

(43) 国際公開日
2007 年 10 月 25 日 (25.10.2007)

PCT

(10) 国際公開番号
WO 2007/119779 A1

- (51) 国際特許分類:
G01N 27/447 (2006.01) C12Q 1/68 (2006.01)
C12M 1/00 (2006.01)
- (21) 国際出願番号: PCT/JP2007/058055
- (22) 国際出願日: 2007 年 4 月 12 日 (12.04.2007)
- (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2006-112600 2006 年 4 月 14 日 (14.04.2006) JP
- (71) 出願人 (米国を除く全ての指定国について): 日本電気株式会社 (NEC CORPORATION) [JP/JP]; 〒1088001 東京都港区芝五丁目 7 番 1 号 Tokyo (JP).
- (72) 発明者; および
- (75) 発明者/出願人 (米国についてのみ): 麻生川 稔 (ASOGAWA, Minoru) [JP/JP]; 〒1088001 東京都港区芝五丁目 7 番 1 号 日本電気株式会社内 Tokyo (JP). 杉澤 正俊 (SUGISAWA, Masatoshi) [JP/JP]; 〒1088001 東京都港区芝五丁目 7 番 1 号 日本電気株式会社内 Tokyo (JP).

奥井 伸司 (OKUI, Shinji) [JP/JP]; 〒1088001 東京都港区芝五丁目 7 番 1 号 日本電気株式会社内 Tokyo (JP).

- (74) 代理人: 宮崎 昭夫, 外 (MIYAZAKI, Teruo et al.); 〒1070052 東京都港区赤坂 1 丁目 9 番 2 0 号 第 1 6 興和ビル 8 階 Tokyo (JP).

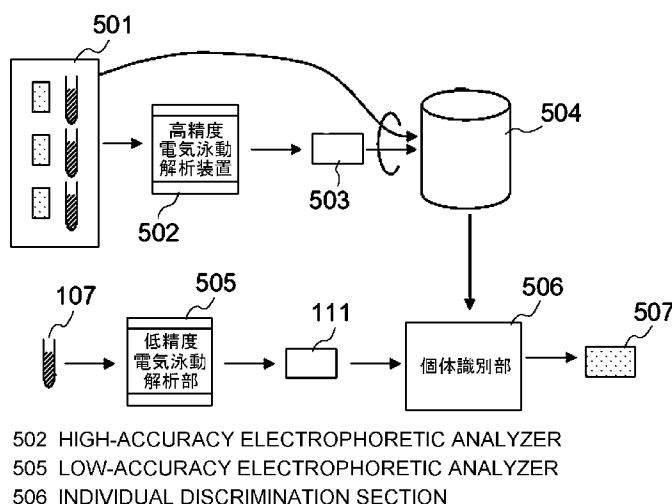
- (81) 指定国 (表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, 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, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), ヨーロッパ (AT, BE, BG,

[続葉有]

(54) Title: INDIVIDUAL DISCRIMINATION METHOD AND APPARATUS

(54) 発明の名称: 個体識別方法および装置



(57) Abstract: An individual discrimination method of discriminating an individual by analyzing a DNA sample using electrophoresis, which comprises: the first analysis step of analyzing a DNA sample having a discrimination mark for an individual attached thereto; the step of storing the results of the analysis on the DNA sample having the discrimination mark attached thereto together with the corresponding discrimination mark; the second analysis step of analyzing a new sample, which is a DNA sample provided as the subject of the individual discrimination, at a lower accuracy than the accuracy in the analysis of the DNA sample having the discrimination mark attached thereto and referring the obtained results as the analysis results of the new sample; and the step of searching for database based on the analysis results of the new sample.

(57) 要約:

DNAサンプルを電気泳動によって解析することによって個体を識別する個体識別方法は、個体に対する識別子が付与された識別子付きDNAサンプルを解析する第1の解析段階と、識別子付きDNAサンプルを解析して得られた結果を対応する識別子とともにデータベースに蓄積する段階と、個体識別対象となるDNAサンプルである新サンプルを、識別子付きDNAサンプルを解析したときの精度よりも低い精度で解析し、結果を新サンプル解析結果とする第2の解析段階と、新サンプル解析結果に基づいてデータベースを検索する段階と、を有する。



CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE,
IS, IT, LT, LU, LV, MC, MT, NL, PL, PT, RO, SE, SI, SK,
TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW,
ML, MR, NE, SN, TD, TG).

2文字コード及び他の略語については、定期発行される
各PCTガゼットの巻頭に掲載されている「コードと略語
のガイダンスノート」を参照。

添付公開書類:

— 国際調査報告書

明 細 書

個体識別方法および装置

技術分野

- [0001] 本発明は、DNA(デオキシリボ核酸)に対して電気泳動法を用いた個体識別方法に関し、特に、低い読み取り性能しか有しない電気泳動解析装置を用いて精度よく個体識別を行う方法および装置に関する。

背景技術

- [0002] 犯罪捜査などを目的として、DNAを用いた個体識別、すなわちいわゆるDNA鑑定を行う場合には、ゲノムのうち個体ごと異なっているDNA領域に対して解析を行っている。DNAを解析する方法の1つとして電気泳動法があり、広く用いられている。電気泳動法は、電場を印加した場合に、DNAの性質の違いによってその流動速度が異なることを利用するものである。
- [0003] ヒトのDNAを用いた個体識別としては、マイクロサテライトと呼ばれる、4～5塩基程度の配列が繰り返し起こる領域を解析することによって行う方法が、FBI(米国連邦捜査局)や日本の警察機関等で使用されている。マイクロサテライト領域の繰り返し回数を計測する一つの方法として、DNAの塩基長を電気泳動法で計測する方法がある。個体識別のためにDNAの電気泳動計測を行う場合、その機材として、DNA解読プロジェクト(あるいはゲノム解読プロジェクト)でも多く使われてきたDNAシーケンサを利用することが多い。
- [0004] DNAシーケンサでは、長さ40cm程度のキャピラリー(毛细管)にゲルを充填したものを電気泳動のための媒体をして用いる。このキャピラリーの一端から、DNAのマイクロサテライトとそれに隣接する領域のみをPCR(ポリメラーゼ連鎖反応: Polymerase Chain Reaction)で増幅させたDNA断片を含んだ溶液試料を導入し、電界に起因する力によって生じる電気泳動によって、キャピラリーのもう一方の端部に向けてこのDNA断片を移動させる。PCRで増幅させたこのDNA断片のことをアンプリコンと呼ぶ。このとき、アンプリコンのサイズすなわちDNAの塩基数によって移動速度が異なるから、アンプリコンごとにキャピラリーのもう一方の端部に到達するまでの

時間に差が出てくる。ここで、キャピラリーのもう一方の端部にどのようなタイミングでアンプリコンが到達したかを計測することによって、そのアンプリコンに係るDNAのサイズを推定でき、ひいてはマイクロサテライト領域での繰り返し回数を計測できる。

[0005] この方法は、ヒトのみならず、個体によって異なるDNA領域を有する生物種に対して適用できる。DNAを用いたヒトの個体識別の方式としてFBIが提唱しているDNA鑑定システムである複合DNAインデックスシステム(CODIS; Combined DNA Index System)などでは、上述したマイクロサテライトを解析する方法を用いているが、ここで使用されている遺伝子座(ローカス; locus)でのマイクロサテライトの繰り返しの塩基数は、4塩基もしくは5塩基単位である。

[0006] マイクロサテライトの解析以外に、DNAを制限酵素で断片化し、長さの異なる断片を解析することによって個体識別する方法もある。制限酵素とは、DNAでの特異な配列を認識して切断する酵素のことである。この方法でも、解析に電気泳動法を用いることができる。

[0007] ところで、PCR増幅によって得られるアンプリコンは、ターゲットとするマイクロサテライトの繰り返し配列部分と、PCRのプライマとハイブリタイズするまでの部分とから構成されている。したがって、マイクロサテライトにおける各繰り返しごとの塩基数が4であるとして、あるアンプリコンのマイクロサテライトでの繰り返し回数が4であったとき、マイクロサテライト部分は $4 \times 4 = 16$ 塩基であり、PCRのプライマとハイブリタイズするまでの塩基数が例えば10塩基であったとすると、アンプリコンの塩基長は、 $10 + 16 = 26$ 塩基となる。同様に、マイクロサテライトの繰り返し回数が5であるとする、30塩基となる。以下、マイクロサテライトの繰り返し回数をSTR(短タンデム繰り返し: Short Tandem Repeat)回数と表現することにする。例えば、計測したアンプリコンのサイズが30塩基であれば、STR回数が5であると判定できる。STR回数は、アンプリコンの塩基長に対応しているから、アンプリコンについての塩基長情報であるといえる。

[0008] 上記の例では、アンプリコンの長さは、30塩基、34塩基など、STR回数が1つ増えると、4塩基(もしくは5塩基)単位で増える。ところが、ヒトのDNA鑑定で使われる遺伝子領域には、繰り返しが4塩基(もしくは5塩基)刻みでないものがある。例えば、ある場合では、通常のSTRに加えて2塩基余計に付け加わったタイプがある。STRの5

回の繰返しに2塩基余計に付け加わったタイプの場合は、“5. 2”といった表記がなされる。STR回数が5の時、アンプリコンが30塩基だったとすると、“5. 2”では、32塩基となる。xx. 2以外に、xx. 1、xx. 3などが存在する。このように繰返しに対して端数となる塩基は、すべてのSTR回数に対して存在するのではなく、種類が限られている、すなわち特定のSTR回数に対して起こることが知られている。

[0009] 例えば、FGAと呼ばれるローカスには、以下の様なバラエティ(多様性)が存在する。表1は、ローカスのバラエティの出現確率を示した例を示したものであって、アメリカ合衆国におけるアフリカ系のヒト200名程度に対してFGAの多様性を調べたデータを示している。ここでは、FGAの種類は18種類であり、すなわちFGAには18通りのSTR回数が存在し、そのうち4種類のみがxx. 2のタイプのものである。表1に示すデータにおいて、合計が200よりも大きいのは、以下に述べるようにSTR回数において父親由来か母親由来の2種類があるためであり、400よりも少ないのは、解析できなかったものがあるからである。また、出現確率の総和が1. 0を上回るのは、STR回数の出現頻度が5以下の場合は、出現確率として一律に0. 014としているからである。表1に示すデータは、Bruce Budowle, “Genotype Profiles for Six Population Groups at the 13 CODIS Short Tandem Repeat Core Loci and Other PCRBBased Loci”, Forensic Science, Volume 1, Number 2, (July 1999) (非特許文献1)において、“dnaloci. txt”として公表されている生データに基づくものである。

[0010] [表1]

表1 ローカスのバラエティの出現頻度の例

STR回数	頻度	確率
17.2	1	0.014
18	3	0.014
18.2	3	0.014
19	19	0.053
19.2	1	0.014
20	26	0.072
21	45	0.125
22	81	0.225
22.2	2	0.014
23	45	0.125
24	67	0.186
25	36	0.100
26	13	0.036
27	8	0.022
28	6	0.017
29	2	0.014
30	1	0.014
30.2	1	0.014
総合	360	1.073

[0011] ヒトのゲノムは2セットあるので、各ローカスに関して父親由来のSTR回数と母親由来のSTR回数が存在することとなり、これは、個人を特定する情報となる。例えば、あるローカスでのSTR回数の種類が10種類であるとする、全ての組み合わせは、100種類(=10×10)存在することになる。このうち10種類のものは、父親由来のローカスにおけるSTR回数と母親由来のローカスにおけるSTR回数が一致している。したがって、そのようなヒトに対してDNA解析を行い、STR回数を求めたとしても、1通りのSTR回数しか得られないことになる。このような場合をホモ接合と呼ぶ。

[0012] ホモ接合を除いた残りの90通りでは、父親由来のSTR回数と母親由来のSTR回数とが異なり、このようなヒトに対してDNA解析を行えば、精度が十分であれば、2通りのSTR回数を得られることになる。このような場合をヘテロ接合と呼ぶ。もっとも、DNAを解析した場合に、どちらのSTR回数が父親由来であって、どちらのSTR回数が母親由来であるかの区別がつかないので、実際には、ヘテロ接合は、90種類の半

分である45種類となる。

- [0013] すなわち、各ゲノムの各セットごとにローカスのSTR回数の種類が10種類あるとき、DNA解析で存在し得る結果は、ホモ接合の10種類とヘテロ接合の45種類とをあわせた全部で55種類であり、これは個人を特定する情報となる。マイクロサテライトを用いたDNAの解析と照会では、この55種類のうちどのタイプかを解析し、その解析結果と完全一致するエントリをデータベース中から検索することになる。
- [0014] DNAを用いた個体識別の現場では、認識精度の向上のために、複数のローカスについて解析を行い、データベース検索を行っている。ヒトにおいてSTR回数は各ローカスごとに独立に決まっているので、複数のローカスの解析をすることによって、認識精度を高めることができる。FBIなどで行われているDNA解析では、13個のローカスを用いている。このようなDNA解析の詳細については、例えば、“Forensic DNA Typing, Second Edition : Biology, Technology, and Genetics of STR Markers”, John M. Butler. (2005). pp. 85–117, 345–370, and 373–386(非特許文献2)に詳しく述べられている。
- [0015] なお、特開2002-253203号公報(特許文献1)は、個人を特定するDNAの塩基配列情報をデジタル化して、バーコードあるいはIC(integrated circuit)カードなどに固定することを開示している。特開2003-245098号公報(特許文献2)は、PCR産物を電気泳動で検出し、塩基長配列のサイズ情報を得ることを開示している。特開2004-073188号公報(特許文献3)は、同定されるべき物体中へマーカーを組み込む方法であって、DNA断片をマーカーとして用いる方法を開示している。特開2005-013226号公報(特許文献4)は、ダイズをDNAより同定する方法であって、PCRの結果を電気泳動等を用いて同定するとともに、ダイズの既知遺伝子配列の検索を行う際に、データベースにアクセスしてサテライトDNAの検索を行うことを開示している。特開2005-160302号公報(特許文献5)は、マイクロサテライト遺伝多型マーカーを用いる遺伝子のマッピング方法を開示している。特開2005-237334号公報(特許文献6)は、テロメア反復配列とこれに相補的な標識プローブとをハイブリダイズさせ、そのDNA一分子の動きの速度を検出することにより、迅速かつ高感度にDNA反復配列を測定する方法を開示している。特開2005-307216号公報(特許文献7

)は、本人認証に利用することのできる合成DNAインクを開示している。特開平11-118760(特許文献8)は、DNA断片の電気泳動パターンの解析法であって、データベース化に適した解析法を開示している。

[0016] WO97/15690(特許文献9)は、DNA配列の定量、同定または判定に関する発明を開示している。WO98/35060(特許文献10)は、多型核酸フラグメントを分析または型分類するためのポリメラーゼを開示している。WO01/14590(特許文献11)は、既知量の、シリカ磁性粒子のような、DNA標的物質を可逆的に結合するための定義できる能力を有するシリカ含有固形支持体と、その粒子の結合能力より多いDNA標的物質とを使用して、媒体中の他の物質から、定義された量のDNA標的物質を単離する方法を開示している。WO02/08469(特許文献12)は、対立遺伝子コール(allele call)を行うためのコンピュータにより実行される方法を開示している。WO02/66650(特許文献13)は、ストレプトコッカス(streptococcus)抗原のフラグメントの解析について開示している。WO03/06692(特許文献14)は、電気泳動分析のための内部校正標準に関する発明を開示している。WO02/86794(特許文献15)は、質量分析に基づいてDNAを解析する方法を開示している。

[0017] 以下、本明細書中で引用した文献を列举する。

特許文献1:特開2003-253203号公報

特許文献2:特開2003-245098号公報

特許文献3:特開2004-073188号公報

特許文献4:特開2005-013226号公報

特許文献5:特開2005-160302号公報

特許文献6:特開2005-237334号公報

特許文献7:特開2005-307216号公報

特許文献8:特開平11-118760号公報

特許文献9:WO97/15690(特表2000-500647)

