配列のセット（Ａ２２）を行い、セットした前記比較元配列と前記比較先配列について、それぞれ、前記各仮想配列の頻度をカウントする。そして、各仮想配列について、前記比較元配列の頻度と前記比較先配列の頻度との差を求め、正数の差のみの総和または負数の差のみの総和を計算する。具体的

には、ｎ個（ｎは正数）の仮想配列が存在する場合、前記比較元配列について、各仮想配列の頻度としてｎ個の頻度（S_1、・・・・、S_n）、前記比較先配列について、ｎ個の頻度（T_1、・・・・、T_n）が得られる。そして、各仮想配列の頻度について、前記比較元配列と前記比較先配列との差、すなわち、（S_1-T_1）、・・・・、（S_n-T_n）を求め、正数の差のみの総和または負数の差のみの総和を計算し、総和の絶対値を求める。前記総和の絶対値が、前記比較元配列に対する前記比較先配列の類似度である。

[0065] （Ａ３）候補配列選択

そして、前記類似度が、類似度の許容値を満たすか否か、つまり、許容値よりも大きいか否かを判断する（Ａ３１）。ＮＯの場合、つまり、前記類似度が許容値よりも小さい場合、前記比較先配列は、前記比較元配列に対して許容できる数のミスマッチを有すると判断して、前記比較元配列と前記比較先配列が類似判断の候補配列であるとの結果を出力する（Ａ３２）。他方、ＹＥＳの場合、つまり、前記類似度が許容値よりも大きい場合、前記比較先配列は、前記比較元配列に対して許容できない数のミスマッチを有すると判断して、前記比較先配列が類似候補配列ではないとの結果を出力する（Ａ３３）。

[0066] その後は、未比較の比較先配列の有無を確認する（Ａ３４）。ＹＥＳの場合、つまり、未比較の比較先配列がある場合、Ａ２２ステップから同様の処理を行う。そして、ＮＯの場合、つまり、未比較の比較先配列がない場合、さらに、未比較の比較元配列の有無を確認する（Ａ３５）。ＹＥＳの場合、つまり、未比較の比較元配列がある場合、Ａ２１ステップから同様の処理を行い、ＮＯの場合、つまり、未比較の比較元配列がない場合、終了する。なお、ある配列を比較元配列とし他の配列を比較先配列として比較済みである場合、前者を比較先配列とし後者を比較元配列とする比較は、省略し、比較済みの結果を使用してもよい。

[0067] 前記Ａ２ステップおよびＡ３ステップについて、さらなる具体例として、前記仮想配列が塩基長１の場合を例にあげて説明する。

[0068] 塩基長Ｎ＝１の仮想配列を下記４種類、比較元配列をＳｅｑ３、比較先配列をＳｅｑ４と仮定する。そして、２つの配列をアラインメントした場合に、類似の判断候補として許容できるミスマッチの塩基数をＭとし、許容値をＮ×Ｍ＝１×Ｍ＝Ｍとする。

仮想配列：Ａ、Ｃ、ＧおよびＴ

比較元配列Ｓｅｑ３：ＡＣＧＴＡＣＧＴ

比較先配列Ｓｅｑ４：ＡＡＧＡＡＣＡＴ

[0069] 比較元配列Ｓｅｑ３および比較先配列Ｓｅｑ４における各仮想配列（Ａ、Ｃ、Ｇ、Ｔ）の頻度｛ｆＡ、ｆＣ、ｆＧ、ｆＴ｝は、それぞれ、ＳＥＱ１が｛２、２、２、２｝およびＳｅｑ２が｛５、１、１、１｝となり、各頻度｛ｆＡ、ｆＣ、ｆＧ、ｆＴ｝の差は、Ａが（２－５＝－３）、Ｃが（２－１＝１）、Ｇが（２－１＝１）、Ｔが（２－１＝１）となる。負数の差の総数（－３＋０＋０＋０＝－３）の絶対値は３であり、正数の差の総数（０＋１＋１＋１＝３）の絶対値は３である。この絶対値３が、比較元配列Ｓｅｑ３に対する比較先配列Ｓｅｑ４の類似度であり、比較先配列Ｓｅｑ４が、比較元配列Ｓｅｑ３とアラインメントした際に、少なくとも３つのミスマッチを有することを示す。前記許容できるミスマッチの上限塩基数Ｍを、例えば、２とした場合、許容値はＮ×Ｍ＝１×２＝２である。このため、計算した類似度と許容値とを対比すると、類似度３＞許容値２であるため、比較先配列Ｓｅｑ４は、比較元配列Ｓｅｑ３の類似判断の候補配列からはずす。他方、前記許容できるミスマッチの上限塩基数Ｍを、例えば、３とした場合、許容値はＮ×Ｍ＝１×３＝３である。このため、計算した類似度と許容値とを対比すると、類似度３＝許容値３であるため、比較先配列Ｓｅｑ４は、比較元配列Ｓｅｑ３の類似判断の候補配列として選択する。

[0070] このようにして、前記比較先配列が前記許容条件を満たす場合には、前記比較先配列は、前記比較元配列と共に、類似判断の候補配列として選択する。つまり、候補配列群として選択する。他方、前記比較先配列が前記許容条件を満たさない場合には、前記比較先配列は、類似判断の候補配列として選

択しない。また、前記比較元配列に対して、前記許容条件を満たす比較先配列が存在しない場合は、前記比較元配列も、類似判断の候補配列として選択しない。

[0071] 本実施形態における候補選択装置１０において、入力手段１１と類似度計算手段１３１、類似度計算手段１３１と候補配列選択手段１３２が、それぞれ電気的に接続されてもよい。また、候補選択装置１０は、例えば、各種記憶部を備えてもよいし、備えていなくてもよい。この場合、例えば、入力手段１１により入力された各配列について、類似度計算手段１３１により類似度を計算し、計算された類似度について、候補配列選択手段１３２により候補配列の選択を行ってもよい。

[0072] ＜本発明の類似選択装置および類似選択方法＞

本発明の類似選択装置は、前述のように、下記（Ａ）および（Ｂ）手段を備え、

前記（Ａ）手段が、前記本発明の候補選択装置であることを特徴とする、配列情報群から、相互に類似する類似配列情報群を選択する類似選択装置である。

（Ａ）配列情報群から、配列情報間の類似の判断候補となる候補配列情報群を選択する工程を実行する手段

（Ｂ）前記候補配列情報群の各候補配列情報を相互に対比し、同一および類似する配列情報を類似配列情報群（Ｇ３）として選択する工程を実行する手段

[0073] 本発明の類似選択装置において、前記（Ａ）手段は、前記本発明の候補選択装置であればよく、前記本発明の候補選択装置の記載を援用できる。

[0074] 本発明の類似選択装置は、前記配列情報群が、同一の配列情報および異なる配列情報からなる配列情報群（Ｇ）から選択された前記異なる配列情報の群であることが好ましい。

[0075] 本発明の類似選択装置は、前記（Ｂ）手段が、下記（Ｂ１）、（Ｂ２）、（Ｂ３）、（Ｂ４）および（Ｂ５）工程を実行する手段であることが好まし

い。

（Ｂ１）前記候補配列情報群から、比較元となる候補配列情報と比較先となる候補配列情報とを選択する工程

（Ｂ２）前記比較元候補配列情報に対する前記比較先候補配列情報の類似の有無を決定する工程

（Ｂ３）前記比較元候補配列情報の重複度と、前記比較元候補配列情報に類似する前記比較先候補配列情報の重複度とを合計し、得られた合計値を、前記比較元候補配列情報の類似重複度とする工程

（Ｂ４）前記候補配列情報群から、異なる候補配列情報を、新たな比較元となる候補配列情報として選択し、前記（Ｂ１）、（Ｂ２）および（Ｂ３）工程を反復する工程

（Ｂ５）前記候補配列情報のうち、最も大きな類似重複度を示した候補配列情報および前記候補配列情報に類似する候補配列情報を、類似配列情報群（Ｇ３）として選択する工程

[0076] 前記（Ｂ２）工程において、前記比較元候補配列と前記比較先候補配列との類似の有無は、特に制限されず、公知の方法で決定できる。具体的には、配列と配列とをアラインメントして、許容できるミスマッチ（異なる要素）の数に基づき、類似と非類似とを判断できる。具体例として、例えば、前記両配列をアラインメントした際、ミスマッチの数が、前記許容できるミスマッチの数を超える場合は非類似、前記許容できるミスマッチの数以下の場合は類似と判断できる。前記許容できるミスマッチの個数は、特に制限されず、任意に決定できる。

[0077] 重複度は、後の工程が繰り返される間に、０に再設定される。そこで、前記（Ｂ３）工程における重複度は、各配列の初期の情報であることから、「初期重複度」ともいう。また、後の工程において再設定した重複度０は、「重複度０」または「再設定重複度」ともいう。

[0078] 本発明の類似選択装置は、前記（Ｂ）手段が、さらに、下記（Ｂ６）、（Ｂ７）および（Ｂ８）工程を実行する手段であることが好ましい。類似重複

度の再算出とは、例えば、すでに得られた類似重複度をリセットし、新たに類似重複度を算出することを意味する。

（Ｂ６）前記候補配列情報のうち、最も大きな類似重複度を示した候補配列情報の重複度および前記候補配列情報に類似する候補配列情報の重複度を０に再設定する工程

（Ｂ７）重複度が０以外である他の候補配列情報について、類似重複度を再算出する工程

（Ｂ８）前記他の候補配列情報のうち、最も大きな類似重複度を示した候補配列情報および前記候補配列情報に類似する候補配列情報を、類似配列情報群として再選択する工程

[0079] 本発明の類似選択装置は、前記（Ｂ）手段が、さらに、下記（Ｂ９）の工程を実行する手段であることが好ましい。

（Ｂ９）前記他の候補配列情報のうち、最も大きな類似重複度を示した候補配列情報および前記候補配列情報に類似する候補配列情報の重複度を０に再設定し、前記（Ｂ７）および（Ｂ８）工程を反復する工程

[0080] このように、最も大きな類似重複度に基づく類似候補群の選択と、類似重複度の再計算とを繰り返すことによって、複数の類似配列情報群が選択できる。前記類似配列情報群の再選択は、例えば、全ての候補配列について重複度が０に再設定されるまで行うことが好ましい。