特許文献10:WO98/35060(特表2001-511018)

特許文献11:WO01/14590(特表2003-507049)

特許文献12:WO02/08469(特表2004-516455)

特許文献13:WO02/66650(特表2004-531235)

特許文献14:WO03/06692(特表2004-535198)

特許文献15:WO02/86794(特表2005-509844)

非特許文献1:Bruce Budowle, "Genotype Profiles for Six Population Groups at the 13 CODIS Short Tandem Repeat Core Loci and Other PCRBBased Loci", Forensic Science, Volume 1, Number 2, (July 1999). (インターネット上でも、次のURLから入手可能<URL:<http://www.fbi.gov/hq/lab/fsc/backissu/july1999/budowle.htm>>

非特許文献2:"Forensic DNA Typing, Second Edition : Biology, Technology, and Genetics of STR Markers", John M. Butler. (2005). pp. 85-117, 345-370, and 373-386

発明の開示

発明が解決しようとする課題

[0018] 個体識別のための上述した従来のDNA解析では、大型の電気泳動装置を使用する必要があり、そのため、電気泳動に要する時間が長くなって解析時間が長くなる、という問題点がある。この原因は、DNAの解析とDNAデータベースとの照合において、アンプリコンの長さの計測を1bp(塩基対)の精度で行っているからである。このように1bpといった高い精度で解析しているのは、例えばヒトのDNAを用いた個体識別の方式としてFBIが提唱しているCODISなどでは、そこで使用されているローカスのアンプリコンのDNAサイズの最小の変化幅が2bp程度であるので、1bp程度の精度で塩基長を認識しないと、データベースとの照合ができなくなるからである。

[0019] アンプリコンの長さの計測精度を確保するためには、電気泳動装置において現在使用されているキャピラリーよりも短いキャピラリーを用いたり、電気泳動の路長をより短くしたりすることができない。そのため、電気泳動装置などの構成を簡素化したり、あるいは短時間で電気泳動の解析を行うことができない。

[0020] 本発明の目的は、読み取り精度の低い電気泳動装置を用いても必要な精度を確保でき、かつ、短時間での解析を可能とする個体識別方法を提供することにある。

[0021] 本発明の別の目的は、読み取り精度の低い電気泳動装置を用いても必要な精度を

確保でき、かつ、短時間での解析を可能とする個体識別装置を提供することにある。

課題を解決するための手段

[0022] 個体識別のためのDNA解析の運用を考えると、あらかじめデータベースを構築した上で、新たに入手した検体のDNAを解析し、その解析結果がデータベースに格納されたものと一致するかどうかを照合することになる。ここで、短時間での処理、あるいは簡素化された装置での処理が要求されるのは、新たに入手した検体のDNAの解析であり、データベースにあらかじめ格納しておくデータのためのDNA解析には、簡素化された装置での処理や短時間での処理に対する要求はあまりない。そこで本発明では、新たに入手した検体のDNA解析を、今までは個体識別のためのDNA解析には使用できなかったような精度の低い電気泳動装置を用いて行えるようにする。以下、新たに入手した検体のことを新サンプルと呼ぶ。

[0023] なお、データベース登録のための試料(サンプル)は、その出所、すなわち誰の、あるいはいつどこで採取されたものか、が明確なものであり、出所を特定するための識別子が付与されている。そこで、以下の説明において、データベース登録のための試料(サンプル)を識別子付きサンプルと呼ぶことにする。識別子付きサンプルをデータベース(すなわち識別子付きサンプル解析データ蓄積部)に蓄積するときのDNA解析には、従来用いられていような相対的に精度の高い電気泳動装置を用いてもよいし、従来は使用できなかったような相対的に精度の低い電気泳動装置を用いてもよい。後述するように、本発明では、識別子付きサンプルの解析及び新サンプルの解析の両方に低精度の電気泳動装置と用いたとしても、データベースでの照合を精度よく行えるようにしている。

[0024] 本発明の目的は、DNAサンプルを電気泳動によって解析することによって個体を識別する個体識別方法であって、個体に対する識別子が付与された識別子付きDNAサンプルを解析する第1の解析段階と、識別子付きDNAサンプルを解析して得られた結果を対応する識別子とともに識別子付きサンプル解析データ蓄積部に蓄積する段階と、個体識別対象となるDNAサンプルである新サンプルを、識別子付きDNAサンプルを解析したときの精度よりも低い精度で解析し、結果を新サンプル解析結果とする第2の解析段階と、新サンプル解析結果に基づいて、識別子付きサンプル

解析データ蓄積部を検索する段階と、を有する固体識別方法によって達成される。

[0025] この個体識別方法においては、例えば、識別子付きDNAサンプル及び新サンプルを解析する際に、電気泳動法によってそのサンプルの塩基長に関する情報が求められ、特に、そのサンプルにおけるマイクロサテライトの繰り返し回数に関する情報が取得される。

[0026] この個体識別方法において、典型的には、第1の解析段階における解析精度は、新サンプルでの塩基長の考え得る最小の変化量だけ塩基長が異なる2つのDNAを識別できる精度であり、第2の解析段階における解析精度は、最小の変化量だけ塩基長が異なる2つのDNAを識別できない精度である。

[0027] また第2の解析段階は、例えば、それぞれ1種のアンプリコンを含んだサンプルの集合から複数のサンプルを任意の組み合わせで選択し、選択されたサンプルを混合して多種アンプリコン・サンプルを生成する段階と、多種アンプリコン・サンプルを電気泳動によって解析する第3の解析段階と、第3の解析段階で得られた結果と、当該多種アンプリコン・サンプルにおける塩基長情報とを組にして多種アンプリコン・データ蓄積部に蓄積する段階と、電気泳動によって新サンプルを解析し、新サンプル電気泳動結果データを得る第4の解析段階と、新サンプル電気泳動結果データに基づいて多種アンプリコン・データ蓄積部を検索し、結果を新サンプル解析結果とする検索段階と、を有する。

[0028] あるいは本発明の目的は、DNAサンプルを電気泳動によって解析することによって個体を識別する個体識別方法であって、個体に対する識別子が付与された識別子付きDNAサンプルを解析して識別子付きDNAサンプルの塩基長に関する情報を得る第1の解析段階と、識別子付きDNAサンプルを解析して得られた結果を対応する識別子とともに識別子付きサンプル解析データ蓄積部に蓄積する段階と、個体識別対象となるDNAサンプルである新サンプルを解析し、新サンプルの塩基長に関する情報を含む結果を新サンプル解析結果とする第2の解析段階と、新サンプル解析結果に基づいて、識別子付きサンプル解析データ蓄積部を検索する段階と、を有し、第1の解析段階および第2の解析段階における解析の精度は、個体識別対象となるDNAサンプルでの塩基長の考え得る最小の変化量だけ塩基長が異なる2つの

DNAを識別できない精度である固体識別方法によっても達成される。

[0029] 本発明の第2の目的は、DNAサンプルを電気泳動によって解析することによって個体を識別する個体識別装置であって、個体に対する識別子が付与された識別子付きDNAサンプルを解析する第1の解析手段と、第1の解析手段で識別子付きDNAサンプルを解析して得られた結果を対応する識別子とともに蓄積する識別子付きサンプル解析データ蓄積部と、第1の解析手段より低い解析精度を有し、個体識別対象となるDNAサンプルである新サンプルを解析して結果を新サンプル解析結果とする第2の解析手段と、新サンプル解析結果に基づいて、識別子付きサンプル解析データ蓄積部を検索し、個体識別結果を得る識別手段と、を有する個体識別装置によって達成される。

[0030] あるいは本発明の第2の目的は、DNAサンプルを電気泳動によって解析することによって個体を識別する個体識別装置であって、個体に対する識別子が付与された識別子付きDNAサンプルを解析して識別子付きDNAサンプルの塩基長に関する情報を得る第1の解析手段と、識別子付きDNAサンプルを解析して得られた結果を対応する識別子とともに蓄積する識別子付きサンプル解析データ蓄積部と、個体識別対象となるDNAサンプルである新サンプルを解析し、新サンプルの塩基長に関する情報を含む結果を新サンプル解析結果とする第2の解析手段と、新サンプル解析結果に基づいて、識別子付きサンプル解析データ蓄積部を検索する識別手段と、を有し、第1の解析手段および第2の解析手段における解析の精度は、個体識別対象となるDNAサンプルでの塩基長の考え得る最小の変化量だけ塩基長が異なる2つのDNAを識別できない精度である個体識別装置によっても達成される。

[0031] 上述した本発明の個体識別方法及び装置は、DNAサンプルを解析することによって個体識別を行うものであるが、さらに、指紋情報、掌紋情報、虹彩情報、顔情報などの他のバイOMETリック情報と組み合わせることによって、個体識別の精度をさらに向上させることができる。

[0032] 本発明によれば、特に、個体識別の対象となるサンプルすなわち新サンプルの電気泳動による解析の際に、現在使用されているキャピラリーよりも短いキャピラリーや、短い電気泳動の路長を用いることができるから、解析に必要な時間が短縮され、そ

の結果、DNAを用いた個体識別を短時間で行うことが可能となる。

[0033] また、このように短いキャピラリーや短い電気泳動の路長を用いるため、以下のような有利な点を享受することができる：

(1) 従来の個体識別装置に比べてサイズを小さくして装置構成を簡略化することができ、その結果、屋内屋外を問わず必要な場所でDNAを用いた個体識別が可能となる；

(2) 装置の全体を覆ったりして外部からの雑物の混入を防ぐことが容易になるとともに、装置構成が簡易であるので、温度や湿度などの電気泳動に影響を与える外部要因を制御するのが容易になり、その結果、外部からの雑物の混入や解析装置の環境の不安定性に起因する誤解析を防ぐことが可能となる；

(3) 装置全体を覆うことなどによって、温度や湿度などの装置に影響を与える外部要因を制御するのが容易になり、保守性の向上や耐故障性の向上が可能となる。

[0034] 上述したように、本発明によれば、短時間でのDNAを用いた個体識別が可能となり、屋内屋外を問わず必要な場所でDNAを用いた個体識別が可能となり、誤解析を防ぐことが可能となる。これにより、他のバイOMETリック情報によって個体識別を行う機器との組み合わせが容易に行えるようになり、他のバイOMETリック情報を用いた個体認識との組み合わせによって認識精度の向上を図ることができる。

図面の簡単な説明

[0035] [図1]本発明の第1の実施形態の個体識別装置の構成を示す図である。

[図2]図1に示す個体識別装置における低精度電気泳動解析部の構成を示す図である。

[図3]2種類のアンプリコンを含んだDNAサンプルを混合したものを解析したシミュレーション結果を示すグラフである。

[図4]本発明の第2の実施形態の個体識別装置における低精度電気泳動解析部の構成を示す図である。

[図5]本発明の第3の実施形態の個体識別装置における低精度電気泳動解析部の構成を示す図である。

[図6]本発明の第4の実施形態の個体識別装置における低精度電気泳動解析部の

構成を示す図である。

[図7]本発明の第5の実施形態の個体識別装置の構成を示す図である。

[図8]本発明の第6の実施形態の個体識別装置の構成を示す図である。

[図9]本発明の第7の実施形態の個体識別装置の構成を示す図である。

[図10]本発明の第8の実施形態の個体識別装置の構成を示す図である。

符号の説明

- [0036]
- 101 1種アンプリコン・サンプル保管部
 - 102 選択サンプル
 - 103 多種アンプリコン・サンプル
 - 104 電気泳動解析部
 - 105 多種アンプリコン電気泳動結果データ
 - 106 多種アンプリコン・データ蓄積部
 - 107 新サンプル
 - 108 新サンプル電気泳動解析部
 - 109 新サンプル電気泳動結果データ
 - 110 新サンプル結果データ解析部
 - 111 新サンプル解析結果
 - 201 補間データ作成部
 - 202 補間データ蓄積部
 - 301 パラメータ推定機能付き新サンプル結果データ解析部
 - 401 1種アンプリコン電気泳動結果データ
 - 402 1種アンプリコン・データ蓄積部
 - 403 補間多種アンプリコン・データ作成部
 - 404 補間多種アンプリコン・データ蓄積部
 - 501 識別子付きサンプル
 - 502 高精度電気泳動解析装置
 - 503 サンプル解析結果
 - 504 識別子付きサンプル解析データ蓄積部

- 505 低精度電気泳動解析部
- 506 個体識別部
- 507 個体識別結果
- 601 低精度識別子付きサンプル解析結果
- 602 低精度識別子付きサンプル解析データ蓄積部
- 603 低精度個体識別部
- 604 低精度個体識別結果
- 701 高精度個体識別部
- 702 高精度個体識別結果
- 901 新サンプル取得対象
- 902 DNAサンプル
- 903 DNA解析に基づく個体識別部
- 904 識別子付きDNA解析データ蓄積部
- 905 DNA解析に基づく個体識別結果
- 906 指紋サンプル
- 907 指紋解析に基づく個体識別部
- 908 識別子付き指紋解析データ蓄積部
- 909 指紋解析に基づく個体識別結果
- 910 複数情報による個体識別部
- 911 複数情報による個体識別結果
- 1001 (5, 5)混合サンプル
- 1002 (5, 5. 2)混合サンプル
- 1003 (5, 6)混合サンプル
- 1004 (5, 6. 2)混合サンプル
- 1005 (5, 7)混合サンプル
- 1006 (5, 7. 2)混合サンプル
- 1007 (5, 8)混合サンプル

発明を実施するための最良の形態

[0037] 《第1の実施形態》

図1は、本発明の第1の実施形態の個体識別装置の構成を示している。この個体識別装置は、識別子付きサンプル501を電気泳動により解析する高精度電気泳動解析装置502と、高精度電気泳動解析装置502から出力されるサンプル解析結果503を蓄積する識別子付きサンプル解析データ蓄積部504と、新サンプル107を電気泳動により解析する低精度電気泳動解析部505と、低精度電気泳動解析部505から出力される新サンプル解析結果111に基づいて識別子付きサンプル解析データ蓄積部504内のデータを検索して新サンプル107の個体識別を行い、個体識別結果507を出力する個体識別部506と、を備えている。新サンプル107は、個体識別を行おうとするDNAのサンプルであり、第1の実施形態の個体識別装置は、新サンプル107でのDNAのSTR回数を計測し、その計測結果に応じてデータベースすなわち識別子付きサンプル解析データ蓄積部504を検索して個体識別を行うものである。

[0038] 識別子付きサンプル501は、個体の識別子がついたサンプルの集合であり、高精度電気泳動解析装置502は、このような識別子付きサンプル501の各サンプルを、従来使用されているような十分な読み取り精度で解析する装置である。サンプル解析結果503は、識別子付きサンプル501を高精度電気泳動解析装置502を用いて解析した結果であり、識別子付きサンプル501に含まれているDNA中の複数のSTR回数のセットを示すデータからなっている。識別子付きサンプル解析データ蓄積部504は、識別子付きサンプル501の個体ごとに、従来使用されているような十分な読み取り精度で解析されたサンプル解析結果503である複数のSTR回数のセットと識別子付きサンプル501の個体の識別子とを組にして蓄積する。

[0039] 新サンプル107を解析する低精度電気泳動解析部505の構成の詳細については後述するが、低精度電気泳動解析部505自体も、電気泳動解析装置を備えて構成されており、第1の実施形態では、この低精度電気泳動解析部505は、高精度電気泳動解析装置502に比べて、読み取り精度が、同程度か低いことを想定している。新サンプル解析結果111は、新サンプル107の解析結果であり、複数のSTR回数のセットを示すデータからなっている。個体識別部506では、新サンプル解析結果111の

複数のSTR回数のセットと識別子付きサンプル解析データ蓄積部504の各エントリのSTR回数のセットとに重なりがある識別子を識別子付きサンプル解析データ蓄積部504内で検索し、個体識別結果507とする。個体識別結果507には、個体識別子が1つ含まれている場合または複数含まれている場合があり、さらに、個体識別子が全く含まれない場合もある。