[0081] 本発明の類似選択装置は、前記（Ｂ）手段が、前記（Ｂ１）工程における前記比較元補配列情報と前記比較先候補配列情報との組合せとして、すでに実行した組合せの除外を実行することが好ましい。

[0082] 本発明の類似選択装置において、配列情報の情報項目として、例えば、配列を構成する要素の順序の他に、前記各配列の重複度を含んでもよい。この場合、前記配列群に含まれる配列は、全て、異なる配列であることが好ましい。また、配列情報の情報項目として、前記重複度を含まない場合、例えば、前記重複度をカウントする工程を実行する、下記（Ｂ’）手段を含んでもよい。この場合、前記配列群に含まれる配列は、例えば、異なる配列の他に

、完全に要素の順序が同じである配列を含んでもよい。

（Ｂ’）前記配列情報群について、完全に同一な配列情報の数を重複度としてカウントする工程を実施する手段

[0083] 本発明の類似選択方法は、前述のように、下記（Ａ）および（Ｂ）工程を含み、

前記（Ａ）工程が、前記本発明の候補選択方法を含むことを特徴とする、配列情報群から、相互に類似する類似配列情報群を選択する類似選択方法である。

（Ａ）配列情報群から、配列情報間の類似の判断候補となる候補配列情報群を選択する工程

（Ｂ）前記候補配列情報群の各候補配列情報を相互に対比し、同一および類似する配列情報を類似配列情報群（Ｇ３）として選択する工程

[0084] 本発明の類似選択方法は、前記（Ｂ）工程が、下記（Ｂ１）、（Ｂ２）、（Ｂ３）、（Ｂ４）および（Ｂ５）工程を含むことが好ましい。

（Ｂ１）前記候補配列情報群から、比較元となる候補配列情報と比較先となる候補配列情報とを選択する工程

（Ｂ２）前記比較元候補配列情報に対する前記比較先候補配列情報の類似の有無を決定する工程

（Ｂ３）前記比較元候補配列情報の重複度と、前記比較元候補配列情報に類似する前記比較先候補配列情報の重複度とを合計し、得られた合計値を、前記比較元候補配列情報の類似重複度とする工程

（Ｂ４）前記候補配列情報群から、異なる候補配列情報を、新たな比較元となる候補配列情報として選択し、前記（Ｂ１）、（Ｂ２）および（Ｂ３）工程を反復する工程

（Ｂ５）前記候補配列情報のうち、最も大きな類似重複度を示した候補配列情報および前記候補配列情報に類似する候補配列情報を、類似配列情報群（Ｇ３）として選択する工程

[0085] 本発明の類似選択方法は、前記（Ｂ）工程が、さらに、下記（Ｂ６）、（

Ｂ７）および（Ｂ８）工程を含むことが好ましい。

（Ｂ６）前記候補配列情報のうち、最も大きな類似重複度を示した候補配列情報の重複度および前記候補配列情報に類似する候補配列情報の重複度を０に再設定する工程

（Ｂ７）重複度が０以外である他の候補配列情報について、類似重複度を再算出する工程

（Ｂ８）前記他の候補配列情報のうち、最も大きな類似重複度を示した候補配列情報および前記候補配列情報に類似する候補配列情報を、類似配列情報群として再選択する工程

[0086] 本発明の類似選択方法は、前記（Ｂ）工程が、さらに、下記（Ｂ９）工程を含むことが好ましい。

（Ｂ９）前記他の候補配列情報のうち、最も大きな類似重複度を示した候補配列情報および前記候補配列情報に類似する候補配列情報の重複度を０に再設定し、前記（Ｂ７）および（Ｂ８）工程を反復する工程

[0087] 本発明の類似選択方法は、前記（Ｂ）工程において、前記（Ｂ１）工程における前記比較元補配列情報と前記比較先候補配列情報との組合せとして、すでに実行した組合せを除外することが好ましい。

[0088] 本発明の類似選択方法は、前記各工程が、全て、コンピュータ上で実行されることが好ましい。本発明の類似選択方法は、例えば、前記各工程が、全て、前記本発明の類似選択装置により実行されてもよい。

[0089] 以下、図面を参照しながら本発明のさらに具体的な実施形態について説明する。ただし、本発明は、以下の実施形態に限定されない。また、本実施形態において、前記候補配列群の選択は、前記実施形態１の記載を援用できる。以下、配列情報を配列、配列情報群を配列群と示す。

[0090] ［実施形態２］

実施形態２は、本発明の類似選択装置および類似選択方法に関する。本実施形態は、前記配列として、核酸の塩基配列を使用する一例である。本実施形態は、特に示さない限り、実施形態１の記載を援用できる。

[0091] 本実施形態によれば、複数の塩基配列からなる塩基配列群から、塩基配列間の類似の判断候補となる候補配列を選択し、複数の前記候補配列からなる候補配列群から、相互に類似する類似配列を類似配列群として選択できる。

[0092] 図４に、本実施形態の類似選択装置の一例を示す。図４において、図１の候補選択装置１０と同じ箇所には、同じ符号を付している。図４に示すように、類似選択装置２０は、入力手段１１、配列記憶部１２１、類似度記憶部１２２、候補配列記憶部１２３および類似配列記憶部１２４、類似度計算手段１３１、候補配列選択手段１３２および類似配列選択手段１３３、ならびに出力手段１４を備える。類似度計算手段１３１、候補配列選択手段１３２および類似配列選択手段１３３は、例えば、図４に示すように、ハードウェアであるデータ処理手段１３に組み込まれてもよく、ソフトウェアまたは前記ソフトウェアが組み込まれたハードウェアでもよい。各記憶部１２１、１２２、１２３、１２４は、例えば、図４に示すように、ハードウェアである記憶手段１２に組み込まれてもよい。データ処理手段１３は、ＣＰＵ等を備えてもよい。

[0093] 候補配列記憶部１２３は、さらに、類似配列選択手段１３３と電気的に接続されており、類似配列記憶部１２４は、類似配列選択手段１３３および出力手段１４と、それぞれ電気的に接続されている。また、候補配列選択手段１３２は、類似配列選択手段１３３と、類似配列選択手段１３３は、出力手段１４と、それぞれ電気的に接続されてよい。類似選択装置２０は、例えば、情報を記憶手段１２に記憶させ、記憶させた情報をデータ処理手段１３に出力してデータ処理を行ってもよいし、前記情報をデータ処理手段１３に入力してデータ処理を行ってもよい。

[0094] 本実施形態において、入力する配列の情報項目は、前述のような、配列を構成する要素の順序の他に、前記各配列の重複度を含むことが好ましい。前記情報項目として、前記重複度を含む場合、前記配列群を構成する配列は、全て、異なる配列であることが好ましい。

[0095] また、前記情報項目として、前記重複度を含まない場合、例えば、前記（

Ｂ’）手段を含んでもよい。前記（Ｂ’）手段により、前記配列群について、完全に同一な配列情報の数を重複度としてカウントできる。

[0096] つぎに、図５および図６のフローチャートを参照し、本実施形態の類似選択方法を説明する。本実施形態の類似選択方法は、Ａ１ステップ（配列入力）、Ａ２ステップ（類似度計算）、Ａ３ステップ（候補配列選択）およびＡ４ステップ（類似配列選択）を含む。図５において、図２と同じステップには同じ符号を付している。

[0097] 前記Ａ１ステップ、前記Ａ２ステップおよび前記Ａ３ステップは、前記実施形態１と同様に行うことができ、具体的には、前述した図３のフローチャートに従って行うことができる。前記配列入力において、前記配列群の情報項目は、例えば、配列における塩基の順序および配列の重複度があげられ、前記仮想配列群の情報項目は、例えば、配列における塩基の順序があげられる。

[0098] （Ａ４）類似配列選択

前記Ａ３ステップで選択された候補配列群から、新しい比較元候補配列のセット（Ａ４１）および新しい比較先候補配列のセット（Ａ４２）を行い、セットした前記比較先候補配列が、前記比較元候補配列に類似するか否かを判断する（Ａ４３）。そして、ＮＯの場合、つまり、前記比較先候補配列が、前記比較元候補配列に類似していない場合、前記比較先候補配列は、前記比較元候補配列との類似配列群ではないとの結果を出力する（Ａ４４）。他方、ＹＥＳの場合、つまり、前記比較先候補配列が、前記比較元候補配列に類似している場合、前記比較先候補配列は、前記比較元候補配列との類似配列群であるとの結果を出力する（Ａ４５）。

[0099] その後は、前記比較元候補配列に対して、未比較の比較先候補配列の有無を確認する（Ａ４６）。ＹＥＳの場合、つまり、未比較の比較先配列がある場合、Ａ４２ステップから同様の処理を行う。そして、ＮＯの場合、つまり、未比較の比較先候補配列がない場合、さらに、未比較の比較元候補配列の有無を確認する（Ａ４７）。ＹＥＳの場合、つまり、未比較の比較元候補配

列がある場合、Ａ４１ステップから同様の処理を行い、ＮＯの場合、つまり、未比較の比較元候補配列がない場合、終了する。なお、ある配列を比較元候補配列とし他の配列を比較先候補配列として比較済みである場合、前者を比較先候補配列とし後者を比較元候補配列とする比較は、省略し、比較済みの結果を使用してもよい。

[0100] このようにして、前記候補配列群における各候補配列から、前記比較元候補配列および前記比較先候補配列を、それぞれ順次セットし、配列間の類似を判断することによって、前記比較元候補配列とそれに類似する比較先候補配列とからなる類似配列群を選択できる。

[0101] 本実施形態における類似選択装置２０において、入力手段１１と類似度計算手段１３１、類似度計算手段１３１と候補配列選択手段１３２、候補配列選択手段１３２と類似配列選択手段１３３とが、それぞれ電気的に接続されてもよい。また、類似選択装置２０は、例えば、各種記憶部を備えてもよいし、備えていなくてもよい。この場合、例えば、入力手段１１により入力された各配列について、類似度計算手段１３１により類似度を計算し、計算された類似度について、候補配列選択手段１３２により候補配列群の選択を行い、さらに、選択された候補配列群について、類似配列選択手段１３３により類似配列群の選択を行ってもよい。

[0102] ［実施形態３］

実施形態３は、実施形態２と同様に、本発明の類似選択装置および類似選択方法に関する。本実施形態は、前記実施形態２の前記類似配列群の選択において、重複度を用いる一例である。本実施形態は、特に示さない限り、実施形態１および２の記載を援用できる。

[0103] 本実施形態によれば、配列間の類似度を用いることによって、簡便に、類似配列群を選択できる。

[0104] 図７に、本実施形態の類似選択装置の一例を示す。図７において、図４の類似選択装置２０と同じ箇所には、同じ符号を付している。図７に示すように、類似選択装置３０は、類似重複度記憶部１２４ａおよび類似配列記憶部

１２４ｂ、類似重複度計算手段１３３ａおよび類似配列選択手段１３３ｂを備える。類似重複度計算手段１３３ａおよび類似配列選択手段１３３ｂは、例えば、図７に示すように、ハードウェアであるデータ処理手段１３に組み込まれてもよく、ソフトウェアまたは前記ソフトウェアが組み込まれたハードウェアでもよい。類似重複度記憶部１２４ａおよび類似配列記憶部１２４ｂは、例えば、図７に示すように、ハードウェアである記憶手段１２に組み込まれてもよい。

[0105] 候補配列記憶部１２３は、類似重複度計算手段１３３ａと電気的に接続されており、類似重複度記憶部１２４ａは、類似重複度計算手段１３３ａおよび類似配列選択手段１３３ｂと電気的に接続されており、類似配列記憶部１２４ｂは、類似配列選択手段１３３ｂおよび出力手段１４と、それぞれ電気的に接続されている。また、候補配列選択手段１３２は、類似重複度計算手段１３３ａと、類似重複度計算手段１３３ａは、類似配列選択手段１３３ｂと、類似配列選択手段１３３ｂは、出力手段１４と、それぞれ電気的に接続されてもよい。

[0106] つぎに、図８および図９のフローチャートを参照し、本実施形態の類似選択方法を説明する。本実施形態の類似選択方法は、Ａ１ステップ（配列入力）、Ａ２ステップ（類似度計算）、Ａ３ステップ（候補配列選択）およびＡ４ステップ（類似配列選択）を含む。本実施形態において、Ａ４ステップは、Ａ４ａステップ（類似重複度計算）と、Ａ４ｂステップ（類似重複度の計算結果に基づく類似配列選択）を含む。図８および図９において、図５および図６と同じステップには同じ符号を付している。

[0107] 前記Ａ１ステップ、前記Ａ２ステップおよび前記Ａ３ステップは、前記実施形態２と同様に行うことができる。本実施形態において、入力する配列の情報項目は、例えば、配列を構成する要素の順序の他に、前記各配列の重複度を含む。

[0108] （Ａ４）類似配列選択

前記Ａ３ステップで選択された候補配列群から、新しい比較元候補配列を

セット（Ａ４１’）し、その重複度が０か否かを判断する（Ａ４２’）。ＮＯの場合、つまり、重複度０の場合（初期重複度が０または再設定重複度０）、再度、新しい比較元候補配列をセットする（Ａ４１’）。他方、ＹＥＳの場合、つまり、重複度が０でない場合（初期重複度≧１）、前記比較元候補配列の重複度をセットする（Ａ４３’）。そして、新しい比較先候補配列をセット（Ａ４４’）し、前記比較先候補配列が、前記比較元候補配列に類似するか否かを判断する（Ａ４５’）。ＹＥＳの場合、つまり、前記比較先候補配列が前記比較元候補配列に類似する場合、前記比較元候補配列の類似度と前記比較先候補配列の類似度とを合計し、その合計値を類似重複度とする（Ａ４６’）。この類似重複度は、前記比較元候補配列の類似重複度という。他方、ＮＯの場合、つまり、前記比較先候補配列が、前記比較元候補配列に類似しない場合、未比較の比較先候補配列の有無を確認する（Ａ４７’）。そして、ＹＥＳの場合、つまり、未比較の比較先候補配列がある場合、Ａ４４’ステップから同様の処理を行う。そして、ＮＯの場合、つまり、未比較の比較先候補配列がない場合、さらに、未比較の比較元候補配列の有無を確認する（Ａ４８’）。ＹＥＳの場合、つまり、未比較の比較元候補配列がある場合、Ａ４１’ステップから同様の処理を行う。ＮＯの場合、つまり、未比較の比較元候補配列がない場合、最も大きい類似重複度の候補配列以外であって、類似重複度が０でない候補配列について、類似重複度をリセット、つまり０に再設定する（Ａ４９’）。さらに、最も大きい類似重複度の候補配列およびそれに類似する候補配列について、重複度を０に再設定する（Ａ４１０’）。つぎに、重複度が０でない候補配列の有無を確認する（Ａ４１１’）。ＹＥＳの場合、つまり、重合度が０でない候補配列（初期重複度≧１）がある場合、これを新しい比較元候補配列とし、Ａ４１’ステップから同様の処理を行う。ＮＯの場合、つまり、重複度が０でない候補配列が存在しない場合、類似重複度が０でない候補配列とそれに類似する候補配列とを類似配列群とし、類似配列群の一覧を出力する（Ａ４１２’）。出力する情報項目は例えば、前記類似配列群に含まれる各配列ならびに類似重複度

等があげられる。

[0109] 前記Ａ４ステップについて、さらなる具体例として、候補配列群に含まれる異なる配列が５種類（Ｓｅｑ１、Ｓｅｑ２、Ｓｅｑ３、Ｓｅｑ４、Ｓｅｑ５）であり、それぞれの重複度（すなわち、出現数）が、｛５、４、３、２、１｝である場合を例にあげて説明する。

[0110] まず、下記表１に、候補配列の種類とその重複度を示す。

[0111] [表1]

		比較先					重複度	類似重複度
		Seq1	Seq2	Seq3	Seq4	Seq5		
比較元	Seq1						5	—
	Seq2						4	—
	Seq3						3	—
	Seq4						2	—
	Seq5						1	—

[0112] つぎに、それぞれの配列間における類似を判断する。下記表２において、類似の関係にあるものを、網掛けで示す。

[0113] [表2]

		比較先					重複度	類似重複度
		Seq1	Seq2	Seq3	Seq4	Seq5		
比較元	Seq1	■			■	■	5	—
	Seq2		■	■	■		4	—
	Seq3		■	■			3	—
	Seq4	■	■		■		2	—
	Seq5	■				■	1	—

[0114] そして、それぞれの比較元候補配列について、前記比較元候補配列の初期重複度とそれに類似する前記比較先候補配列の初期重複度とを合計し、この合計値を比較元候補配列の類似重複度とする。下記表３に、類似重複度を示す。そして、前記比較元候補配列のうち、最も大きい類似重複度を示す比較元候補配列を選択し、前記比較元候補配列とそれに類似する比較先候補配列とを、類似配列群とする。下記表３において、最も大きい類似重複度１１を

示すＳｅｑ４ならびにそれに類似するＳｅｑ１およびＳｅｑ２が、同じ類似配列群となる。

[0115] [表3]

		比較先					重複度	類似重複度
		Seq1	Seq2	Seq3	Seq4	Seq5		
比較元	Seq1	5			2	1	5	8
	Seq2		4	3	2		4	9
	Seq3		4	3			3	7
	Seq4	5	4		2		2	11
	Seq5	5				1	1	6

[0116] 続いて、最も大きい類似重複度を示す比較元候補配列以外であって、類似重複度が０ではない候補配列について、類似重複度をリセットし、最も大きい類似重複度を示す比較元候補配列の初期重複度とそれに類似する比較先候補配列の初期重複度とを、０に再設定する（再設定重複度０）。下記表４において、最も大きい類似重複数１１を示すＳｅｑ４以外の配列について、類似重複度をリセットし、且つ、Ｓｅｑ４と、それに類似するＳｅｑ１およびＳｅｑ２の初期重複度を、０に再設定する（再設定重複度０）。

[0117] [表4]

		比較先					重複度	類似重複度
		Seq1	Seq2	Seq3	Seq4	Seq5		
比較元	Seq1	0			0	0	5→0	—
	Seq2		0	0	0		4→0	—
	Seq3		0	3			3	—
	Seq4	0	0		0		2→0	11
	Seq5	0				1	1	—

[0118] そして、重複度が０以外（初期重複度≧１）の比較元候補配列について、同様にして、類似重複度の計算、最も大きい類似重複度に基づく類似候補群の選択を行う。類似候補群の選択は、全ての候補配列の初期重複度が０に再設定されるまで、繰り返し行うことが好ましい。下記表５において、重複度が０ではない候補配列のうち、最も大きい類似重複度３を示すＳｅｑ３を、

類似配列群とする。

[0119] [表5]

		比較先					重複度	類似重複度
		Seq1	Seq2	Seq3	Seq4	Seq5		
比較元	Seq1	0			0	0	0	—
	Seq2		0	0	0		0	—
	Seq3		0	3			3	3
	Seq4	0	0		0		0	11
	Seq5	0				1	1	1

[0120] なお、配列間の類似について、一方の配列を比較元候補配列とし、他方の配列を比較先候補配列とするのと、前記一方の配列を比較先候補配列とし、前記他方の配列を比較元候補配列とするのは、実質的に同じといえる。そこで、前記類似配列群の選択をより促進できるため、例えば、比較元候補配列と比較先候補配列との組合せから、すでに実行した組合せを除外することが好ましい。この場合、例えば、下記表６のように、異なる配列間の組合せを半分にできる（セル数の半減）。

[0121] [表6]

		比較先					重複度	類似重複度
		Seq1	Seq2	Seq3	Seq4	Seq5		
比較元	Seq1						5	
	Seq2						4	
	Seq3						3	
	Seq4						2	
	Seq5						1	

[0122] これらの処理を繰り返すことによって、候補配列群を類似配列群に分類することができる。

[0123] ＜目的の類似配列群の濃縮を判定する装置＞

本発明の濃縮の判定装置は、前述のように、下記（Ｘ）および（Ｙ）手段を備え、前記（Ｘ）手段が、前記本発明の類似選択装置であることを特徴とする、目的の類似配列情報群の濃縮の判定装置である。