[0040] 次に、低精度電気泳動解析部505の構成について、図2を用いて説明する。

[0041] 第1の実施形態では、低精度電気泳動解析部505は、1種アンプリコン・サンプル保管部101と、1種アンプリコン・サンプル保管部101から選択されたDNAサンプルすなわち選択サンプル102を混合して得られる多種アンプリコン・サンプル103を電気泳動によって解析する電気泳動解析部104と、電気泳動解析部104から出力される多種アンプリコン電気泳動結果データ105を蓄積する多種アンプリコン・データ蓄積部106と、新サンプル107を電気泳動によって解析する新サンプル電気泳動解析部108と、新サンプル電気泳動解析部108が出力する新サンプル電気泳動結果データ109に基づいて多種アンプリコン・データ蓄積部106内を検索し、検索結果を新サンプル解析結果111として出力する新サンプル結果データ解析部110と、を備えている。

[0042] ここで、1種アンプリコン・サンプル保管部101は、それぞれが1種類ずつのアンプリコンを含んだDNAのサンプルである1種アンプリコン・サンプルを複数個保管するものであり、アンプリコン・サンプルごとにそれらのサンプルにおけるSTR回数も保持している。選択サンプル102は、1種アンプリコン・サンプル保管部101の中から複数のサンプルを任意の組み合わせで選択したもの(集合)である。このように選択された複数種類の選択サンプル102を混合することによって、多種アンプリコン・サンプル103が得られる。したがって、多種アンプリコン・サンプル103は、単一のサンプル中に複数種類の、すなわちSTR回数が異なる複数のアンプリコンが含まれていることになる。

[0043] 低精度電気泳動解析部505では、電気泳動解析部104によって、多種アンプリコン・サンプル103が解析され、その結果として多種アンプリコン電気泳動結果データ105が得られることになる。多種アンプリコン・データ蓄積部106は、多種アンプリコン

電気泳動結果データ105と、その多種アンプリコン電気泳動データ105に対応する多種アンプリコン・サンプル103を構成する各アンプリコンのSTR回数とを組にして蓄積する。また、新サンプル107を新サンプル電気泳動解析部108によって解析した結果が新サンプル電気泳動結果データ109であり、新サンプル結果データ解析部110は、新サンプル電気泳動結果データ109に基づいて多種アンプリコン・データ蓄積部106のデータを検索することにより、新サンプル107のSTR回数を解析し、そのSTR回数解析の結果を新サンプル解析結果111として出力する。

[0044] 次に、この個体識別装置の動作について説明する。

[0045] まず、データベース(すなわち識別子付きサンプル解析データ蓄積部504)にデータを蓄積するために、高精度電気泳動解析装置502によって、識別子付きサンプル501の各サンプルを十分な読み取り精度で解析し、これらのサンプルにおけるSTR回数の情報を読み取る。その結果、識別子付きサンプル501に含まれているDNA中の複数のSTR回数がサンプル解析結果503として得られるので、識別子付きサンプル解析データ蓄積部504は、サンプル解析結果503である情報と、識別子付きサンプル501に対応する個体の識別子とを組にして蓄積する。

[0046] 次に、個体識別の対象である新サンプル107を低精度電気泳動解析部505で解析し、その解析結果として、複数のSTR回数のセットである新サンプル解析結果111を得る。以下、低精度電気泳動解析部505での処理について、図2を参照して説明する。

[0047] 上述したように、複数のDNAサンプルとそれらのSTR回数とが1種アンプリコン・サンプル保管部101に保管されているから、1種アンプリコン・サンプル保管部101より2種またはそれ以上のサンプルを任意の組み合わせで選択し選択サンプル102とし、これらの選択サンプル102のDNAサンプルを混合して多種アンプリコン・サンプル103を作成する。そしてこの多種アンプリコン・サンプル103を電気泳動解析部104において電気泳動によって解析し、結果として多種アンプリコン電気泳動結果データ105を得る。電気泳動解析部104での電気泳動の結果としては、山型波形のピーク位置および山型波形の形状特徴、または、そのいずれかを用いる。山型波形の形状特徴としては、(a)ピーク高さ、(b)ピーク幅、(c)山型波形の面積、(d)波形の変曲

点、のうちの1つ以上が含まれる。電気泳動結果の解析手法は、当業者にとってよく知られており、また本発明とは直接関係しないので、その詳細な説明は省略する。

[0048] 多種アンプリコン電気泳動結果データ105が得られたら、多種アンプリコン・データ蓄積部106に、その多種アンプリコン電気泳動結果データ105と選択サンプル102でのSTR回数とを組にして蓄積する。上述したようにSTR回数はアンプリコンの塩基長情報であるから、多種アンプリコン・データ蓄積部106には、多種アンプリコンの塩基長情報が蓄積されることになる。これらの処理により、複数種類のアンプリコンのDNAのサンプルの組み合わせによって電気泳動の解析結果がどのようなバリエーションになるかが測定され、統計データとして得られたことになる。なお、多種アンプリコン電気泳動結果データは、STR回数とは関連付けられるものであるが、複数のSTR回数を有する対照用のサンプルとしてのみ用いられるものであって、実在の個体とは直接的には関連付けられるものではない。

[0049] 個体識別対象である新サンプル107は、新サンプル電気泳動解析部108によって、電気泳動法を用いて解析される。ここで新サンプル電気泳動解析部108は、上述した電気泳動解析部104とは、同一またはほぼ同等の解析性能を有するものである。単一の電気泳動解析装置を電気泳動解析部104及び新サンプル電気泳動解析部108として共用するようにしてもよい。新サンプル107を新サンプル電気泳動解析部108によって解析することにより結果として新サンプル電気泳動結果データ109が得られるから、新サンプル結果データ解析部110は、多種アンプリコン・データ蓄積部106に蓄積されている多種アンプリコン電気泳動結果データ105のうち、新サンプル電気泳動結果データ109と類似するものを検索することによって、新サンプル107のSTR回数を解析し、その解析結果を新サンプル解析結果111として出力する。

[0050] 第1の実施形態では、上述したように、低精度電気泳動解析部505の読み取り精度は、高精度電気泳動解析装置502に比べて同程度か低いことを想定している。その後、個体識別部506(図1参照)は、新サンプル解析結果111の複数のSTR回数のセットと識別子付きサンプル解析データ蓄積部504の各エントリのSTR回数のセットとに重なりがある識別子を検索し、個体識別結果507とする。個体識別結果507には、個体識別子が、1つ含まれている場合または複数含まれている場合があり、さらに、

個体識別子が全く含まれない場合もある。

[0051] 以下、読み取り精度が十分でない低精度電気泳動解析部505を用いて新サンプル107を解析しても、十分な精度で個体識別結果が得られることについて説明する。

[0052] 第1の実施形態では、低精度電気泳動解析部505の新サンプル電気泳動解析部108もしくは電気泳動解析部104として、従来から個体識別に用いられている電気泳動解析装置と比較して簡便なものを用いることを想定している。この状況では、PCRによって1種類のマイクロサテライト領域を増幅しようとしても、DNAの不完全な複製によって多少違うアンプリコンが生成される場合があることや、同じサイズのアンプリコンであってもキャピラリー中での移動の際に拡散の影響を受けることなどから、1種類のマイクロサテライト領域からのサンプルを電気泳動させた場合であっても、キャピラリー他端に到達する際にはその到達時間が前後することになる。その結果、電気泳動結果において、到達時間に対して、アンプリコンが幅をもって分布するようになり、その濃度を観察すると山型の波形になる。この現象は、電気泳動の媒体としてキャピラリーなどを用いる場合にもゲル板を用いる場合でも発生し、アンプリコンのサイズ解析の精度を低下させる。

[0053] ヘテロ接合のヒトのDNAマイクロサテライト領域において同程度のサイズのアンプリコンが2種類ある場合、電気泳動の結果として、それぞれが別個のアンプリコンに対応する2つの山型波形が重なり合ってしまい、見かけ上、1つの山型波形として観察されることがある。しかしながら、2種類のアンプリコンのサイズが、山型波形の幅に比べて、大きく異なる場合には、2つの山型波形は別の部位に発生するため、重なることは起こらず、各々の山型波形の位置や濃度は各アンプリコンのサイズに起因するものとして、正しく解析される。

[0054] 近接している波形が分離されなくなるのは、電気泳動時の拡散等に起因する解像度の問題であり、このような現象の発生を抑えた解像度の高い装置、すなわち読み取り精度の高い電気泳動解析装置では、解析結果における山型波形の幅が狭くなるため、2種類のアンプリコンのサイズがほぼ同じでも、波形を分離して観測することができる。

[0055] 2種類のアンプリコンの濃度がほとんど同じである場合、2種類のアンプリコンによっ

て合成によって発生する山型波形のピーク位置は、それぞれのアンプリコンの電気泳動によって生成されるであろうと考えられる2つの山型波形のそれぞれのピークの中間に位置する。例えば、2種類のアンプリコンが、5のSTR回数すなわち30塩基と、5. 2のSTR回数すなわち32塩基とであった場合は、31塩基にピークを有する山型波形として観察される。2bpの読み取り誤差があるとする、このサンプルは、30～32塩基として認識される。したがって、STR回数が5のものか、STR回数が5. 2のものか、どちらかのものであるかの判定がつかない。

[0056] この状況を説明するために、山型波形の形状をガウス分布で近似し、STR回数が5のものと、STR回数が5～8の2種類のアンプリコンを含んだDNAサンプルを混合したものを解析した時の結果をシミュレーションした結果を図3に示す。図3において、x軸はDNAのサイズである。ここでSTR繰り返しの単位が4bp (塩基対)であることを仮定すると、(5, 5)混合サンプル1001は(5, 5)の混合サンプルの山型波形の形状を示している。ここで(x, y)混合サンプルとは、STR回数がxのサンプルとSTR回数がyのサンプルとを混合していることを意味する。x=yであればホモ接合であり、x≠yであればヘテロ接合である。図3のシミュレーションから分かるように、拡散等の影響によりアンプリコンを電気泳動させたときの到達時間にばらつきが生じるため、山型波形の幅が広がっており、STR回数が1違うところにも影響が出ている。以下、STR回数が1異なることを、「1STRの違い」のように表記する。

[0057] このように1STR違うところに影響が出る程度、山形波形の幅が広がるような解像度では、例えば、(5, 7)混合サンプル1005、(5, 7. 2)混合サンプル1006、(5, 8)混合サンプル1007では、ヘテロ結合のDNAサンプルであるとはっきり認識できる。このため、ここで示したような解像度では、2STR以上の違いがある混合サンプルでは、ヘテロ接合であると正しく認識できると思われる。なお、1. 2STR(すなわち6bp)違いの(5, 6. 2)混合サンプル1004でも、解像度が2bp以下であれば、正しくヘテロ接合が認識できると思われる。1STR(すなわち4bp)違いの(5, 6)混合サンプル1003は、他に類似の山型波形の波形がなく、ピーク位置もずれていないことから、ピーク位置を手がかりにすれば、正しく認識できると思われる。

[0058] しかしながら、このような解像度において、2つのアンプリコンの違いが2bpである(5

、5. 2)混合サンプル1002では、単峰性の山形波形しか得られていない。この山形波形は、ホモ接合((5, 5)混合サンプル1001)の山形波形と形状が似ているために、山形波形の形状からは、ヘテロ接合なのか、ホモ接合なのかを特定することができない。ホモ接合である(5, 5)混合サンプル1001のピークと、ヘテロ接合である(5, 5. 2)混合サンプル1002のピークとは、位置が約1bp程度ずれているので、DNAの長さに対して精度の高い解析ができる場合には両者の違いを認識することは可能であるが、DNAの長さに対する精度が低い場合には、正しい認識を行うことは難しい。

[0059] 結論として、上記のシミュレーション条件では、山形波形の形状の特異性と電気泳動の正確性を考え合わせると、1STR以上の違いは区別できるが、1STR未満の違いでは、ホモ接合かヘテロ接合かの区別ができない、ということになる。第1の実施形態では、低精度電気泳動解析部505内の新サンプル電気泳動解析部108もしくは電気泳動解析部104は、従来の電気泳動装置と比較して読み取り精度が低いとしている。したがって、高い精度での解析に基づく結果を蓄積した識別子付きサンプル解析データ蓄積部504を、このような低精度電気泳動解析部505によって得られたデータに基づいて直接検索し、照合を行ったとしても、完全一致するデータを検索することができないことがあり得る。これは、2bp程度やそれよりも悪い精度を有する読み取り装置で解析を行うと、ヒトにおいてSTR回数として出現し得るxxとxx. 2などを区別できないためである。

[0060] しかしながら、上記の例のように解像度が2dp程度であるとして、測定された塩基長が30塩基程度であれば、STR回数は5もしくは5. 2であると判断することはできる。そこで、このような場合、5または5. 2のSTR回数として取り扱ってデータベースを照合することにより、正しいSTR回数を含むエントリの検索が可能となる。もっともこの場合、真のSTR回数が5であるとする、STR回数が5. 2のものも余分に検索されてしまうことが問題となる。すなわち、解像度の悪い機器で測定された新サンプル電気泳動結果データ109が得られた場合に、STR回数として可能性があるものを考慮して、高い解像度のデータから生成されたデータベースを検索した場合、新サンプル107に含まれているSTR回数のエントリを含んで検索できるが、間違ったエントリも余分に検索されてしまう。このように余分な結果も得られてしまう、という問題は、後述する他

の手法、例えば多種のローカス情報を用いる方法などと組み合わせることによって、実際の使用の場面においては、実はほとんど問題にならない。

[0061] 以下、余分な結果が得られることが個体識別のための障害とはなりえないことについて説明する。

[0062] ここでは、説明のため、状況として起こりえるSTR回数の種類が、4, 5, 5. 2, 6, 7の5通りであり、それぞれの塩基長は、26, 30, 32, 34, 38塩基であるとする。

[0063] 例題として、DNAサンプル(新サンプル107)のアンプリコンの真のSTR回数が(5, 5. 2)であるとする。低い解像度(2bp程度)で新サンプル電気泳動結果データが得られているのであれば、1bpなどのような高い精度で読み取りを行って作成されたデータベース中の{(5, 5), (5, 5. 2), (5. 2, 5. 2), (5. 2, 6)}の4種類のSTR回数のうちのどれと一致するのか区別できない場合がある。すなわち、1bpの解像度を有する機器を用いた場合に比較して、STR回数を特定する能力が低下する。しかしながら、解析したDNAサンプル(新サンプル107)は、上述の4種類以外のものではないことは、認識できる。

[0064] 同様に、新サンプルのアンプリコンの真のSTR回数が(5. 2, 5. 2)であるとする。このときは、データベースとの照合の結果、{(5, 5), (5, 5. 2), (5. 2, 5. 2), (5. 2, 6), (6, 6)}のいずれかであるとの認識がなされることになる。真のSTR回数(ここでは、(5. 2, 5. 2))は含まれてはいるものの、真のもの以外を含めて多めに検索される。ここで、山型波形のピークの読み取り誤差、つまり正確性は、およそ1塩基として考える。言い換えれば、34塩基と読み取られたとしても、実際のDNAサイズとしては、33, 34, 35塩基の可能性があるということである。解像度は2bp程度であるとする。つまり、2bpだけ異なるアンプリコンを有するヘテロ接合の場合、2つの山型波形が重なり合ってしまい、1つの山型波形として読み取られる、として考える。例えば、図3に示した例のように、塩基長が30bpと32bpのアンプリコンからなるヘテロ接合の場合には、31bp辺りにピークがある単峰型の山型波形として読み取られる。ヘテロ接合の場合、それぞれ独立して塩基長に関して読み取り誤差が発生するとしているが、アンプリコンサイズの差が4bp以下の場合、電気泳動結果のグラフにおいて隣接するアンプリコンなので、これらの2つのアンプリコンの相対的な読み取り誤差は、な