（Ｘ）配列情報群から、目的配列情報とそれに類似する配列情報とを目的の類似配列情報群として選択する工程を実行する手段

（Ｙ）前記類似配列情報群における前記目的配列情報と前記類似する配列情報との重複度の合計から、前記類似配列情報群の濃縮を判定する工程を実行する手段

[0124] 本発明の判定装置において、前記（Ｘ）手段は、前記本発明の類似選択装置であればよく、前記本発明の類似選択装置の記載を援用できる。

[0125] 本発明の濃縮の判定装置は、前記（Ｘ）手段が、比較元となる類似配列情報群と、比較先となる類似配列情報群を、それぞれ選択する工程を実行し、前記（Ｙ）手段が、下記（Ｙ１）および（Ｙ２）工程を実行する手段であることが好ましい。

（Ｙ１）前記比較元の類似配列情報群における目的の配列情報とそれに類似する配列情報との重複度の合計と、前記比較先の類似配列情報群における目的の配列情報とそれに類似する配列情報との重複度の合計とを、比較する工程

（Ｙ２）前記比較元の類似配列情報群における前記重複度の合計が、前記比較先の類似配列情報群における前記重複度の合計よりも大きい場合に、前記比較元の類似配列情報群が、前記比較先の配列情報群よりも、濃縮されていると判断する工程

[0126] 本発明において、濃縮の判定は、例えば、同じ配列情報群に含まれる異なる配列情報について、前記配列情報の間における濃縮度合いの違いを比較することにより行ってもよい。この場合、例えば、前記比較元の類似配列情報群と前記比較先の類似配列情報群とが、同じ配列群から選択された類似配列情報群であり、前記比較元の類似配列情報群の前記目的の配列情報と、前記比較先の類似配列情報群の前記目的の配列情報とが、異なる配列情報である。これによって、例えば、同じ配列情報群から、相対的に濃縮度合いの高い配列情報およびその類似配列情報を選択することが可能となる。具体例としては、例えば、アプタマーの調製において、特定のラウンドのライブラリー

に含まれる複数の類似配列情報群から、相対的に濃縮度の高い類似配列情報群の選択、つまり濃縮度が高いアプタマー類似配列群の選択を行うことができる。

[0127] また、前記濃縮の判定は、例えば、異なる配列情報群に含まれる同じ配列情報について、前記配列情報群の間における濃縮度合いの違いを比較することにより行ってもよい。この場合、例えば、前記比較元の類似配列情報群と前記比較先の類似配列情報群とが、異なる配列群から選択された類似配列情報群であり、前記比較元の類似配列情報群の前記目的の配列情報と、前記比較先の類似配列情報群の前記目的の配列情報とが、同じ配列情報である。これによって、例えば、特定の配列情報の類似配列情報群について、相対的に濃縮度合いの高い配列情報群を選択することができる。具体例としては、例えば、アプタマーの調製において、各ラウンドのライブラリーのうち、特定のアプタマー類似配列群の濃縮度が相対的に高いライブラリーを選択することができる。

[0128] 本発明の濃縮の判定方法は、下記（Ｘ）および（Ｙ）工程を含み、前記（Ｘ）工程が、前記本発明の類似選択方法を含むことを特徴とする、類似配列情報群の濃縮の判定方法である。本発明の濃縮の判定方法は、特に示さない限り、前記本発明の濃縮の判定装置における記載を援用できる。

（Ｘ）配列情報群から、目的の配列情報とそれに類似する配列情報とを判定対象の類似配列情報群として選択する工程

（Ｙ）前記類似配列情報群における前記目的の配列情報と前記類似する配列情報との重複度の合計から、前記類似配列情報群の濃縮を判定する工程

[0129] 本発明の濃縮の判定方法は、前記（Ｘ）工程が、比較元となる類似配列情報群と、比較先となる類似配列情報群を、それぞれ選択する工程であり、

前記（Ｙ）工程が、下記（Ｙ１）および（Ｙ２）工程を含むことが好ましい。

（Ｙ１）前記比較元の類似配列情報群における目的の配列情報とそれに類似する配列情報との重複度の合計と、前記比較先の類似配列情報群における目

的の配列情報とそれに類似する配列情報との重複度の合計とを、比較する工程

（Ｙ２）前記比較元の類似配列情報群における前記重複度の合計が、前記比較先の類似配列情報群における前記重複度の合計よりも大きい場合に、前記比較元の類似配列情報群が、前記比較先の配列情報群よりも、濃縮されていると判断する工程

[0130] 本発明の濃縮の判定方法は、前記比較元の類似配列情報群と前記比較先の類似配列情報群とが、同じ配列群から選択された類似配列情報群であり、前記比較元の類似配列情報群の前記目的の配列情報と、前記比較先の類似配列情報群の前記目的の配列情報とが、異なる配列情報であってもよい。

[0131] 本発明の濃縮の判定方法は、前記比較元の類似配列情報群と前記比較先の類似配列情報群とが、異なる配列群から選択された類似配列情報群であり、前記比較元の類似配列情報群の前記目的の配列情報と、前記比較先の類似配列情報群の前記目的の配列情報とが、同じ配列情報であってもよい。

[0132] 本発明の用途は、特に制限されないが、例えば、アプタマーの調製における濃縮の判定に適用することが好ましい。本発明によれば、前述のように、例えば、同じライブラリー内における異なるアプタマー類似配列情報群の濃縮度合いの比較、または、異なるライブラリー内における同じアプタマー類似配列情報群の濃縮度合いの比較が可能である。

実施例

[0133] つぎに、本発明の実施例について説明する。ただし、本発明は、下記の実施例により制限されない。

[0134] ［実施例１］

本実施例では、低分子化合物をターゲットとするライブラリーについて、本発明の類似選択方法により、類似配列群の分類を行った。

[0135] 配列群として、４０塩基長の８５，８００個の核酸配列群を使用した。仮想配列群の条件、許容できるミスマッチの塩基数および許容条件を下記表７に示す。

[0136] [表7]

	仮想配列		許容ミスマッチ数	許容条件	計算時間
	塩基長（Ｎ）	配列数	（Ｍ）	（Ｎ×Ｍ）	（時間）
比較例	－	－	－	－	８３
実施例	１	４	５	５	１７
	２	4^2	５	１０	９
	３	4^3	５	１５	１
	４	4^4	５	２０	２

[0137] 実施例は、前記条件に従い、前記表６に示すセル数の半減を行って、候補配列群の選択、類似配列群の選択を行った。これらの計算にかかった時間を前記表７にあわせて示す。なお、比較例は、前記配列群の全ての核酸配列について、アラインメントによる類似の判断を行い、類似配列群を選択した。その結果、実施例によれば、比較例よりも格段に短い計算時間で類似配列群の選択を行うことができた。

[0138] 以上、実施形態を参照して本願発明を説明したが、本願発明は、上記実施形態に限定されるものではない。本願発明の構成や詳細には、本願発明のスコープ内で当業者が理解しうる様々な変更をすることができる。

[0139] この出願は、２０１３年２月１５日に出願された日本出願特願２０１３－０２７８５１を基礎とする優先権を主張し、その開示の全てをここに取り込む。

産業上の利用可能性

[0140] 本発明によれば、配列情報間の類似を判断するにあたって、まず、類似を判断するための候補配列群が選択される。このため、例えば、全ての配列情報間の類似を確認する従来の方法とは異なり、簡便に効率よく類似の判断を行うことができる。このため、例えば、アプタマーの濃縮の判定等についても、労力、時間およびコストの軽減が可能となる。

符号の説明

[0141] １０ 候補選択装置

２０、３０ 類似選択装置

１１	入力手段
１２	記憶手段
１２１	配列記憶部
１２２	類似度記憶部
１２３	候補配列記憶部
１２４	類似配列記憶部
１２４ａ	類似重複度記憶部
１２４ｂ	類似配列記憶部
１３	データ処理手段
１３１	類似度計算手段
１３２	候補配列選択手段
１３３	類似配列選択手段
１３３ａ	類似重複度計算手段
１３３ｂ	類似配列選択手段
１４	出力手段

請求の範囲

[請求項1] 下記（ａ）、（ｂ）、（ｃ）および（ｄ）手段を備えることを特徴とする、配列情報群から、配列情報間の類似の判断候補となる候補配列情報群を選択する候補選択装置。

（ａ）配列情報群の各配列情報について、仮想配列情報群の各仮想配列情報の頻度をカウントする工程を実行する手段

（ｂ）前記配列情報群から、比較元となる配列情報と比較先となる配列情報とを選択する工程を実行する手段

（ｃ）前記比較元配列情報の前記各仮想配列情報の頻度と、前記比較先配列情報の前記各仮想配列情報の頻度との相違を、前記比較元配列情報に対する前記比較先配列情報の類似度として計算する工程を実行する手段

（ｄ）前記比較元配列情報に対する前記比較先配列情報の類似度が、前記仮想配列情報群に設定した類似度の許容条件を満たす場合、前記比較元配列情報および前記比較先配列情報を、配列情報間の類似を判断する候補配列情報群として選択する工程を実行する手段

[請求項2] 前記仮想配列情報群が、配列情報を構成する要素から構築された仮想配列情報の群である、請求項１記載の候補選択装置。

[請求項3] 前記（ｃ）手段が、下記（ｃ１）および（ｃ２）工程を実行する手段である、請求項１または２記載の候補選択装置。

（ｃ１）前記仮想配列情報ごとに、前記比較元配列情報における頻度と前記比較先配列情報における頻度との差を求める工程

（ｃ２）前記各仮想配列情報の頻度の差のうち、正数の差のみの総和の絶対値または負数の差のみの総和の絶対値を求め、前記絶対値を、前記比較元配列情報に対する前記比較先配列情報の類似度として計算する工程

[請求項4] 前記類似度の許容条件が、２つの配列情報を対比した場合に許容できるミスマッチの個数に基づき設定された条件である、請求項１から３

のいずれか一項に記載の候補選択装置。

[請求項5] 前記配列情報が、塩基配列であり、前記配列情報を構成する要素が、Ａ、Ｇ、Ｃ、ＴおよびＵの塩基である、請求項１から４のいずれか一項に記載の候補選択装置。