いとして考える。

[0065] 表2は、2種類のアンプリコンの混合物の電気泳動結果の例を示すものであって、上記の状況において2種類のアンプリコンの混合物が、電気泳動によってどのように解析されるかを示したものである。行頭の番号が**であるのは、観測されるDNAサイズの対が同じであるものが存在することを示す。例えば、真のSTR回数が(4, 5)の場合と(4, 5. 2)の場合では、電気泳動解析装置によって、両方とも(25bp, 31bp)の組み合わせと解析される可能性があることを示している。

[0066] [表2]

表2 2種類のアンプリコンの混合物の電気泳動結果の例

1: (25bp, 25bp) ← (26bp, 26bp)=(4, 4)	42: (30bp, 33bp) ← (30bp, 34bp)=(5, 6)
2: (25bp, 29bp) ← (26bp, 30bp)=(4, 5)	43: (30bp, 34bp) ← (30bp, 34bp)=(5, 6)
3: (25bp, 30bp) ← (26bp, 30bp)=(4, 5)	44: (30bp, 35bp) ← (30bp, 34bp)=(5, 6)
4: (25bp, 31bp) ← (26bp, 30bp)=(4, 5)	45: (30bp, 37bp) ← (30bp, 38bp)=(5, 7)
**:(25bp, 31bp) ← (26bp, 32bp)=(4, 5.2)	46: (30bp, 38bp) ← (30bp, 38bp)=(5, 7)
5: (25bp, 32bp) ← (26bp, 32bp)=(4, 5.2)	47: (30bp, 39bp) ← (30bp, 38bp)=(5, 7)
6: (25bp, 33bp) ← (26bp, 32bp)=(4, 5.2)	48: (31bp, 31bp) ← (30bp, 30bp)=(5, 5)
**:(25bp, 33bp) ← (26bp, 34bp)=(4, 6)	**:(31bp, 31bp) ← (30bp, 32bp)=(5, 5.2)
7: (25bp, 34bp) ← (26bp, 34bp)=(4, 6)	**:(31bp, 31bp) ← (32bp, 32bp)=(5.2, 5.2)
8: (25bp, 35bp) ← (26bp, 34bp)=(4, 6)	49: (31bp, 33bp) ← (30bp, 34bp)=(5, 6)
9: (25bp, 37bp) ← (26bp, 38bp)=(4, 7)	50: (31bp, 34bp) ← (30bp, 34bp)=(5, 6)
10: (25bp, 38bp) ← (26bp, 38bp)=(4, 7)	51: (31bp, 35bp) ← (30bp, 34bp)=(5, 6)
11: (25bp, 39bp) ← (26bp, 38bp)=(4, 7)	52: (31bp, 37bp) ← (30bp, 38bp)=(5, 7)
12: (26bp, 26bp) ← (26bp, 26bp)=(4, 4)	**:(31bp, 37bp) ← (32bp, 38bp)=(5.2, 7)
13: (26bp, 29bp) ← (26bp, 30bp)=(4, 5)	53: (31bp, 38bp) ← (30bp, 38bp)=(5, 7)
14: (26bp, 30bp) ← (26bp, 30bp)=(4, 5)	**:(31bp, 38bp) ← (32bp, 38bp)=(5.2, 7)
15: (26bp, 31bp) ← (26bp, 30bp)=(4, 5)	54: (31bp, 39bp) ← (30bp, 38bp)=(5, 7)
**:(26bp, 31bp) ← (26bp, 32bp)=(4, 5.2)	**:(31bp, 39bp) ← (32bp, 38bp)=(5.2, 7)
16: (26bp, 32bp) ← (26bp, 32bp)=(4, 5.2)	55: (32bp, 32bp) ← (30bp, 32bp)=(5, 5.2)
17: (26bp, 33bp) ← (26bp, 32bp)=(4, 5.2)	**:(32bp, 32bp) ← (32bp, 32bp)=(5.2, 5.2)
**:(26bp, 33bp) ← (26bp, 34bp)=(4, 6)	**:(32bp, 32bp) ← (32bp, 34bp)=(5.2, 6)
18: (26bp, 34bp) ← (26bp, 34bp)=(4, 6)	56: (32bp, 37bp) ← (32bp, 38bp)=(5.2, 7)
19: (26bp, 35bp) ← (26bp, 34bp)=(4, 6)	57: (32bp, 38bp) ← (32bp, 38bp)=(5.2, 7)
20: (26bp, 37bp) ← (26bp, 38bp)=(4, 7)	58: (32bp, 39bp) ← (32bp, 38bp)=(5.2, 7)
21: (26bp, 38bp) ← (26bp, 38bp)=(4, 7)	59: (33bp, 33bp) ← (32bp, 32bp)=(5.2, 5.2)
22: (26bp, 39bp) ← (26bp, 38bp)=(4, 7)	**:(33bp, 33bp) ← (32bp, 34bp)=(5.2, 6)
23: (27bp, 27bp) ← (26bp, 26bp)=(4, 4)	**:(33bp, 33bp) ← (34bp, 34bp)=(6, 6)
24: (27bp, 29bp) ← (26bp, 30bp)=(4, 5)	60: (33bp, 37bp) ← (32bp, 38bp)=(5.2, 7)
25: (27bp, 30bp) ← (26bp, 30bp)=(4, 5)	**:(33bp, 37bp) ← (34bp, 38bp)=(6, 7)
26: (27bp, 31bp) ← (26bp, 30bp)=(4, 5)	61: (33bp, 38bp) ← (32bp, 38bp)=(5.2, 7)
**:(27bp, 31bp) ← (26bp, 32bp)=(4, 5.2)	**:(33bp, 38bp) ← (34bp, 38bp)=(6, 7)
27: (27bp, 32bp) ← (26bp, 32bp)=(4, 5.2)	62: (33bp, 39bp) ← (32bp, 38bp)=(5.2, 7)
28: (27bp, 33bp) ← (26bp, 32bp)=(4, 5.2)	**:(33bp, 39bp) ← (34bp, 38bp)=(6, 7)
**:(27bp, 33bp) ← (26bp, 34bp)=(4, 6)	63: (34bp, 34bp) ← (32bp, 34bp)=(5.2, 6)
29: (27bp, 34bp) ← (26bp, 34bp)=(4, 6)	**:(34bp, 34bp) ← (34bp, 34bp)=(6, 6)
30: (27bp, 35bp) ← (26bp, 34bp)=(4, 6)	64: (34bp, 37bp) ← (34bp, 38bp)=(6, 7)
31: (27bp, 37bp) ← (26bp, 38bp)=(4, 7)	65: (34bp, 38bp) ← (34bp, 38bp)=(6, 7)
32: (27bp, 38bp) ← (26bp, 38bp)=(4, 7)	66: (34bp, 39bp) ← (34bp, 38bp)=(6, 7)
33: (27bp, 39bp) ← (26bp, 38bp)=(4, 7)	67: (35bp, 35bp) ← (34bp, 34bp)=(6, 6)
34: (29bp, 29bp) ← (30bp, 30bp)=(5, 5)	68: (35bp, 37bp) ← (34bp, 38bp)=(6, 7)
35: (29bp, 33bp) ← (30bp, 34bp)=(5, 6)	69: (35bp, 38bp) ← (34bp, 38bp)=(6, 7)
36: (29bp, 34bp) ← (30bp, 34bp)=(5, 6)	70: (35bp, 39bp) ← (34bp, 38bp)=(6, 7)
37: (29bp, 35bp) ← (30bp, 34bp)=(5, 6)	71: (37bp, 37bp) ← (38bp, 38bp)=(7, 7)
38: (29bp, 37bp) ← (30bp, 38bp)=(5, 7)	72: (38bp, 38bp) ← (38bp, 38bp)=(7, 7)
39: (29bp, 38bp) ← (30bp, 38bp)=(5, 7)	73: (39bp, 39bp) ← (38bp, 38bp)=(7, 7)
40: (29bp, 39bp) ← (30bp, 38bp)=(5, 7)	
41: (30bp, 30bp) ← (30bp, 30bp)=(5, 5)	
**:(30bp, 30bp) ← (30bp, 32bp)=(5, 5.2)	

(観測されるDNAサイズの対) ← (真のサイズの対) = (真のSTR回数) のフォーマット
 行頭の番号が ** であるのは、同じ (観測されるDNAサイズの対) が存在する。

[0067] FBIによる複合DNAインデックスシステム(CODIS)等で使用されているローカス

のDNAデータを調べてみると、xx. 1、xx. 2またはxx. 3と記載される、STR回数が非整数であるものの発現回数が少ない。以下の説明において、xx. 1、xx. 2およびxx. 3のいずれかのことを、xx. {1, 2, 3}と表記する。例えば、上述した表1に示したローカスFGAの多様性を示すデータでは、STR回数には18通りのSTR回数が存在し、そのうち4種類のみが非整数の回数のも、ここではxx. 2のタイプのものである。

[0068] 上述したBudowleらの論文(非特許文献1)中で公表されている“dnaloci. txt”には、アフリカ系アメリカ人におけるローカスFGAの多様性を示すデータだけでなく、他のローカス、他の人口グループ(population group)における同様の多様性に関するデータを含んでいる。以下、Budowleらの論文に付属する生データ“dnaloci. txt”を用いて、第1の実施形態の個体識別装置が個体を正確に識別できることを説明する。表3は、以下の説明において用いるデータの概要を示したものであって、ローカスごとのSTR回数と出現頻度との関係を示している。ここでは、4bp程度の精度で解析することが難しいと思われるSTR回数のものに限って示している。

[0069] [表3]

表 3 ローカスごとのSTR回数と出現頻度

ローカス	xx.{1,2,3} の STR 回数	出現頻度
CSF1PO	10.3	0.11
D18S51	13.2	0.17
D18S51	14.2	0.01
D18S51	15.2	0.01
D18S51	21.2	0.02
D21S11	24.2	0.36
D21S11	24.3	0.03
D21S11	29.2	0.12
D21S11	30.2	2.79
D21S11	30.3	0.01
D21S11	31.2	8.72
D21S11	32.1	0.01
D21S11	32.2	10.44
D21S11	33.2	3.42
D21S11	34.2	0.19
D21S11	35.2	0.12
D3S1358	15.2	0.02
D7S820	10.1	0.01
D7S820	11.3	0.01
FGA	17.2	0.09
FGA	18.2	0.32
FGA	19.2	0.10
FGA	20.2	0.17
FGA	21.2	0.06
FGA	22.2	0.71
FGA	22.3	0.01
FGA	23.2	0.16
FGA	24.2	0.01
FGA	24.3	0.01
FGA	30.2	0.08
THO1	8.3	0.11
THO1	9.3	22.25
総合		50.65% (13 ローカス:1300%中)

4 b p 程度の精度で解析することが難しいと思われるSTR
回数のものに限る

[0070] ここで用いるデータには、アメリカ合衆国における6つの人口グループ(アフリカ系ア

メリカ人 (AFRICAN AMERICAN)、米国カフカス人 (U. S. CAUCASIAN)、南部ヒスパニック (SOUTHWESTERN HISPANIC)、バハマ人 (BAHAMIAN)、ジャマイカ人 (JAMAICAN)、トリニダード人 (TRINIDADIAN)) についてのデータを含んでいる。以下、平均的な能力を知るために、アメリカ合衆国における人口グループの構成比として、アフリカ系アメリカ人が25%、米国カフカス人が45%、南部ヒスパニックが20%、残り10%がバハマ人、ジャマイカ人、トリニダード人であるとする。仮に、バハマ人、ジャマイカ人、トリニダード人の割合をそれぞれ4, 4, 2%と仮定し、データを作成して、統計的な解析を進める。また、上記生データ中で“<xx”, “>xx”などの表記は、xxよりもSTR回数が小さいものや大きいものの確率を示すものであるが、処理が煩雑になり、また、出現回数が少ないために、省略した。

- [0071] xx. {1, 2, 3}のタイプのSTR回数は、7つのローカス (CSF1PO, D18S51, D21S11, D3S1358, D7S820, FGA, THO1) に含まれ、合計32種類である。STR回数の種類を全ローカスにわたって総計を求めると163種類であるので、種類の比率としては、xx. {1, 2, 3}型のデータは、19%となる。xx. {1, 2, 3}の出現比率は、3.85%である。
- [0072] CODIS自体は13種類のローカスを用いるものであって、これらのローカスにおいてxx. {1, 2, 3}の出現頻度は、合計すると50.65%となる。なお全13ローカスなので、頻度の総計は1300%となる。頻度データに注目すると、ローカスD21S11ではxx. {1, 2, 3}の頻度が多いが、それ以外のローカスでは、xx. {1, 2, 3}頻度が非常に少なく、xx. {1, 2, 3}型のSTR回数にあまり遭遇しないといえる。つまり、xx. {1, 2, 3}とxxの区別ができない装置を用いて、STR回数がおおよそ18と判定された場合、真のSTR回数は18もしくは18.2であるが、18.2の発現回数が0.014と少ないので、18と18.2を1つにまとめて判断しても、判別能力がほとんど変わらない、と予想される。なお、正確な見積もりについては後述する。
- [0073] 認識能力を正確に見積もるために、各人の各STR回数が独立して出現するものとする。その時に偶然に2人のSTR回数が一致する確率について考える。これは、識別力 (Discrimination Power) と呼ばれる値であり、ある解析手法がどのくらいの認識能力を有するかを示す量である。偶然に2人のSTR回数が一致する確率が低い

程、認識能力が高いとされる。

- [0074] 表3に示したアメリカ合衆国での6つの人口グループのデータの混合の例で考えると、FGAの1つのSTR回数が25である確率は0.100、24である確率が0.186となる。したがって、ランダムに選んだ1人のFGAが(24, 25)である確率は、 $0.100 \times 0.186 \times 2$ となる。ここで、ランダムに選んだ2人のFGAが、偶然にともにFGAが(24, 25)である確率は、 $(0.100 \times 0.186 \times 2)^2$ となる。FGAを用いた場合の識別力は、全てのFGAに関するSTR回数の組み合わせについてランダムに選んだ2人がともに同じSTR回数である確率であるので、以下の総和で求めることができる。ただし、注意しないといけないのは、上記のヘテロ接合の例とは違うホモ接合の場合であり、ホモ接合の場合には、出現確率は、例えば、(24, 24)の場合は、 0.186×0.186 のように、ヘテロ接合の場合と異なって、2倍にするための項が出てこない。

- [0075] ヘテロ接合とホモ接合のそれぞれについての出現確率は以下ようになる。

- [0076] ヘテロ接合の場合：

- [0077] [数1]

$$\sum_{i, j, i \neq j} ((\text{STR回数が } i \text{ の出現確率}) \times (\text{STR回数が } j \text{ の出現確率}) \times 2)^2$$

- [0078] ホモ接合の場合：

- [0079] [数2]

$$\sum_i (\text{STR回数が } i \text{ の出現確率})^4$$

- [0080] 正確な見積もりとしては、上述したように、STR回数の組み合わせが5人以下の場合は5人いたとして確率を計算するが、このデータでは、6通りの人口グループのデータをミックスして仮想しているために、そこまでの正確性は必要ないものとして、そのような計算を行うことは省略した。

- [0081] ローカスFGAの場合には、識別力、すなわち2人のSTR回数が偶然にマッチする確率は、0.30391となる。CODISで用いられている他のローカスについて、識別力は表4に示すようになる。表4は、ローカスごとの識別力と、13種類のローカスすべてを用いたときにランダムに選んだ2人のSTR回数のすべてが一致する確率とを示して

いる。

[0082] [表4]

表4 ローカスごとの識別力と、13種類のローカスのすべてを用いたときにランダムに選んだ2人のSTR回数の全てが一致する確率

ローカス	識別力	$-\log_{10}$ 表記	種類数
D13S317	0.080887	(1.092124)	9
CSF1PO	0.104326	(0.981606)	11
D16S539	0.075676	(1.121040)	8
D18S51	0.025047	(1.601245)	19
D21S11	0.037495	(1.426024)	23
D3S1358	0.085779	(1.066618)	9
D5S818	0.126088	(0.899327)	10
D7S820	0.070993	(1.148787)	11
D8S1179	0.066310	(1.178421)	11
FGA	0.030391	(1.517258)	26
TH01	0.078208	(1.106751)	8
TPOX	0.143799	(0.842243)	8
vWA	0.061754	(1.209334)	10
総合	6.444986×10^{-16}	(15.190778)	163

[0083] 表4の左から3カラム目のカッコ内の数字は、「識別力(すなわち、偶然にマッチする確率)」の $-\log_{10}$ 表記である。したがって、カッコ内の数字が1.0の場合、10人に1人の割合で偶然にSTR回数が合致することを意味する。最後のカラムの数字は、ローカスごとのSTR回数の種類数である。STR回数の種類数が多いと、偶然にマッチする確率は低くなる。もっとも、STR回数の種類数が同じであったとしても、STR回数の分布に偏りがあるあるため、「偶然にマッチする確率」が同じになるわけではない。

[0084] 上記の13通りのローカスにおける全163種類のSTR回数を用いた時、ランダムに選んだ2人が偶然に一致する確率は、ローカスごとの識別力の積であるから、 $6.444986 \times 10^{-16} (= 10^{-15.190778})$ となり、 $1.551594 \times 10^{+15}$ 人($= 1 / 6.444986 \times 10^{-16}$)に1人の確率で合致することになる。

[0085] 第1の実施形態では、上述したような条件で、低精度電気泳動解析部505内の新サンプル電気泳動解析部108もしくは電気泳動解析部104を用いて、DNAサンプルすなわち新サンプル107の解析結果を新サンプル解析結果111として得て、この

解析結果に基づいてデータベースを検索することにより、DNAサンプルに含まれているSTR回数のエントリを検索できる。この場合、上述したように、間違っただエントリも余分に検索されてしまう。そこで、間違っただエントリも余分に検索されてしまうような状況において「ランダムに選んだ2人が偶然に一致する確率」を考える。

[0086] 低精度電気泳動解析部505を用いたことによる低い精度読み取り精度のために、 x と xx 、 2 と $xx+1$ の区別がつかないとする。つまり、これらが「1種類のSTR回数として認識される」といった状況を考える。表5は、このような状況における識別力を、電気泳動における解析精度が1bp程度である場合の識別力と対比して示したものであり、低精度電気泳動解析装置を用いた場合における、ローカスごとの識別力と、13種類のローカスすべてを用いたときにランダムに選んだ2人のSTR回数のすべてが一致する確率とを示している。

[0087] [表5]

表5 低精度電気泳動解析装置を用いた場合における、ローカスごとの識別力と、13種類のローカスのすべてを用いたときにランダムに選んだ2人のSTR回数がすべて一致する確率

ローカス	高精度電気泳動解析装置		低精度電気泳動解析装置		差
	識別力	$-\log_{10}$ 表記	識別力	$-\log_{10}$ 表記	
D13S317	0.080887	(1.092124)	0.080887	(1.092124)	0
CSF1PO	0.104326	(0.981606)	0.105320	(0.977490)	0.004116
D16S539	0.075676	(1.121040)	0.075676	(1.121040)	0
D18S51	0.025047	(1.601245)	0.025463	(1.594093)	0.007152
D21S11	0.037495	(1.426024)	0.059666	(1.224274)	0.20175
D3S1358	0.085779	(1.066618)	0.085946	(1.065774)	0.000844
D5S818	0.126088	(0.899327)	0.126088	(0.899327)	0
D7S820	0.070993	(1.148787)	0.071117	(1.148027)	0.00076
D8S1179	0.066310	(1.178421)	0.066310	(1.178421)	0
FGA	0.030391	(1.517258)	0.034277	(1.465002)	0.052256
THO1	0.078208	(1.106751)	0.129458	(0.887872)	0.218879
TPOX	0.143799	(0.842243)	0.143799	(0.842243)	0
vWA	0.061754	(1.209334)	0.061754	(1.209334)	0
総合		(15.190778)		(14.75020)	0.485758

[0088] 表5の1番左のカラムはローカス名を示し、左から2番目と3番目のカラムは、解析精度が1bp程度である高精度電気泳動解析装置を用いたときの識別力とその $-\log_{10}$

表記を示している。なお、左から2番目と3番目のカラムにおける数値は、表4に示したものと同一である。低精度電気泳動解析装置と記載されている表5の左から4番目のカラムは、上述したような解像度が4bpである低精度電気泳動解析部505を用いたことによる識別力を示し、5番目のカラムは、4番目のカラムの識別力を $-\log_{10}$ 表記したものである。表5の右端のカラムは、3番目のカラムと5番目のカラムとの間の差を示している。右端のカラムの数値をcとして、その10のべき乗、すなわち 10^c を考えると、低精度電気泳動解析部505を用いたことにより、識別力が 10^c だけ低下することになる。

[0089] ローカスD12S317, D16S539, D5S818, D8S1179, TPOX, vWAでは、STR回数においてxx. {1, 2, 3}のタイプのものが存在しないことにより、電気泳動における精度が上述したように低下したとしても、異なるSTR回数を正確に識別することができ、したがって、識別力の低下は起こらない。このことは、表5において、右端のカラムの値が0であることによって示されている。これに対し、ローカスD21S11, TH01では、右端のカラムに示された値が、 $-\log_{10}$ 表記での差として、0.2程度であることから、 $0.63 (=10^{-0.2})$ 倍に、識別力が低下することが分かる。

[0090] 表5の最下行に示すように、ここで示した13ローカスの全部のSTR回数を用い、上述したような低精度電気泳動解析部505を使用したとすると、ランダムに選んだ2人が偶然に一致する確率は、 $1.972332 \times 10^{-15} (=10^{-14.705020})$ となり、 $5.07014 \times 10^{+14}$ 人に1人の確率で合致することになる。ここで用いている13ローカスは、CODISで用いられている13ローカスと同じである。

[0091] これに対し、従来用いられているように、解析精度が1dpである電気泳動解析装置を用い、13ローカスの全部のSTR回数を用いた時、ランダムに選んだ2人が偶然に一致する確率は、 $6.444986 \times 10^{-16} (=10^{-15.190778})$ となり、 $1.551594 \times 10^{+15}$ 人 $(=1/6.444986 \times 10^{-16})$ に1人の確率で合致することになる。つまり、低精度電気泳動解析部505を用いたことによって、CODISなどで使用されている13ローカスでは、識別力が $1/(1.551594 \times 10^{+15})$ から $1/(5.07014 \times 10^{+14})$ に悪化することが分かる。すなわち、0.3267699倍に識別力が悪化することが分かる。

[0092] 平均的に、検索条件としてSTR回数を1回使用することに対象を1/10程度の絞り

込めることを考えると、CODISの13ローカスを用いて個体認識を行う場合、分解能が1bpである電気泳動解析装置を用いるときと分解能が4bpである電気泳動解析装置を用いるときとの認識能力差(0.32677699倍)は、「ある1つのローカスの情報を使わなかった」と同程度か、それ以下の認識能力の低下とみなすことができる。

[0093] 識別力を用いて、「目の前にあるサンプルのSTR回数がデータベースのうちのあるエントリと一致したとき、それはどのくらい起こりえるか？」の指標を計算することができる。この数字は、鑑定の証明力などを立証するために法廷などで使用される数字である。識別力は「ランダムに選んだ2人がともに同じ個体の遺伝子型である確率」であるのに対し、この指標は、「目の前にあるサンプルのSTR回数がデータベースのうちのエントリと一致して、他のn人のサンプルのSTR回数はデータベースと一致しない」確率を示している。ここで、「ランダムに選んだ2人がともに同じ個体の遺伝子型である確率」をpとすると、「データベースと一致しない」確率は、 $1-p$ となる。それがn人なので、n人の全てがデータベースと一致しない確率は、 $(1-p)^n$ となる。このようなことそれ自体が減多に起こらないよう、1%以下の危険率で計算すると、 $(1-p)^n \geq 1-0.01$ が得られる。

[0094] この状況をアメリカ合衆国規模の人口に当てはめると、nは3億人となり、 $(1-p)^n \approx 1-np$ と近似できることに注意すると、 $p \leq 3.33 \times 10^{-11}$ となる。

[0095] この数字と、1bpの解像度で読み取った時のランダムに選んだ2人のSTR回数が偶然に一致する確率の $6.444986 \times 10^{-16} (=10^{-15.190778})$ と、4bpの解像度で読み取った時のランダムに選んだ2人のSTR回数が偶然に一致する確率の $1.972332 \times 10^{-15} (=10^{-14.705020})$ とを比較する必要がある。

[0096] $(1-p)^n \geq 1-\text{危険率}$

$$1-np \geq 1-\text{危険率}$$

$$\text{危険率} \geq np$$

に注意すると、1bpの解像度で読み取った時の危険率と、上記の条件に示すような低い解像度(4bp)での危険率は、それぞれ、 $1.933496 \times 10^{-7} (=6.444986 \times 10^{-16} \times 3 \times 10^8)$ 、 $5.916996 \times 10^{-7} (=1.972332 \times 10^{-15} \times 3 \times 10^8)$ となる。低い解像度の場合の危険率は、1bpの場合の危険率に比べて約3倍大きい。

[0097] 以上まとめると、従来の高い解像度で解析した時には、「目の前にあるサンプルのSTR回数がデータベースのうちのエン트리と一致して、他のn人のサンプルのSTR回数はデータベースと一致しない」と言うことに対して、 $99.99998\% (=1.0 - 1.933496 \times 10^{-7})$ の確からしさを表明できる。これに対し、低精度電気泳動解析部505の低い読み取り精度のために、xxとxx.2とxx+1の区別がつかない状況では、「目の前にあるサンプルのSTR回数がデータベースのうちのエン트리と一致して、他のn人のサンプルのSTR回数はデータベースと一致しない」と言うことに対して、 $99.99994\% (=1.0 - 5.916996 \times 10^{-7})$ の確からしさを表明するかの違いになる。つまり、「ランダムに選んだ2人がともに同じ個体の遺伝子型である確率」が、小数点第5位の値が多少変化するだけなので、実用上ほとんど問題にならないことがわかる。

[0098] 《第2の実施形態》

次に、本発明の第2の実施形態の個体識別装置を説明する。この個体識別装置は、図1に示した第1の実施形態の個体識別装置と同様のものであるが、低精度電気泳動解析部505の構成において異なっている。図4は、第2の実施形態の個体識別装置における低精度電気泳動解析部505の構成を示している。

[0099] 上述した第1の実施形態では、多種アンプリコン・データ蓄積部106に蓄積されるべきデータの作成において、全ての組み合わせの多種アンプリコン・サンプル103を用意したが、この第2の実施形態では、すべての組み合わせの多種アンプリコン・サンプル103を用意するのではなく、適当な組み合わせでSTR回数のDNAサンプル(選択サンプル102)を用意してこれらを混合して多種アンプリコン・サンプル103を生成し、電気泳動解析部104によって解析し、解析により得られた多種アンプリコン電気泳動結果データ105を多種アンプリコン・データ蓄積部106に保存する。この場合、アンプリコンの考えられる組み合わせのうち、実際に多種アンプリコン・サンプル103が生成されていないSTR回数の組み合わせが存在するが、このようなSTR回数の組み合わせに関しては、測定済みの多種アンプリコン・データ蓄積部106のデータから、シミュレーション方法などを用いた補間などにより、データを生成する。このため、第2の実施形態では、低精度電気泳動解析部505は、測定されて多種アンプリコン・データ蓄積部106に蓄積されたデータから、補間によってデータを生成する補間デ

ータ作成部201と、補間データ生成部201で生成されたデータを補間する補間データ蓄積部202と、を備えている。新サンプル結果データ解析部110は、新サンプル107を電気泳動で解析した結果である新サンプル電気泳動結果データ109と多種アンプリコン・データ蓄積部106に蓄積されたデータと補間データ蓄積部202に蓄積されたデータとを比較解析することによって、新サンプル107のSTR回数を推定し、新サンプル解析結果111として出力する。

[0100] 《第3の実施形態》

次に、本発明の第3の実施形態の個体識別装置を説明する。この個体識別装置は、図1に示した第1の実施形態の個体識別装置と同様のものであるが、低精度電気泳動解析部505の構成において異なっている。図5は、第3の実施形態の個体識別装置における低精度電気泳動解析部505の構成を示している。

[0101] 上述した第1の実施形態では、多種アンプリコン・データ蓄積部106に蓄積されるべきデータの作成において、全ての組み合わせの多種アンプリコン・サンプル103を用意したが、この第3の実施形態では、すべての組み合わせの多種アンプリコン・サンプル103を用意するのではなく、適当な組み合わせでSTR回数のDNAサンプル(選択サンプル102)を用意してこれらを混合して多種アンプリコン・サンプル103を生成し、電気泳動解析部104によって解析し、解析により得られた多種アンプリコン電気泳動結果データ105を多種アンプリコン・データ蓄積部106に保存する。この場合、アンプリコンの考えられる組み合わせのうち、実際に多種アンプリコン・サンプル103が生成されていないSTR回数の組み合わせが存在するが、第3の実施形態では、新サンプル結果データ解析部として、パラメータ推定機能を有するパラメータ推定機能付き新サンプル結果データ解析部301を使用する。

[0102] パラメータ推定機能付き新サンプル結果データ解析部301は、新サンプル107を新サンプル電気泳動解析部108で解析した結果である新サンプル電気泳動結果データ109に基づいて多種アンプリコン・データ蓄積部106内のデータを検索し、新サンプル電気泳動結果データ109を解析する際に、多種アンプリコン・データ蓄積部106に既に蓄積されているデータを用いて、STR回数の変化に基づく新サンプル電気泳動結果データ109の変化の様子をパラメータ化し、解析に用いる。パラメータ推

定機能付き新サンプル結果データ解析部301は、新サンプル電気泳動結果データ109のSTR回数を解析して新サンプル解析結果111を出力する。

[0103] 《第4の実施形態》

次に、本発明の第4の実施形態の個体識別装置を説明する。この個体識別装置は、図1に示した第1の実施形態の個体識別装置と同様のものであるが、低精度電気泳動解析部505の構成において異なっている。図6は、第4の実施形態の個体識別装置における低精度電気泳動解析部505の構成をしている。

[0104] 第1の実施形態では、多種アンプリコン・サンプル103を生成して電気泳動によって解析し、解析結果を多種アンプリコン・データ蓄積部106に格納するようにしていたが、第4の実施形態では、多種アンプリコン・サンプルを生成することなく1種アンプリコン・サンプルのままで電気泳動解析を行い、その電気泳動解析から、複数のアンプリコンを含むサンプルの解析結果を補間によって求めて蓄積し、蓄積された結果に基づいて新サンプル電気泳動結果データ109を解析することによって、新サンプル107に対する解析結果を新サンプル解析結果111として得るようにしている。

[0105] すなわち、第4の実施形態において、低精度電気泳動解析部505は、低精度電気泳動解析部505は、1種アンプリコン・サンプル保管部101と、1種アンプリコン・サンプル保管部101から選択されたDNAサンプル(選択サンプル102)を電気泳動によって解析する電気泳動解析部104と、電気泳動解析部104から出力される1種アンプリコン電気泳動結果データ401を蓄積する1種アンプリコン・データ蓄積部402と、1種アンプリコン・データ蓄積部402に蓄積されたデータに基づいて補間多種アンプリコン・データを作成する補間多種アンプリコン・データ作成部403と、作成された補間多種アンプリコン・データを蓄積する補間多種アンプリコン・データ蓄積部404と、新サンプル107を電気泳動によって解析する新サンプル電気泳動解析部108と、新サンプル電気泳動解析部108が出力する新サンプル電気泳動結果データ109に基づいて1種アンプリコン・データ蓄積部402及び／または補間多種アンプリコン・データ蓄積部404内を検索し、検索結果を新サンプル解析結果111として出力する新サンプル結果データ解析部110と、を備えている。