[請求項6] 前記仮想配列情報の塩基長が、１～１０塩基長である、請求項５記載の候補選択装置。

[請求項7] 前記仮想配列情報群の各仮想配列情報が、すべて同じ塩基長である、請求項５または６記載の候補選択装置。

[請求項8] 前記類似度の許容条件が、２つの配列情報を対比した場合に許容できるミスマッチの塩基数に基づき設定された条件である、請求項３から７のいずれか一項に記載の候補選択装置。

[請求項9] 前記類似度の許容条件が、２つの配列情報を対比した場合に許容できるミスマッチの塩基数（Ｍ）に前記仮想配列情報の塩基長（Ｎ）を乗じた値である、請求項５から８のいずれか一項に記載の候補選択装置。

[請求項10] さらに、下記（ｅ）手段を有する、請求項１から９のいずれか一項に記載の候補選択装置。

（ｅ）前記（ｂ）、（ｃ）および（ｄ）手段による各工程の反復を実行する手段

[請求項11] 前記（ｂ）手段は、前記工程の実行ごとに、前記配列情報群から、異なる配列情報を前記比較元配列情報として選択する、請求項１０記載の候補選択装置。

[請求項12] 下記（Ａ）および（Ｂ）手段を備え、

前記（Ａ）手段が、請求項１から１１のいずれか一項に記載の候補選択装置であることを特徴とする、配列情報群から、相互に類似する類似配列情報群を選択する類似選択装置。

（Ａ）配列情報群から、配列情報間の類似の判断候補となる候補配列情報群を選択する工程を実行する手段

（Ｂ）前記候補配列情報群の各候補配列情報を相互に対比し、同一および類似する配列情報を類似配列情報群（Ｇ３）として選択する工程を実行する手段

[請求項13] 前記（Ｂ）手段が、下記（Ｂ１）、（Ｂ２）、（Ｂ３）、（Ｂ４）および（Ｂ５）工程を実行する手段である、請求項１２記載の類似選択装置

（Ｂ１）前記候補配列情報群から、比較元となる候補配列情報と比較先となる候補配列情報とを選択する工程

（Ｂ２）前記比較元候補配列情報に対する前記比較先候補配列情報の類似の有無を決定する工程

（Ｂ３）前記比較元候補配列情報の重複度と、前記比較元候補配列情報に類似する前記比較先候補配列情報の重複度とを合計し、得られた合計値を、前記比較元候補配列情報の類似重複度とする工程

（Ｂ４）前記候補配列情報群から、異なる候補配列情報を、新たな比較元となる候補配列情報として選択し、前記（Ｂ１）、（Ｂ２）および（Ｂ３）工程を反復する工程

（Ｂ５）前記候補配列情報のうち、最も大きな類似重複度を示した候補配列情報および前記候補配列情報に類似する候補配列情報を、類似配列情報群（Ｇ３）として選択する工程

[請求項14] 前記（Ｂ）手段が、さらに、下記（Ｂ６）、（Ｂ７）および（Ｂ８）工程を実行する手段である、請求項１３記載の類似選択装置。

（Ｂ６）前記候補配列情報のうち、最も大きな類似重複度を示した候補配列情報の重複度および前記候補配列情報に類似する候補配列情報の重複度を０に再設定する工程

（Ｂ７）重複度が０以外である他の候補配列情報について、類似重複度を再算出する工程

（Ｂ８）前記他の候補配列情報のうち、最も大きな類似重複度を示した候補配列情報および前記候補配列情報に類似する候補配列情報を、

類似配列情報群として再選択する工程

[請求項15] 前記（Ｂ）手段が、さらに、下記（Ｂ９）の工程を実行する手段である、請求項１４記載の類似選択装置。

（Ｂ９）前記他の候補配列情報のうち、最も大きな類似重複度を示した候補配列情報および前記候補配列情報に類似する候補配列情報の重複度を０に再設定し、前記（Ｂ７）および（Ｂ８）工程を反復する工程

[請求項16] 前記（Ｂ）手段が、前記（Ｂ１）工程における前記比較元補配列情報と前記比較先候補配列情報との組合せとして、すでに実行した組合せの除外を実行する、請求項１３から１５のいずれか一項に記載の類似選択装置。

[請求項17] 下記（Ｘ）および（Ｙ）手段を備え、前記（Ｘ）手段が、請求項１２から１６のいずれか一項に記載の類似選択装置であることを特徴とする、目的の類似配列情報群の濃縮の判定装置。

（Ｘ）配列情報群から、目的配列情報とそれに類似する配列情報とを目的の類似配列情報群として選択する工程を実行する手段

（Ｙ）前記類似配列情報群における前記目的配列情報と前記類似する配列情報との重複度の合計から、前記類似配列情報群の濃縮を判定する工程を実行する手段

[請求項18] 前記（Ｘ）手段が、比較元となる類似配列情報群と、比較先となる類似配列情報群を、それぞれ選択する工程を実行し、

前記（Ｙ）手段が、下記（Ｙ１）および（Ｙ２）工程を実行する手段である、請求項１７記載の判定装置。

（Ｙ１）前記比較元の類似配列情報群における目的の配列情報とそれに類似する配列情報との重複度の合計と、前記比較先の類似配列情報群における目的の配列情報とそれに類似する配列情報との重複度の合計とを、比較する工程

（Ｙ２）前記比較元の類似配列情報群における前記重複度の合計が、

前記比較先の類似配列情報群における前記重複度の合計よりも大きい場合に、前記比較元の類似配列情報群が、前記比較先の配列情報群よりも、濃縮されていると判断する工程

[請求項19] 前記比較元の類似配列情報群と前記比較先の類似配列情報群とが、同じ配列群から選択された類似配列情報群であり、

前記比較元の類似配列情報群の前記目的の配列情報と、前記比較先の類似配列情報群の前記目的の配列情報とが、異なる配列情報である、請求項１８記載の判定装置。

[請求項20] 前記比較元の類似配列情報群と前記比較先の類似配列情報群とが、異なる配列群から選択された類似配列情報群であり、

前記比較元の類似配列情報群の前記目的の配列情報と、前記比較先の類似配列情報群の前記目的の配列情報とが、同じ配列情報である、請求項１８記載の判定装置。

[請求項21] 下記（ａ）、（ｂ）、（ｃ）および（ｄ）工程を含むことを特徴とする、配列情報群から、配列情報間の類似の判断候補となる候補配列情報群を選択する候補選択方法。

（ａ）配列情報群の各配列情報について、仮想配列情報群の各仮想配列情報の頻度をカウントする工程

（ｂ）前記配列情報群から、比較元となる配列情報と比較先となる配列情報とを選択する工程

（ｃ）前記比較元配列情報の前記各仮想配列情報の頻度と、前記比較先配列情報の前記各仮想配列情報の頻度との相違を、前記比較元配列情報に対する前記比較先配列情報の類似度として計算する工程

（ｄ）前記比較元配列情報に対する前記比較先配列情報の類似度が、前記仮想配列情報群に設定した類似度の許容条件を満たす場合、前記比較元配列情報および前記比較先配列情報を、配列情報間の類似を判断する候補配列情報群として選択する工程

[請求項22] 前記仮想配列情報群が、配列情報を構成する要素から構築された仮想

配列情報の群である、請求項２１記載の候補選択方法。

[請求項23] 前記（ｃ）工程が、下記（ｃ１）および（ｃ２）工程を含む、請求項２１または２２記載の候補選択方法。

（ｃ１）前記仮想配列情報ごとに、前記比較元配列情報における頻度と前記比較先配列情報における頻度との差を求める工程

（ｃ２）前記各仮想配列情報の頻度の差のうち、正数の差のみの総和の絶対値または負数の差のみの総和の絶対値を求め、前記絶対値を、前記比較元配列情報に対する前記比較先配列情報の類似度として計算する工程

[請求項24] 前記類似度の許容条件が、２つの配列情報を対比した場合に許容できるミスマッチの個数に基づき設定された条件である、請求項２１から２３のいずれか一項に記載の候補選択方法。

[請求項25] 前記配列情報が、塩基配列であり、前記配列情報を構成する要素が、Ａ、Ｇ、Ｃ、ＴおよびＵの塩基である、請求項２１から２４のいずれか一項に記載の候補選択方法。

[請求項26] 前記仮想配列情報の塩基長が、１～１０塩基長である、請求項２５記載の候補選択方法。

[請求項27] 前記仮想配列情報群の各仮想配列情報が、すべて同じ塩基長である、請求項２５または２６記載の候補選択方法。

[請求項28] 前記類似度の許容条件が、２つの配列情報を対比した場合に許容できるミスマッチの塩基数に基づき設定された条件である、請求項２３から２７のいずれか一項に記載の候補選択方法。

[請求項29] 前記類似度の許容条件が、２つの配列情報を対比した場合に許容できるミスマッチの塩基数（Ｍ）に前記仮想配列情報の塩基長（Ｎ）を乗じた値である、請求項２５から２８のいずれか一項に記載の候補選択方法。

[請求項30] さらに、下記（ｅ）工程を含む、請求項２１から２９のいずれか一項に記載の候補選択方法。

（ｅ）前記（ｂ）、（ｃ）および（ｄ）工程を反復する工程

[請求項31] 前記（ｂ）工程において、前記工程の実行ごとに、前記配列情報群から、異なる配列情報を前記比較元配列情報として選択する、請求項３０記載の候補選択方法。

[請求項32] 下記（Ａ）および（Ｂ）工程を含み、

前記（Ａ）工程が、請求項２１から３１のいずれか一項に記載の候補選択方法を含むことを特徴とする、配列情報群から、相互に類似する類似配列情報群を選択する類似選択方法。

（Ａ）配列情報群から、配列情報間の類似の判断候補となる候補配列情報群を選択する工程

（Ｂ）前記候補配列情報群の各候補配列情報を相互に対比し、同一および類似する配列情報を類似配列情報群（Ｇ３）として選択する工程

[請求項33] 前記（Ｂ）工程が、下記（Ｂ１）、（Ｂ２）、（Ｂ３）、（Ｂ４）および（Ｂ５）工程を含む、請求項３２記載の類似選択方法

（Ｂ１）前記候補配列情報群から、比較元となる候補配列情報と比較先となる候補配列情報とを選択する工程

（Ｂ２）前記比較元候補配列情報に対する前記比較先候補配列情報の類似の有無を決定する工程

（Ｂ３）前記比較元候補配列情報の重複度と、前記比較元候補配列情報に類似する前記比較先候補配列情報の重複度とを合計し、得られた合計値を、前記比較元候補配列情報の類似重複度とする工程