[0106] ここで、1種アンプリコン・サンプル保管部101は、それぞれが1種類ずつのアンプリ

コンを含んだDNAのサンプルである1種アンプリコン・サンプルを複数個保管するものであり、アンプリコン・サンプルごとにそれらのサンプルにおけるSTR回数も保持している。選択サンプル102は、1種アンプリコン・サンプル保管部101の中から1種のサンプルを選択したものである。選択サンプル102を電気泳動解析部104によって電気泳動によって解析した結果が1種アンプリコン電気泳動結果データ401があり、1種アンプリコン・データ蓄積部402には、1種アンプリコン電気泳動結果データ401と、その1種アンプリコン電気泳動結果データ401に対応するアンプリコンのSTR回数を組にして蓄積する。第4の実施形態では、1種アンプリコン・サンプル保管部101より全てのサンプルを選択してそれぞれ選択サンプル102とし、電気泳動の結果がどのようなバリエーションになるかを測定し、統計データとする。

[0107] 第4の実施形態では、複数のアンプリコンを含むサンプルが示すであろう電気泳動結果を補間法で求める。そのため、補間多種アンプリコン・データ作成部403は、1種アンプリコン・データ蓄積部402に蓄積されているデータから、シミュレーション方法などを用いてそのようなデータ、すなわち補間多種アンプリコン・データを作成し、その作成された補間多種アンプリコン・データを補間多種アンプリコン・データ蓄積部404に保管する。

[0108] 新サンプル結果データ解析部110は、新サンプル107を電気泳動で解析した結果である新サンプル電気泳動結果データ109と補間多種アンプリコン・データ蓄積部404に蓄積されたデータとを比較解析することによって、新サンプル107のSTR回数を推定し、新サンプル解析結果111として出力する。

[0109] 第4の実施形態の別の例では、1種アンプリコン・サンプル保管部101に格納されている全ての1種アンプリコン・サンプルをそれぞれ選択サンプル102とするのではなく、1種アンプリコン・サンプル保管部101から一部のサンプルを選択し選択サンプル102としてもよい。

[0110] なお第4の実施形態では、新サンプル結果データ解析部110は、新サンプル電気泳動結果データ109を解析する際に、補間多種アンプリコン・データ蓄積部404に蓄積されたデータに加え、1種アンプリコン・データ蓄積部402に蓄積されている1種アンプリコン電気泳動結果データ401を使用するようにしてもよい。

[0111] 《第5の実施形態》

図7は、本発明の第5の実施形態の個体識別装置の構成を示している。この個体識別装置は、第1の実施形態のものと類似しているが、識別子付きサンプル501の解析に高精度電気泳動解析装置ではなく、低精度電気泳動解析部505を用いる点で、第1の実施形態のものと大きく異なっている。低精度電気泳動解析部505は、識別子付きサンプル501の各サンプルを、低い精度で解析し、その結果を低精度識別子付きサンプル解析結果601として出力する。低精度識別子付きサンプル解析結果601は、識別子付きサンプル501の個体ごとに、識別子とともに低精度識別子付きサンプル解析データ蓄積部602に蓄積される。

[0112] 個体認識の対象である新サンプル107は、第1の実施形態と同様に低精度電気泳動解析部505で解析され、その結果、新サンプル解析結果111が得られる。低精度個体識別部603は、低精度識別子付きサンプル解析データ蓄積部602を参照して、新サンプル解析結果111と共通のSTR回数を有するエントリを検索し、見つかったエントリを低精度個体識別結果604として出力する。

[0113] 電気泳動における解析精度が低下すれば、その分、識別力も低下するが、第1の実施形態において説明したように、識別力がどれだけ低下するかを評価することができる。識別子付きサンプル501に対する電気泳動解析の解像度の低下は、実効的に、識別子付きサンプル501に対する解析精度は低下しなかったものの新サンプル解析結果111における解析精度がさらに低下したものとして扱うことができる。そのように新サンプル解析結果111における解析精度がさらに低下したとして識別力を算出し、それが許容できるものであれば、識別子付きサンプル501を低精度電気泳動解析部505で解析したとしても、問題は生じない。

[0114] 《第6の実施形態》

図8は、本発明の第6の実施形態の個体識別装置の構成を示している。この個体識別装置は、第1の実施形態のものと類似しているが、新サンプル107を低精度電気泳動解析部505で解析して識別子付きサンプル解析データ蓄積部504内のデータと照合し、個体識別結果507を得た際に、その個体識別結果507によれば新サンプル107が複数の個体のSTR回数と一致する可能性がある場合、新サンプル107

を改めて高精度電気泳動解析装置502で解析する点で、第1の実施形態とは異なっている。新サンプル107を高精度電気泳動解析装置502で解析するとサンプル結果503が得られる。この個体識別装置は、高精度個体識別部701を備えており、高精度個体識別部701は、新サンプル107から得られたサンプル解析結果503に基づき、識別子付きサンプル解析データ蓄積部504内のエントリのうち、サンプル解析結果503と共通のSTR回数を有するエントリを検索し、検索結果を高精度個体識別結果702として出力する。

[0115] 《第7の実施形態》

図9は、本発明の第7の実施形態の個体識別装置の構成を示している。この個体識別装置は、第5の実施形態(図7参照)の場合と同様の手順によって低精度識別子付きサンプル解析結果601が蓄積される低精度識別子付きサンプル解析データ蓄積部602と、第6の実施形態(図8参照)の場合と同様の手順によって識別子付きサンプル501のサンプル解析結果503が蓄積される識別子付きサンプル解析データ蓄積部504と、を備えている。そしてこの個体識別装置では、第5の実施形態の場合と同様にして、新サンプル107をまず低精度電気泳動解析部505で解析して新サンプル解析結果111を取得し、低精度個体識別部603は、新サンプル解析結果111に基づいて低精度識別子付きサンプル解析データ蓄積部602を検索して、低精度個体識別結果604を出力する。この低精度個体識別結果604において、個体識別されたエントリが1つまたは複数ある場合には、今度は、第1の実施形態などの場合と同様に、先に求めた新サンプル解析結果111に基づき、個体識別部506によって識別子付きサンプル解析データ蓄積部504内を参照し、新サンプル解析結果111のSTR回数のセットと識別子付きサンプル解析データ蓄積部504の各エントリのSTR回数のセットとに重なりがある識別子を検索し、その検索結果を個体識別結果507とする。

[0116] この個体識別結果507において検索されたエントリが1または複数存在する場合には、識別子付きサンプル解析データ蓄積部504に完全一致するエントリが存在する可能性がある。これを調査するために、次に、第6の実施形態の場合と同様にして新サンプル107を高精度電気泳動解析装置502で解析し、サンプル解析結果503を

得る。高精度個体識別部701は、新サンプル107から得たサンプル解析結果503に基づいて、識別子付きサンプル解析データ蓄積部504中のエントリのうちサンプル解析結果503と共通のSTR回数を有するエントリを検索し、検索結果を高精度個体識別結果702として出力する。

[0117] 《第8の実施形態》

図10は、本発明の第6の実施形態の個体識別装置の構成を示している。この個体識別装置は、DNA解析を行うとともに、指紋などの他の個体識別情報（バイオメトリック情報）を用いて個体識別を行うものである。ここでは、DNAサンプル902と指紋サンプル906などを採取できる対象を新サンプル取得対象901として、新サンプル取得対象901の個体識別を行う場合を説明する。

[0118] この個体識別装置は、DNA解析に基づく個体識別部903と、識別子付きDNA解析データ蓄積部904と、指紋解析に基づく個体識別部907と、識別子付き指紋解析データ蓄積部908と、複数情報に基づく個体識別部910と、を備えている。ここで、DNA解析に基づく個体識別部903は、上述の各実施形態のいずれかにおける個体識別装置と同様のものであり、DNAサンプル902（上述の各実施形態における新サンプル107）を解析し、その解析結果に基づいて識別子付きDNA解析データ蓄積部904を検索し、検索結果をDNA解析に基づく個体識別結果905として出力する。識別子付きDNA解析データ蓄積部904は、上述の実施形態における識別子付きサンプル解析データ蓄積部504（あるいは低精度識別子付きサンプル解析データ蓄積部602）に該当するものであり、識別子付き、すなわち起源が明確である、DNAサンプルにおける解析結果を蓄積したものである。

[0119] 同様に、識別子付き指紋解析データ蓄積部908は、識別子付き、すなわち起源が明確である、指紋データを解析した結果を蓄積するものである。指紋解析に基づく個体識別部907は、新サンプル取得対象901から採取した指紋サンプル906について、指紋解析を行い、識別子付き指紋解析データ蓄積部908を参照して、指紋サンプル906がどの個体に識別されたかを示す情報を指紋解析に基づく個体識別結果909として出力する。なお、指紋解析の技法に関しては、当業者にとってよく知られており、また本発明とは直接関係しないので、その詳細な説明は省略する。

- [0120] このように、DNA解析に基づく個体識別結果905と指紋解析に基づく個体識別結果909が得られると、複数情報による個体識別部910は、これらの個体識別結果905、906を組み合わせることにより、複数情報による個体識別結果911を出力する。第8の実施形態の個体識別装置では、DNA解析による結果と指紋解析などによる結果とを組み合わせることで個体識別を行うので、個体識別能力を向上させることができる。
- [0121] 第8の実施形態では、DNA解析の結果と組み合わせられる他の個体識別情報としては、上述した指紋解析情報のほか、虹彩、掌紋あるいは顔などを利用した個体識別技術によって得られる情報を用いることが可能である。また、これらの技術の複数の組み合わせも、可能である。これらの個々の解析技術に関しては、当業者にとってよく知られており、また本発明とは直接関係しないので、その詳細な説明は省略する。

請求の範囲

- [1] DNAサンプルを電気泳動によって解析することによって個体を識別する個体識別方法であって、
- 個体に対する識別子が付与された識別子付きDNAサンプルを解析する第1の解析段階と、
- 前記識別子付きDNAサンプルを解析して得られた結果を対応する識別子とともに識別子付きサンプル解析データ蓄積部に蓄積する段階と、
- 個体識別対象となるDNAサンプルである新サンプルを、前記識別子付きDNAサンプルを解析したときの精度よりも低い精度で解析し、結果を新サンプル解析結果とする第2の解析段階と、
- 前記新サンプル解析結果に基づいて、前記識別子付きサンプル解析データ蓄積部を検索する段階と、
- を有する方法。
- [2] 前記識別子付きDNAサンプル及び前記新サンプルを解析する際に、電気泳動法によって当該サンプルの塩基長に関する情報を求める、請求項1に記載の方法。
- [3] 前記識別子付きDNAサンプル及び前記新サンプルを解析する際に、当該サンプルにおけるマイクロサテライトの繰り返し回数に関する情報を取得する、請求項2に記載の方法。
- [4] 前記第1の解析段階における解析精度は、前記新サンプルでの塩基長の考え得る最小の変化量だけ塩基長が異なる2つのDNAを識別できる精度であり、前記第2の解析段階における解析精度は、前記最小の変化量だけ塩基長が異なる前記2つのDNAを識別できない精度である、請求項2または3に記載の方法。
- [5] 前記第2の解析段階は、
- それぞれ1種のアンプリコンを含んだサンプルの集合から複数のサンプルを任意の組み合わせで選択し、選択されたサンプルを混合して多種アンプリコン・サンプルを生成する段階と、
- 前記多種アンプリコン・サンプルを電気泳動によって解析する第3の解析段階と、
- 前記第3の解析段階で得られた結果と、当該多種アンプリコン・サンプルにおける

塩基長情報とを組にして多種アンプリコン・データ蓄積部に蓄積する段階と、

電気泳動によって前記新サンプルを解析し、新サンプル電気泳動結果データを得る第4の解析段階と、

前記新サンプル電気泳動結果データに基づいて前記多種アンプリコン・データ蓄積部を検索し、結果を前記新サンプル解析結果とする検索段階と、

を有する、請求項1乃至4のいずれか1項に記載の方法。

- [6] アンプリコンの組み合わせのうち、前記多種アンプリコン・データ蓄積部に解析結果が蓄積されていない組み合わせに関して補間により塩基長情報を生成して補間データ蓄積部に蓄積する段階をさらに有し、

前記検索段階において、前記新サンプル電気泳動結果データに基づいて前記多種アンプリコン・データ蓄積部及び前記補間データ蓄積部を検索し、結果を前記新サンプル解析結果とする、請求項5に記載の方法。

- [7] 前記第2の解析段階は、

それぞれ1種のアンプリコンを含んだサンプルの集合から得られる組み合わせのうちの一部の組み合わせに基づいてサンプルを選択して混合し多種アンプリコン・サンプルを生成する段階と、

前記多種アンプリコン・サンプルを電気泳動によって解析する第3の解析段階と、

前記第3の解析段階で得られた結果と、当該多種アンプリコン・サンプルにおける塩基長情報とを組にして多種アンプリコン・データ蓄積部に蓄積する段階と、

電気泳動によって前記新サンプルを解析し、新サンプル電気泳動結果データを得る第4の解析段階と、

前記多種アンプリコン・データ蓄積部を参照しながら、前記新サンプルにおける塩基長の変化に対する前記新サンプル電気泳動結果データをパラメータ化し、そのパラメータ化の結果と前記新サンプル電気泳動結果データとに基づいて前記多種アンプリコン・データ蓄積部を検索し、検索の結果を前記新サンプル解析結果とする段階と、

を有する、請求項1乃至4のいずれか1項に記載の方法。

- [8] 前記新サンプル解析結果に基づいて前記識別子付きサンプル解析データ蓄積部

を検索した結果、前記識別子付きサンプル解析データ蓄積部から該当するエントリが見つかった場合に、前記第1の解析段階で用いたものと同程度の精度で前記新サンプルを解析し、当該解析による結果を用いて前記識別子付きサンプル解析データ蓄積部を検索して個体識別結果を得る段階をさらに有する、請求項1乃至4のいずれか1項に記載の方法。

- [9] DNAサンプルを電気泳動によって解析することによって個体を識別する個体識別方法であって、

個体に対する識別子が付与された識別子付きDNAサンプルを解析して前記識別子付きDNAサンプルの塩基長に関する情報を得る第1の解析段階と、

前記識別子付きDNAサンプルを解析して得られた結果を対応する識別子とともに識別子付きサンプル解析データ蓄積部に蓄積する段階と、

個体識別対象となるDNAサンプルである新サンプルを解析し、前記新サンプルの塩基長に関する情報を含む結果を新サンプル解析結果とする第2の解析段階と、

前記新サンプル解析結果に基づいて、前記識別子付きサンプル解析データ蓄積部を検索する段階と、

を有し、

前記第1の解析段階および前記第2の解析段階における解析の精度は、個体識別対象となるDNAサンプルでの塩基長の考え得る最小の変化量だけ塩基長が異なる2つのDNAを識別できない精度である、方法。

- [10] 前記識別子付きDNAサンプル及び前記新サンプルを解析する際に、当該サンプルにおけるマイクロサテライトの繰り返し回数に関する情報を取得する、請求項9に記載の方法。

- [11] 新サンプル取得対象から得たDNAを前記個体識別対象となるDNAサンプルとして、請求項1乃至10のいずれか1項に記載の方法を実施する第1の個体識別段階と、