（Ｂ４）前記候補配列情報群から、異なる候補配列情報を、新たな比較元となる候補配列情報として選択し、前記（Ｂ１）、（Ｂ２）および（Ｂ３）工程を反復する工程

（Ｂ５）前記候補配列情報のうち、最も大きな類似重複度を示した候補配列情報および前記候補配列情報に類似する候補配列情報を、類似配列情報群（Ｇ３）として選択する工程

[請求項34] 前記（Ｂ）工程が、さらに、下記（Ｂ６）、（Ｂ７）および（Ｂ８）

工程を含む、請求項３３記載の類似選択方法。

（Ｂ６）前記候補配列情報のうち、最も大きな類似重複度を示した候補配列情報の重複度および前記候補配列情報に類似する候補配列情報の重複度を０に再設定する工程

（Ｂ７）重複度が０以外である他の候補配列情報について、類似重複度を再算出する工程

（Ｂ８）前記他の候補配列情報のうち、最も大きな類似重複度を示した候補配列情報および前記候補配列情報に類似する候補配列情報を、類似配列情報群として再選択する工程

[請求項35] 前記（Ｂ）工程が、さらに、下記（Ｂ９）工程を含む、請求項３４記載の類似選択方法。

（Ｂ９）前記他の候補配列情報のうち、最も大きな類似重複度を示した候補配列情報および前記候補配列情報に類似する候補配列情報の重複度を０に再設定し、前記（Ｂ７）および（Ｂ８）工程を反復する工程

[請求項36] 前記（Ｂ）工程において、前記（Ｂ１）工程における前記比較元補配列情報と前記比較先候補配列情報との組合せとして、すでに実行した組合せを除外する、請求項３３から３５のいずれか一項に記載の類似選択方法。

[請求項37] 下記（Ｘ）および（Ｙ）工程を含み、前記（Ｘ）工程が、請求項３２から３６のいずれか一項に記載の類似選択方法を含むことを特徴とする、類似配列情報群の濃縮の判定方法。

（Ｘ）配列情報群から、目的の配列情報とそれに類似する配列情報とを判定対象の類似配列情報群として選択する工程

（Ｙ）前記類似配列情報群における前記目的の配列情報と前記類似する配列情報との重複度の合計から、前記類似配列情報群の濃縮を判定する工程

[請求項38] 前記（Ｘ）工程が、比較元となる類似配列情報群と、比較先となる類

似配列情報群を、それぞれ選択する工程であり、

前記（Ｙ）工程が、下記（Ｙ１）および（Ｙ２）工程を含む、請求項３７記載の判定方法。

（Ｙ１）前記比較元の類似配列情報群における目的の配列情報とそれに類似する配列情報との重複度の合計と、前記比較先の類似配列情報群における目的の配列情報とそれに類似する配列情報との重複度の合計とを、比較する工程

（Ｙ２）前記比較元の類似配列情報群における前記重複度の合計が、前記比較先の類似配列情報群における前記重複度の合計よりも大きい場合に、前記比較元の類似配列情報群が、前記比較先の配列情報群よりも、濃縮されていると判断する工程

[請求項39] 前記比較元の類似配列情報群と前記比較先の類似配列情報群とが、同じ配列群から選択された類似配列情報群であり、

前記比較元の類似配列情報群の前記目的の配列情報と、前記比較先の類似配列情報群の前記目的の配列情報とが、異なる配列情報である、請求項３８記載の判定方法。

[請求項40] 前記比較元の類似配列情報群と前記比較先の類似配列情報群とが、異なる配列群から選択された類似配列情報群であり、

前記比較元の類似配列情報群の前記目的の配列情報と、前記比較先の類似配列情報群の前記目的の配列情報とが、同じ配列情報である、請求項３８記載の判定方法。

[請求項41] 請求項２１から３１のいずれか一項に記載の候補選択方法を、コンピュータ上で実行可能なことを特徴とするプログラム。

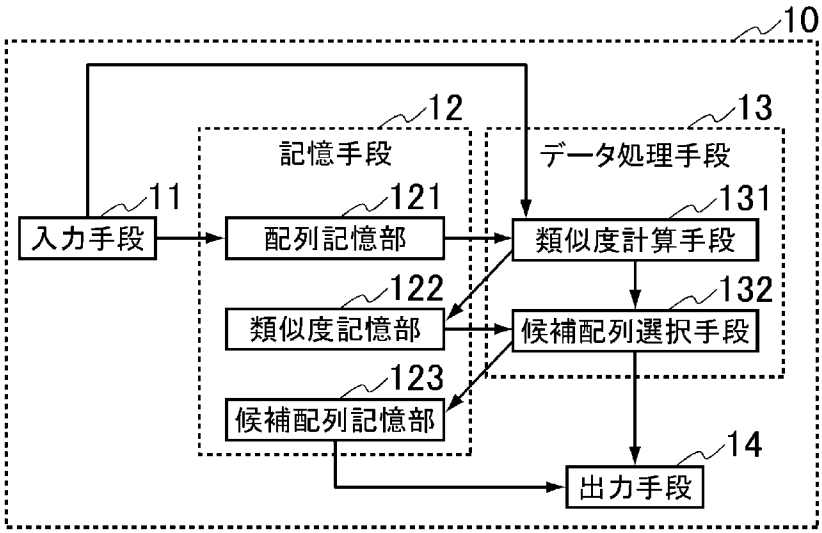
[請求項42] 請求項３２から３６のいずれか一項に記載の類似選択方法を、コンピュータ上で実行可能なことを特徴とするプログラム。

[請求項43] 請求項３７から４０のいずれか一項に記載の判定方法を、コンピュータ上で実行可能なことを特徴とするプログラム。

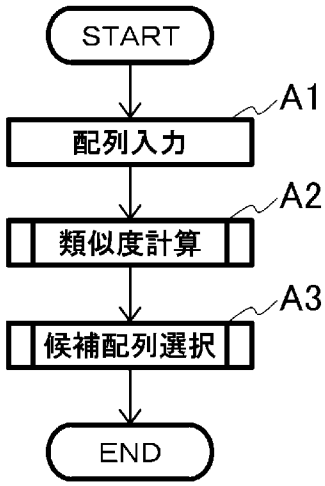
[請求項44] 請求項４１から４３のいずれか一項に記載のプログラムを記録してい

ることを特徴とする記録媒体。

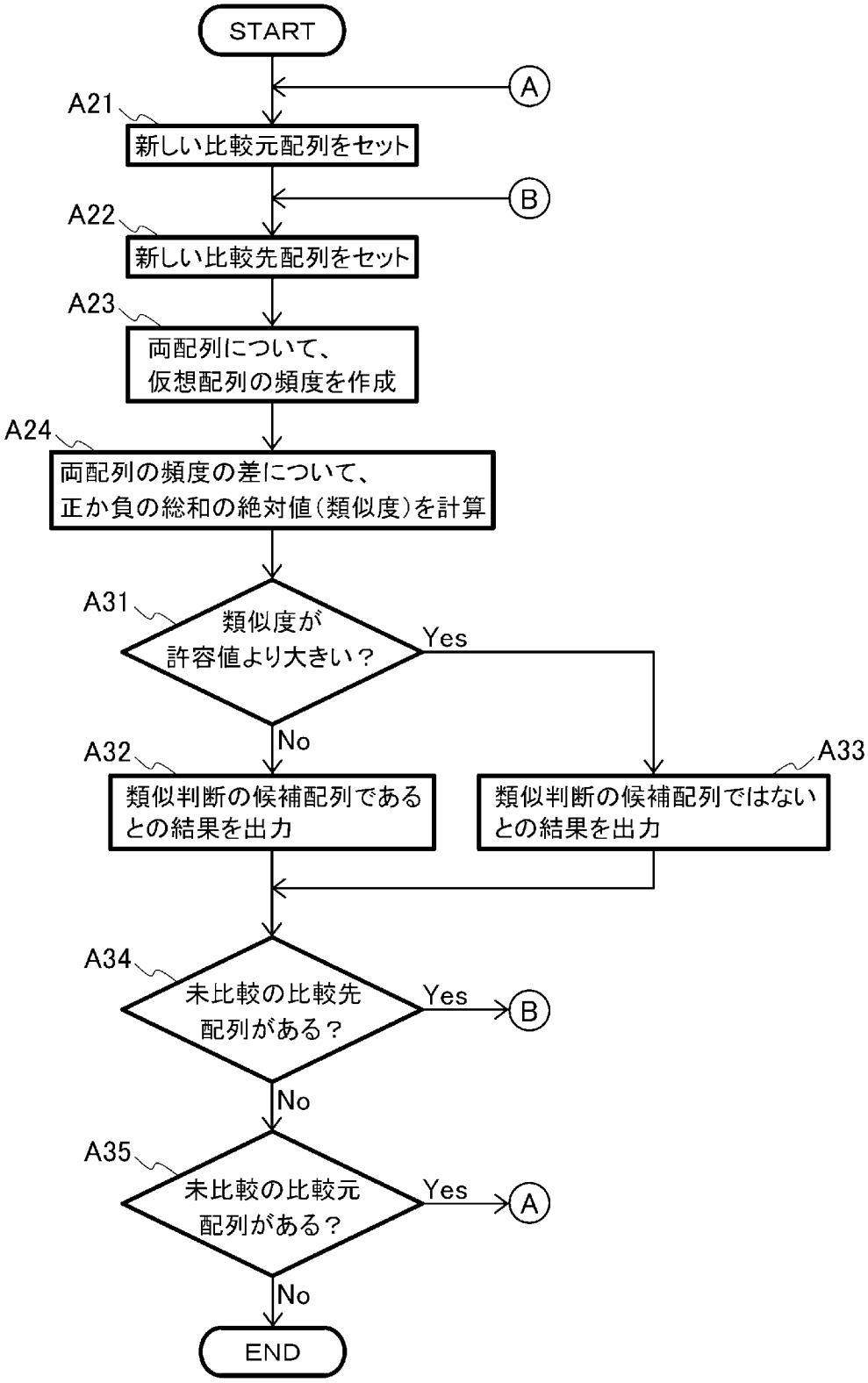
[図1]