前記新サンプル取得対象から得たDNA以外のバイオメトリック情報を用いて個体を識別する第2の個体識別段階と、

前記第1の個体識別段階で得られた結果と前記第2の個体識別段階で得られた結

果とに基づいて個体識別を行う第3の個体識別段階と、
を有する、複数情報による個体識別方法。

[12] 前記バイオメトリック情報が指紋情報である、請求項11に記載の複数情報による個体識別方法。

[13] DNAサンプルを電気泳動によって解析することによって個体を識別する個体識別装置であって、

個体に対する識別子が付与された識別子付きDNAサンプルを解析する第1の解析手段と、

前記第1の解析手段で識別子付きDNAサンプルを解析して得られた結果を対応する識別子とともに蓄積する識別子付きサンプル解析データ蓄積部と、

前記第1の解析手段より低い解析精度を有し、個体識別対象となるDNAサンプルである新サンプルを解析して結果を新サンプル解析結果とする第2の解析手段と、

前記新サンプル解析結果に基づいて、前記識別子付きサンプル解析データ蓄積部を検索し、個体識別結果を得る識別手段と、

を有する装置。

[14] 前記第1及び第2の解析手段は、いずれも、電気泳動法によってサンプルの塩基長に関する情報を求めるものである、請求項13に記載の装置。

[15] 前記第1及び前記第2の解析手段は、それぞれ、前記識別子付きDNAサンプル及び前記新サンプルを解析する際に、当該サンプルにおけるマイクロサテライトの繰り返し回数に関する情報を取得する、請求項14に記載の装置。

[16] 前記第1の解析手段における解析精度は、前記新サンプルでの塩基長の考え得る最小の変化量だけ塩基長が異なる2つのDNAを識別できる精度であり、前記第2の解析手段における解析精度は、前記最小の変化量だけ塩基長が異なる前記2つのDNAを識別できない精度である、請求項14または15に記載の装置。

[17] 前記第2の解析手段は、

それぞれ1種のアンプリコンを含んだサンプルの集合を備える1種アンプリコン・サンプル保管部と、

前記1種アンプリコン・サンプル保管部から複数のサンプルを任意の組み合わせで

選択し、選択したサンプルを混合して得られる多種アンプリコン・サンプルを電気泳動によって解析する第1の電気泳動解析部と、

前記第1の電気泳動解析部で得られた結果と、当該多種アンプリコン・サンプルにおける塩基長情報とを蓄積する多種アンプリコン・データ蓄積部と、

電気泳動によって前記新サンプルを解析し、新サンプル電気泳動結果データを得る第2の電気泳動解析部と、

前記新サンプル電気泳動結果データに基づいて前記多種アンプリコン・データ蓄積部を検索し、結果を前記新サンプル解析結果とするデータ解析部と、

を有する、請求項13乃至16のいずれか1項に記載の装置。

- [18] アンプリコンの組み合わせのうち、前記多種アンプリコン・データ蓄積部に解析結果が蓄積されていない組み合わせに関して補間により塩基長情報を生成する補間データ作成部と、

前記補間データ作成部によって作成された塩基長情報を蓄積する補間データ蓄積部と、

をさらに備え、

前記データ解析部は、前記新サンプル電気泳動結果データに基づいて前記多種アンプリコン・データ蓄積部及び前記補間データ蓄積部を検索し、結果を前記新サンプル解析結果とする、請求項17に記載の装置。

- [19] 前記第2の解析手段は、

それぞれ1種のアンプリコンを含んだサンプルの集合から得られる組み合わせのうちの一部の組み合わせに基づいてサンプルを選択して混合して生成された多種アンプリコン・サンプルを電気泳動によって解析する第1の電気泳動解析部と、

前記第1の電気泳動解析部で得られた結果と、当該多種アンプリコン・サンプルにおける塩基長情報とを組にして蓄積する多種アンプリコン・データ蓄積部と、

電気泳動によって前記新サンプルを解析し、新サンプル電気泳動結果データを得る第2の電気泳動解析部と、

前記多種アンプリコン・データ蓄積部を参照しながら、前記新サンプルにおける塩基長の変化に対する前記新サンプル電気泳動結果データをパラメータ化し、そのパ

ラメータ化の結果と前記新サンプル電気泳動結果データとに基づいて前記多種アンプリコン・データ蓄積部を検索し、検索の結果を前記新サンプル解析結果とするデータ解析部と、

を有する、請求項13乃至16のいずれか1項に記載の装置。

- [20] 前記第1の解析手段と同程度の解析精度を有する第3の解析手段と、
前記第3の解析手段による結果を用いて前記識別子付きサンプル解析データ蓄積部を検索して個体識別結果を得る高精度識別手段と、

をさらに備え、前記新サンプル解析結果に基づいて前記識別手段が前記識別子付きサンプル解析データ蓄積部を検索した結果、前記識別子付きサンプル解析データ蓄積部から該当するエントリが見つかった場合に、前記第3の解析手段が前記新サンプルを解析する、請求項13乃至16のいずれか1項に記載の装置。

- [21] DNAサンプルを電気泳動によって解析することによって個体を識別する個体識別装置であって、

個体に対する識別子が付与された識別子付きDNAサンプルを解析して前記識別子付きDNAサンプルの塩基長に関する情報を得る第1の解析手段と、

前記識別子付きDNAサンプルを解析して得られた結果を対応する識別子とともに蓄積する識別子付きサンプル解析データ蓄積部と、

個体識別対象となるDNAサンプルである新サンプルを解析し、前記新サンプルの塩基長に関する情報を含む結果を新サンプル解析結果とする第2の解析手段と、

前記新サンプル解析結果に基づいて、前記識別子付きサンプル解析データ蓄積部を検索する識別手段と、

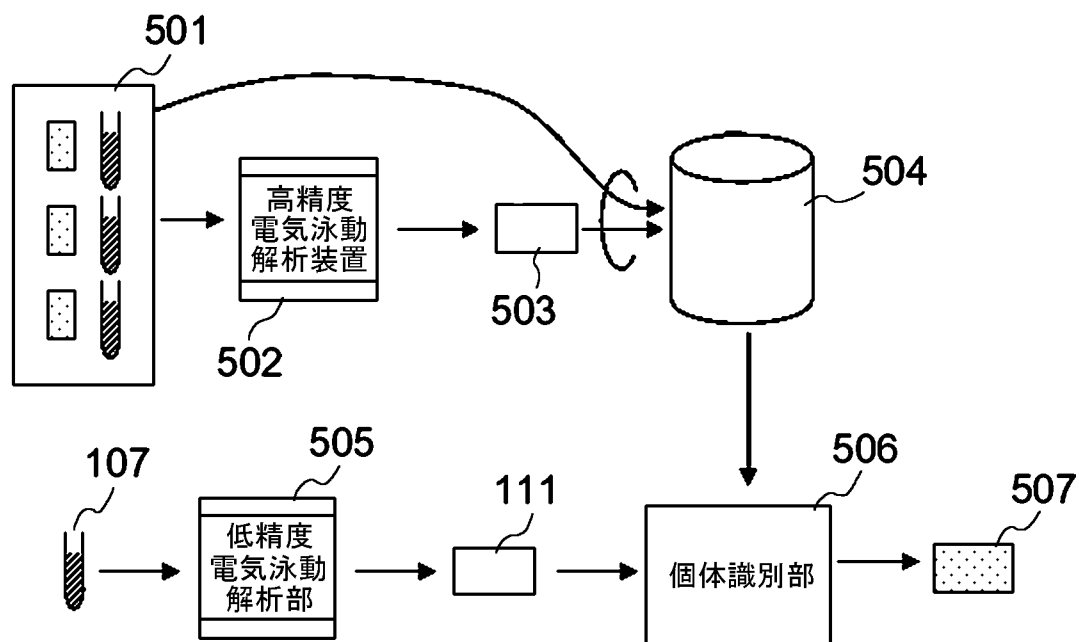
を有し、

前記第1の解析手段および前記第2の解析手段における解析の精度は、個体識別対象となるDNAサンプルでの塩基長の考え得る最小の変化量だけ塩基長が異なる2つのDNAを識別できない精度である、装置。

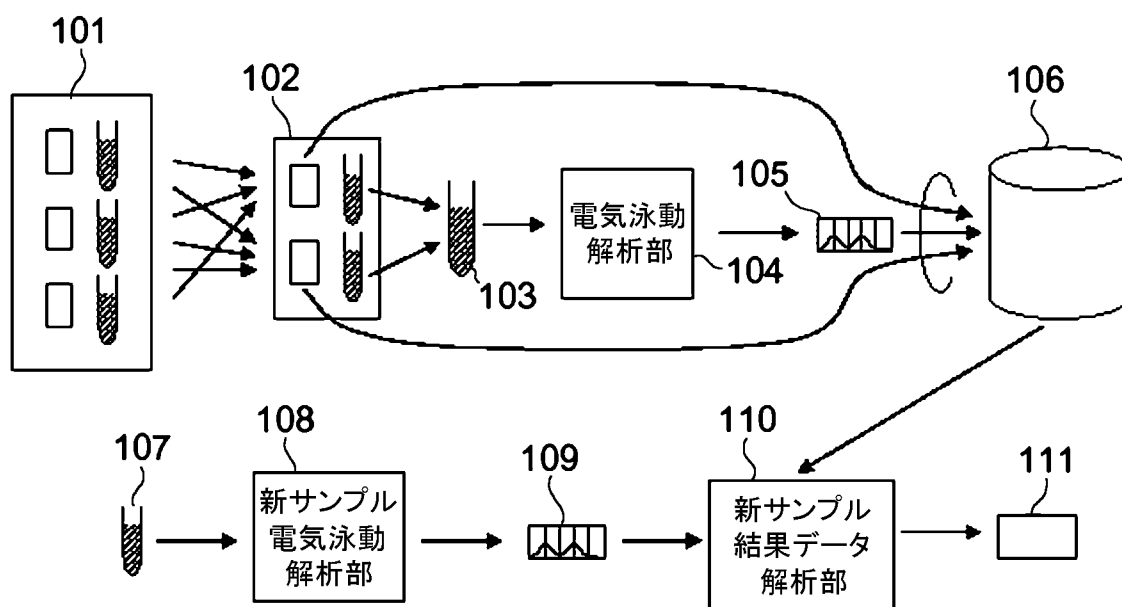
- [22] 前記第1及び第2の解析手段は、それぞれ、前記識別子付きDNAサンプル及び前記新サンプルを解析する際に、当該サンプルにおけるマイクロサテライトの繰り返し回数に関する情報を取得する、請求項21に記載の装置。

- [23] 請求項13乃至22のいずれか1項に記載の装置からなり、新サンプル取得対象から得たDNAを前記個体識別対象となるDNAサンプルとする第1の個体識別手段と、
前記新サンプル取得対象から得たDNA以外のバイOMETリック情報を用いて個体を識別する第2の個体識別手段と、
前記第1の個体識別手段で得られた結果と前記第2の個体識別手段で得られた結果とに基づいて個体識別を行う第3の個体識別手段と、
を有する、複数情報による個体識別装置。
- [24] 前記バイOMETリック情報が指紋情報である、請求項23に記載の複数情報による個体識別装置。

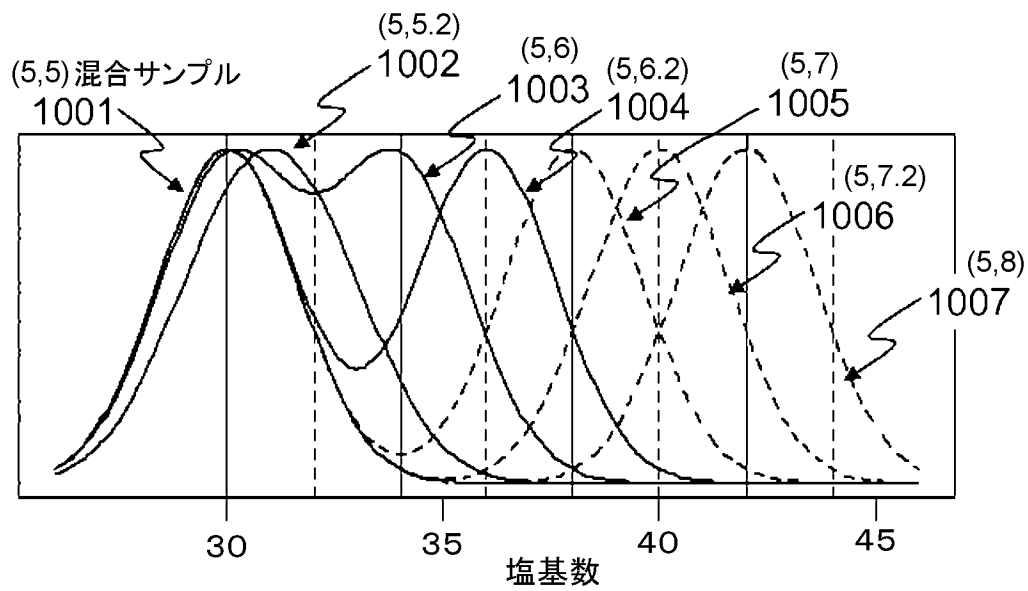
[図1]



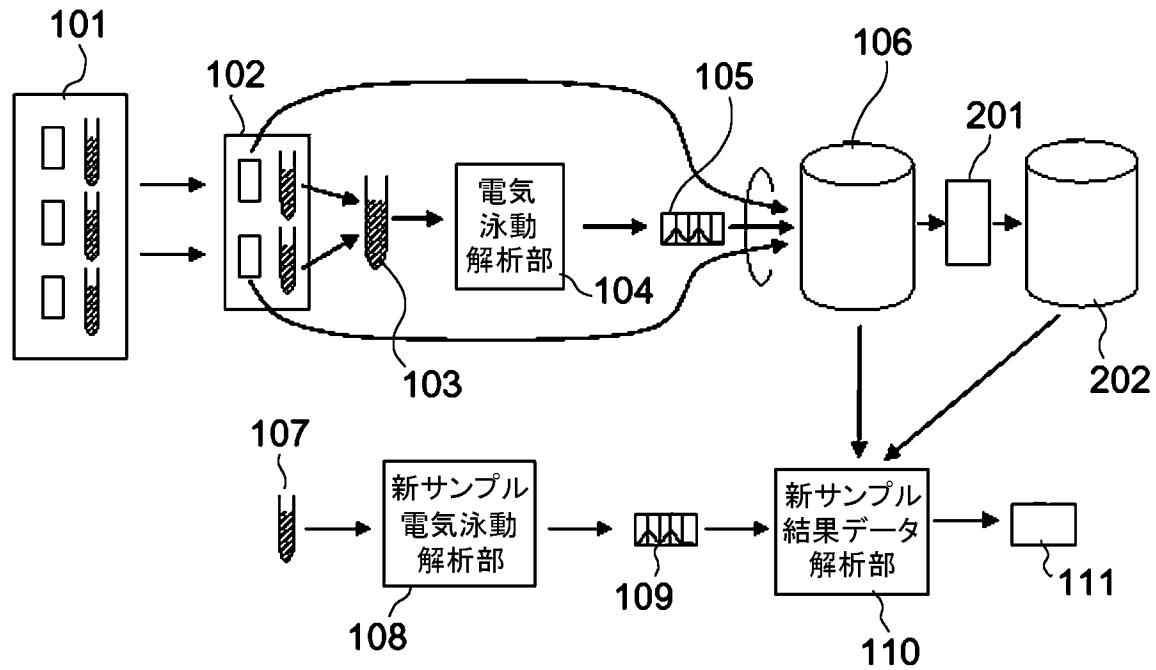
[図2]