10
12
記憶手段
13
データ処理手段
11
入力手段
121
配列記憶部
131
類似度計算手段
122
類似度記憶部
132
候補配列選択手段
123
候補配列記憶部
14
出力手段

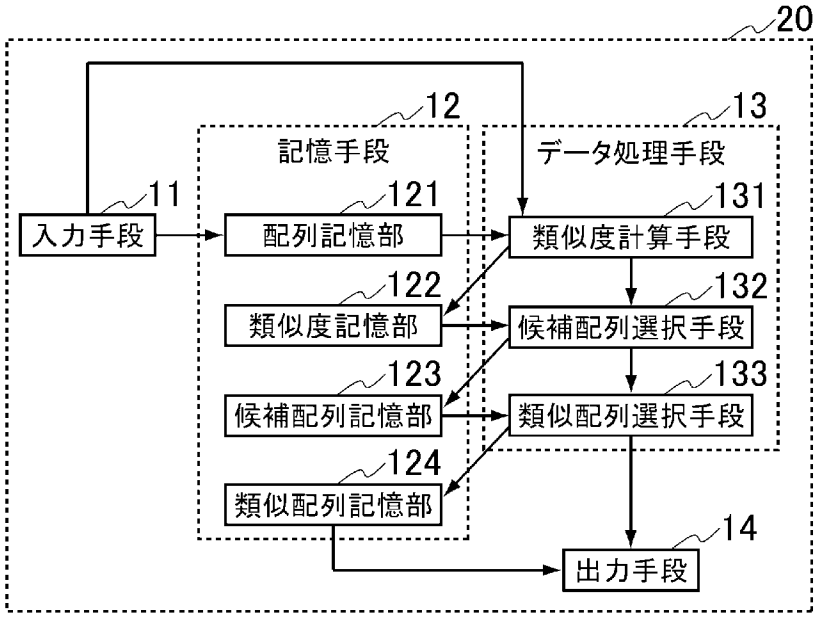
[図2]


START
A1
配列入力
A2
類似度計算
A3
候補配列選択
END

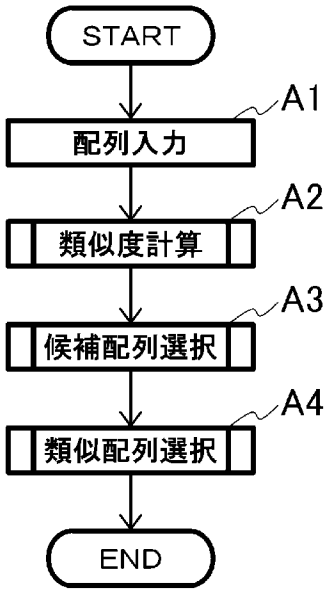
[図3]


START
A
A21
新しい比較元配列をセット
B
A22
新しい比較先配列をセット
A23
両配列について、
仮想配列の頻度を作成
A24
両配列の頻度の差について、
正か負の総和の絶対値（類似度）を計算
A31
類似度が
許容値より大きい？
Yes
No
A32
類似判断の候補配列である
との結果を出力
A33
類似判断の候補配列ではない
との結果を出力
A34
未比較の比較先
配列がある？
Yes
B
No
A35
未比較の比較元
配列がある？
Yes
A
No
END

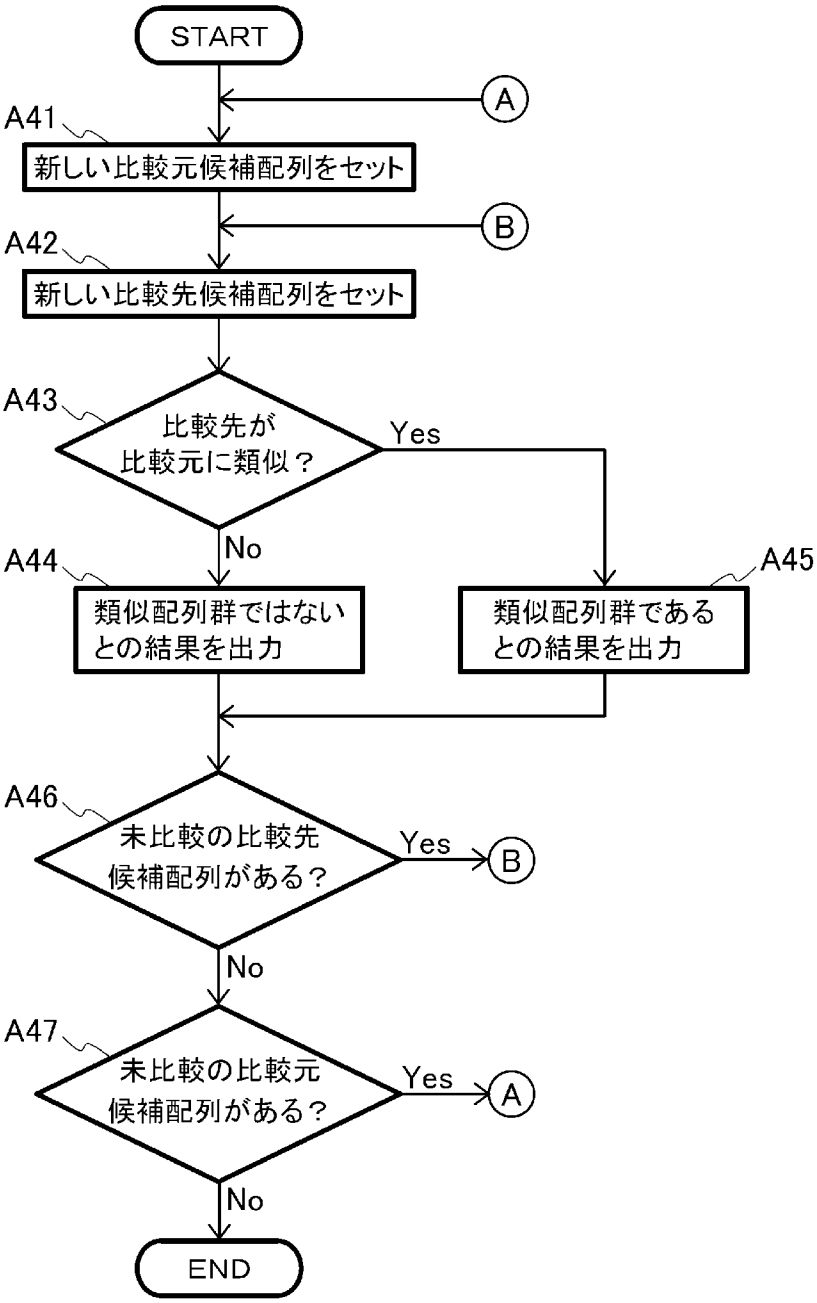
[図4]


20
12
記憶手段
13
データ処理手段
11
入力手段
121
配列記憶部
131
類似度計算手段
122
類似度記憶部
132
候補配列選択手段
123
候補配列記憶部
133
類似配列選択手段
124
類似配列記憶部
14
出力手段

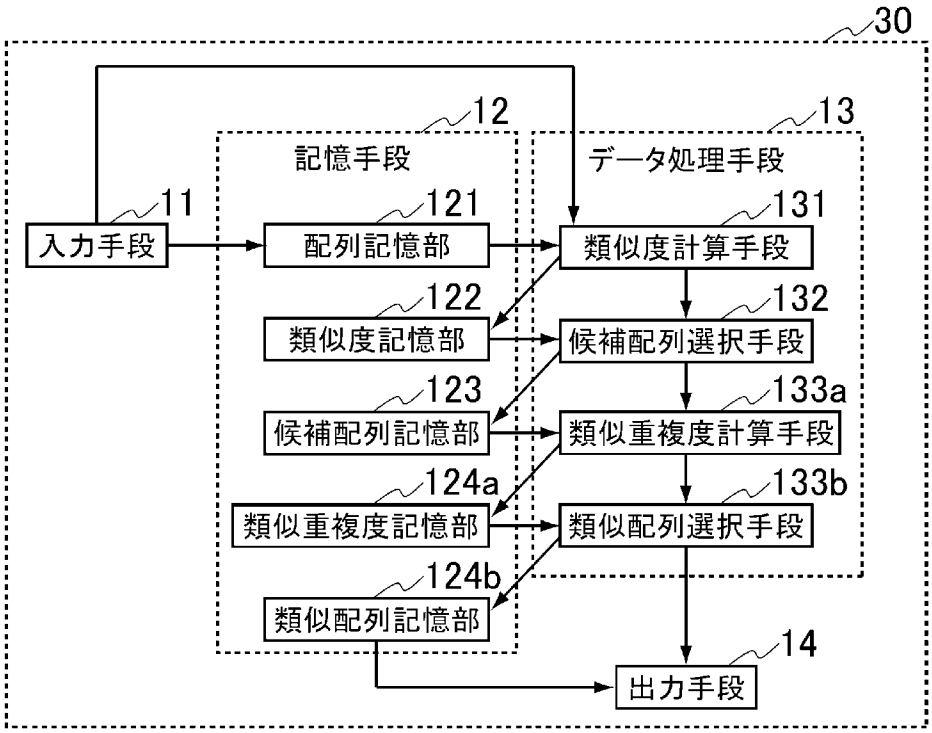
[図5]


START
A1
配列入力
A2
類似度計算
A3
候補配列選択
A4
類似配列選択
END

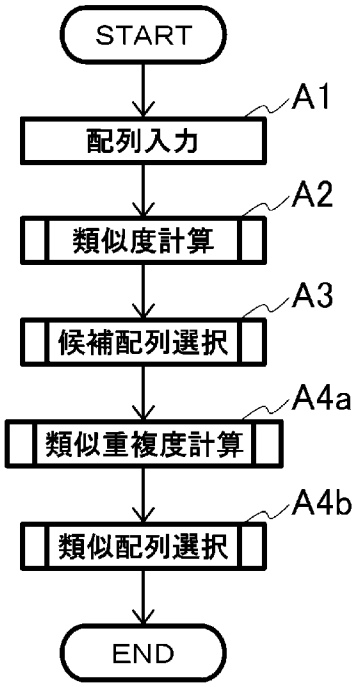
[図6]


START
A
A41
新しい比較元候補配列をセット
B
A42
新しい比較先候補配列をセット
A43
比較先が
比較元に類似？
Yes
No
A44
類似配列群ではない
との結果を出力
A45
類似配列群である
との結果を出力
A46
未比較の比較先
候補配列がある？
Yes
B
No
A47
未比較の比較元
候補配列がある？
Yes
A
No
END

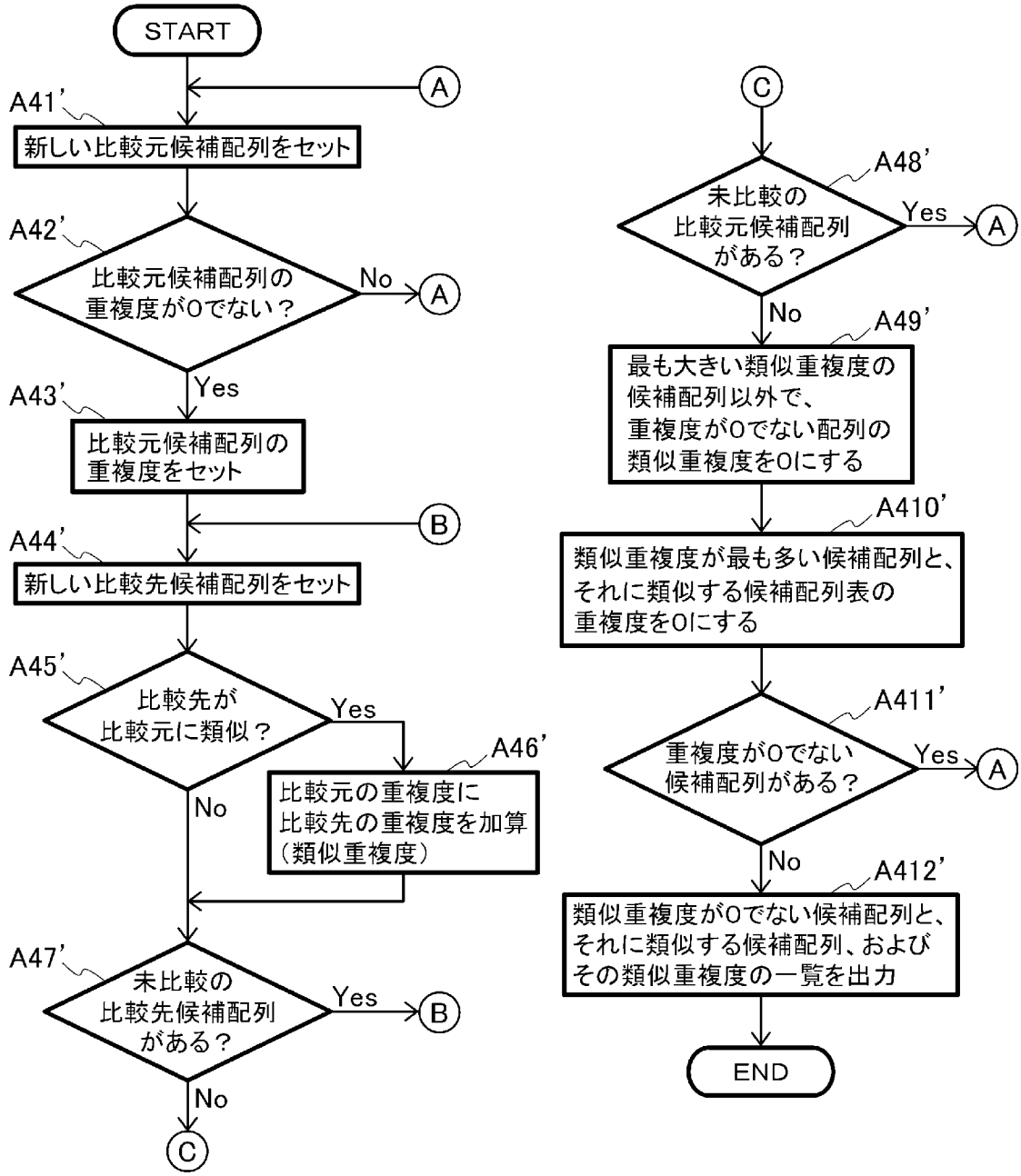
[図7]


30
12
記憶手段
13
データ処理手段
11
入力手段
121
配列記憶部
131
類似度計算手段
122
類似度記憶部
132
候補配列選択手段
123
候補配列記憶部
133a
類似重複度計算手段
124a
類似重複度記憶部
133b
類似配列選択手段
124b
類似配列記憶部
14
出力手段

[図8]


START
A1
配列入力
A2
類似度計算
A3
候補配列選択
A4a
類似重複度計算
A4b
類似配列選択
END

[図9]


START
A41'
新しい比較元候補配列をセット
A42'
比較元候補配列の
重複度が0でない？
No
A
Yes
A43'
比較元候補配列の
重複度をセット
B
A44'
新しい比較先候補配列をセット
A45'
比較先が
比較元に類似？
Yes
A46'
比較元の重複度に
比較先の重複度を加算
（類似重複度）
No
A47'
未比較の
比較先候補配列
がある？
Yes
B
No
C
C
A48'
未比較の
比較元候補配列
がある？
Yes
A
No
A49'
最も大きい類似重複度の
候補配列以外で、
重複度が0でない配列の
類似重複度を0にする
A410'
類似重複度が最も多い候補配列と、
それに類似する候補配列表の
重複度を0にする
A411'
重複度が0でない
候補配列がある？
Yes
A
No
A412'
類似重複度が0でない候補配列と、
それに類似する候補配列、および
その類似重複度の一覧を出力
END

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2014/053516

A. CLASSIFICATION OF SUBJECT MATTER
G06F19/22(2011.01)i, *C12N15/115*(2010.01)n

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G06F19/22, C12N15/115

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
Jitsuyo Shinan Koho 1922-1996 Jitsuyo Shinan Toroku Koho 1996-2014
Kokai Jitsuyo Shinan Koho 1971-2014 Toroku Jitsuyo Shinan Koho 1994-2014

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X Y	JP 2012-146067 A (Japan Software Management Co., Ltd.), 02 August 2012 (02.08.2012), claims 1 to 10; paragraphs [0010] to [0172] & WO 2012/096016 A1 & EP 2665009 A1 & US 2014/019062 A1 & CN 103339632 A	1-11,21-31, 41,44 12-20,32-40, 42,43
Y	JP 2012-146066 A (Japan Software Management Co., Ltd.), 02 August 2012 (02.08.2012), claims 1 to 6; paragraphs [0010] to [0165] & WO 2012/096015 A1 & EP 2665010 A1 & US 2014/058682 A1 & CN 103348350 A	12-20,32-40, 42,43

☒ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
09 April, 2014 (09.04.14)	22 April, 2014 (22.04.14)
Name and mailing address of the ISA/ Japanese Patent Office Facsimile No.	Authorized officer Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2014/053516

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	JP 2005-102695 A (National Institute of Advanced Industrial Science and Technology), 21 April 2005 (21.04.2005), paragraphs [0034] to [0042] (Family: none)	17-20,37-40, 43

国際調査報告

国際出願番号　ＰＣＴ／ＪＰ２０１４／０５３５１６

Ａ．　発明の属する分野の分類（国際特許分類（ＩＰＣ））

Int.Cl.　G06F19/22(2011.01)i, C12N15/115(2010.01)n

Ｂ．　調査を行った分野

調査を行った最小限資料（国際特許分類（ＩＰＣ））

Int.Cl.　G06F19/22, C12N15/115

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	１９２２－１９９６年
日本国公開実用新案公報	１９７１－２０１４年
日本国実用新案登録公報	１９９６－２０１４年
日本国登録実用新案公報	１９９４－２０１４年

国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）

Ｃ．　関連すると認められる文献

引用文献のカテゴリー＊	引用文献名　及び一部の箇所が関連するときは、その関連する箇所の表示	関連する請求項の番号
X	JP 2012-146067 A (日本ソフトウェアマネジメント株式会社) 2012.08.02, 【請求項1】-【請求項10】, 【0010】-【0172】 & WO 2012/096016 A1 & EP 2665009 A1 & US 2014/019062 A1 & CN 103339632 A	1-11, 21-31, 41, 44
Y		12-20, 32-40, 42, 43
Y	JP 2012-146066 A (日本ソフトウェアマネジメント株式会社) 2012.08.02, 【請求項1】-【請求項6】, 【0010】-【0165】 & WO 2012/096015 A1 & EP 2665010 A1 & US 2014/058682 A1 & CN 103348350 A	12-20, 32-40, 42, 43

☑　Ｃ欄の続きにも文献が列挙されている。

☐　パテントファミリーに関する別紙を参照。

＊　引用文献のカテゴリー

「Ａ」特に関連のある文献ではなく、一般的技術水準を示すもの

「Ｅ」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの

「Ｌ」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）

「Ｏ」口頭による開示、使用、展示等に言及する文献

「Ｐ」国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献

「Ｔ」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの

「Ｘ」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの

「Ｙ」特に関連のある文献であって、当該文献と他の１以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの

「＆」同一パテントファミリー文献

国際調査を完了した日	国際調査報告の発送日
０９．０４．２０１４	２２．０４．２０１４

国際調査機関の名称及びあて先
日本国特許庁（ＩＳＡ／ＪＰ）
郵便番号１００－８９１５
東京都千代田区霞が関三丁目４番３号

特許庁審査官（権限のある職員）　５Ｌ　４０６４
山内　裕史
電話番号　０３－３５８１－１１０１　内線　３５６２

C（続き）．　関連すると認められる文献

引用文献のカテゴリー＊	引用文献名　及び一部の箇所が関連するときは、その関連する箇所の表示	関連する請求項の番号
Y	JP 2005-102695 A（独立行政法人産業技術総合研究所）2005.04.21, 【0034】-【0042】（ファミリーなし）	17-20, 37-40, 43