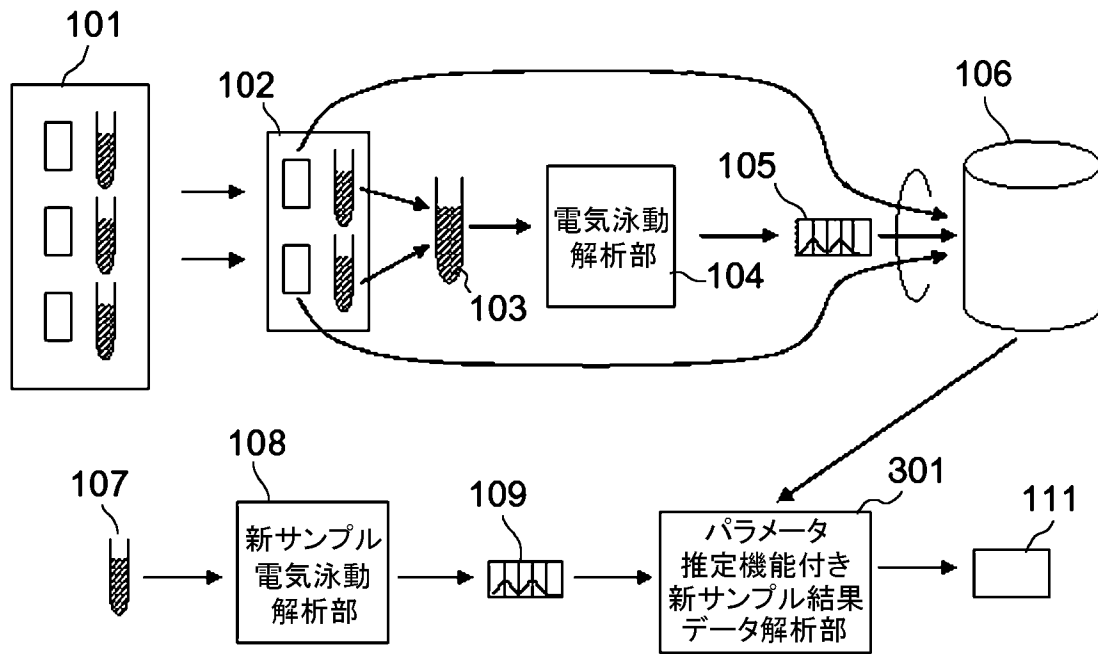
[図3]



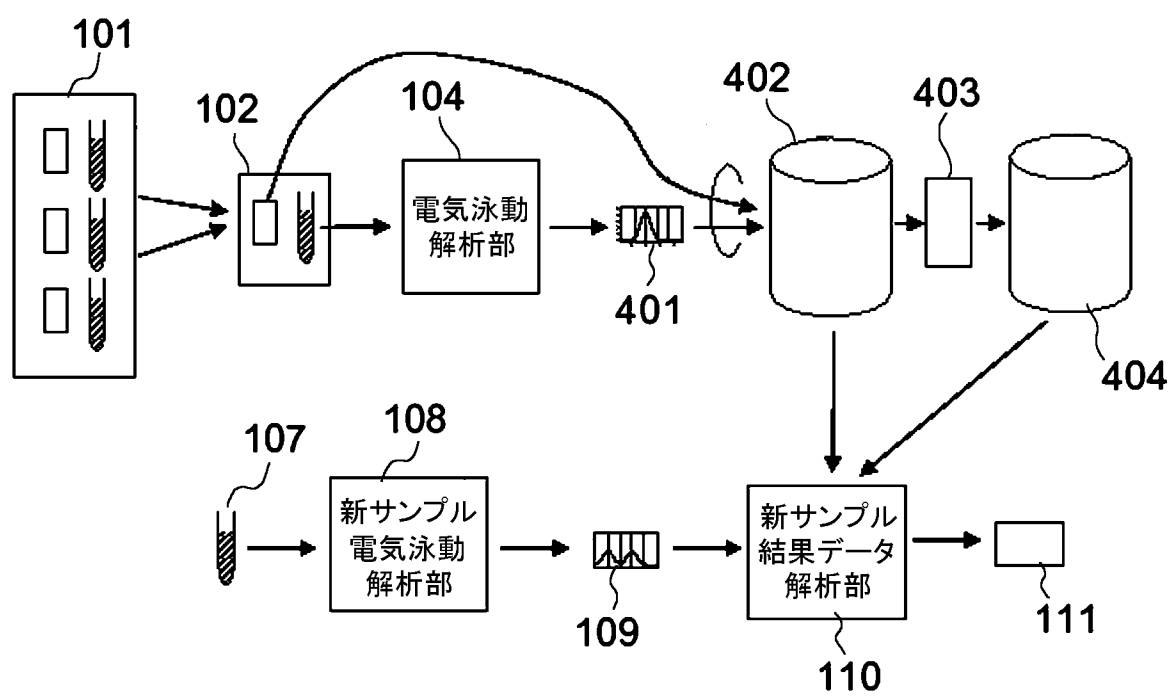
[図4]



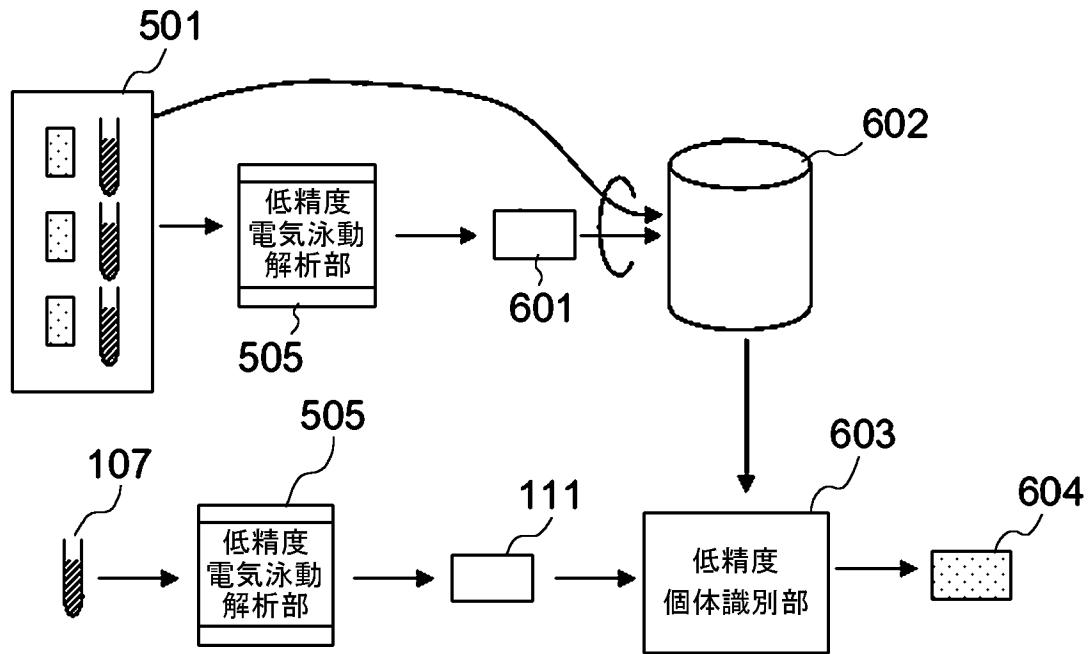
[図5]



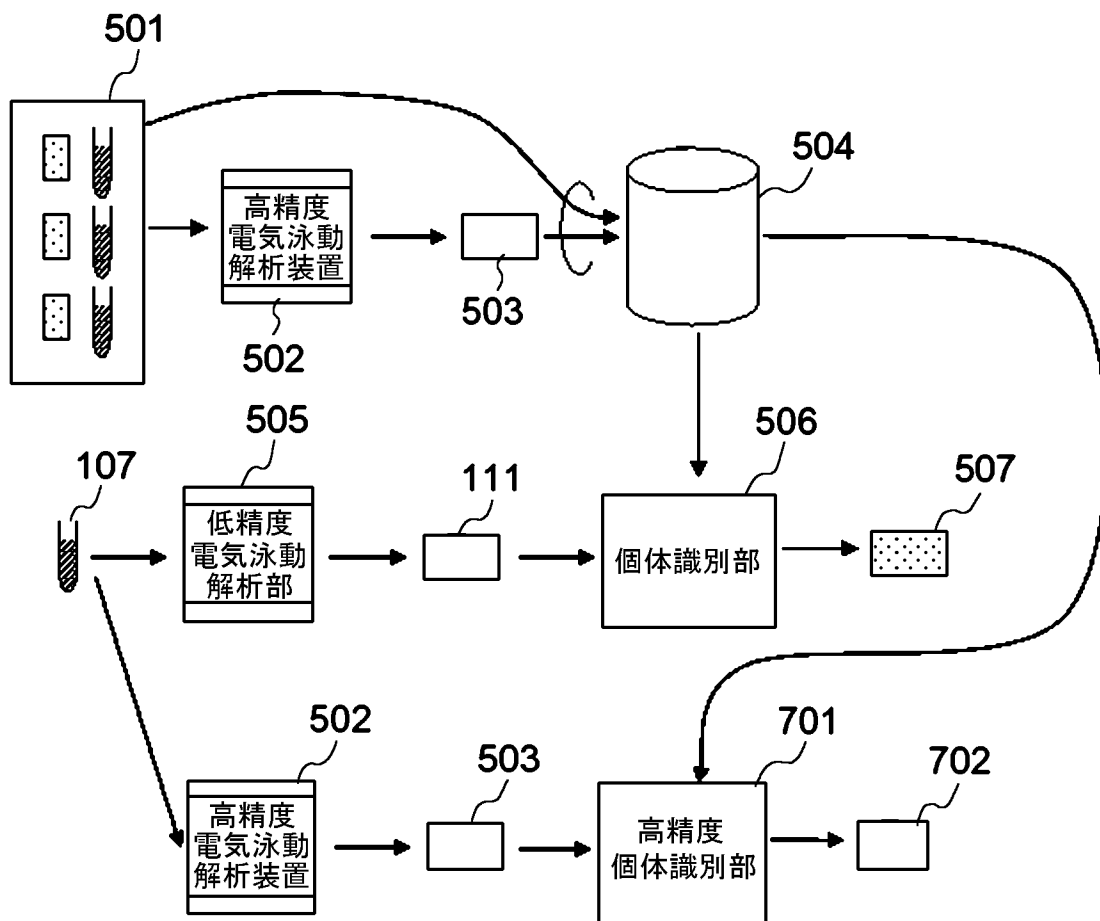
[図6]



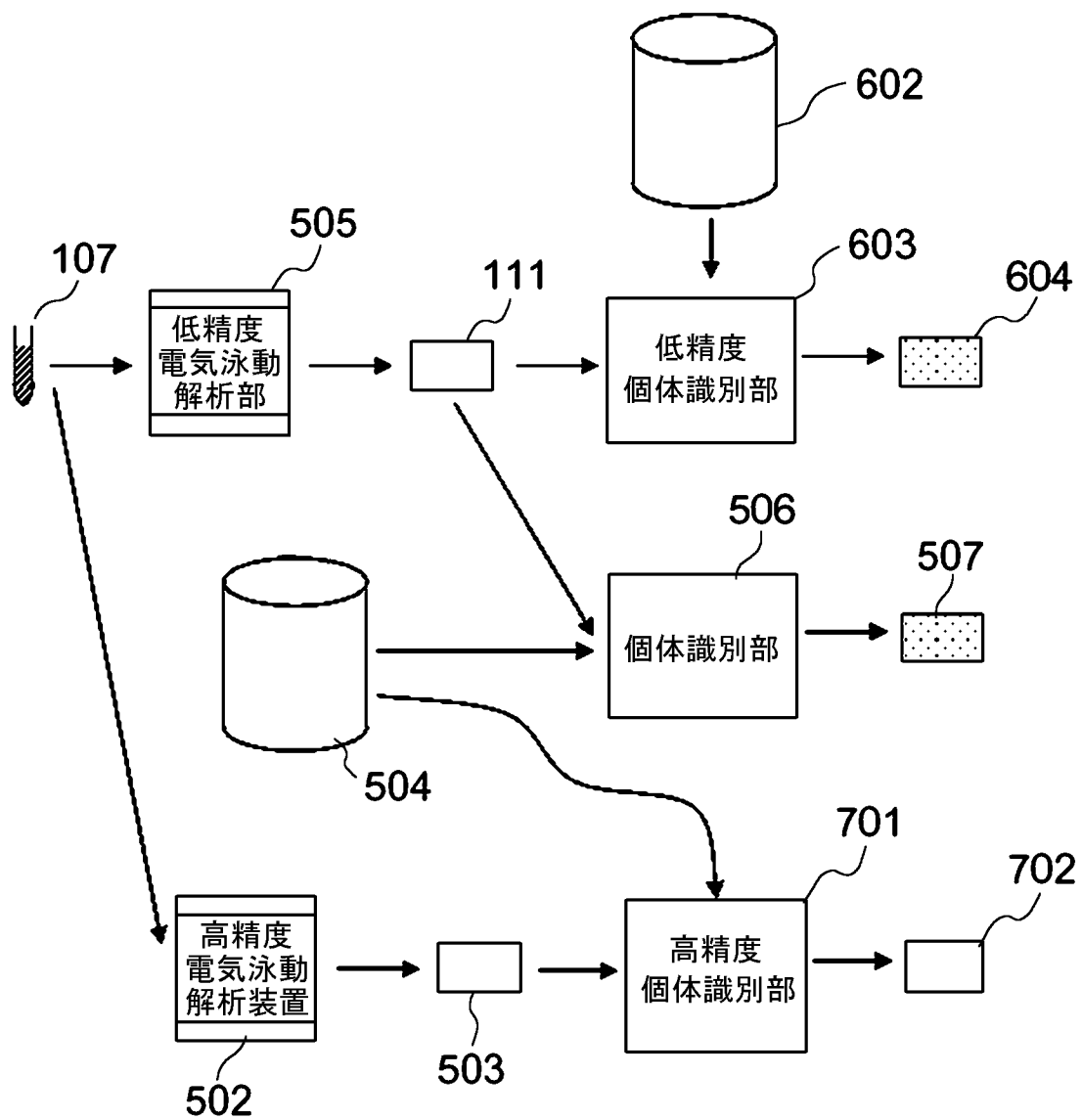
[図7]



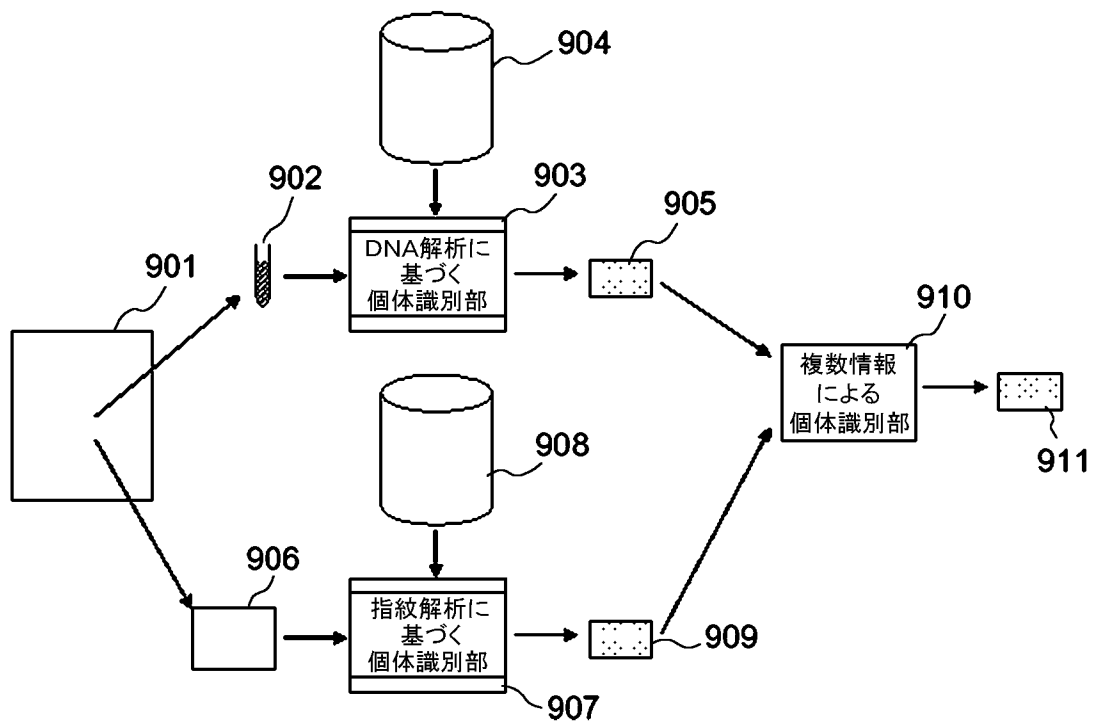
[図8]



[図9]



[図10]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2007/058055

A. CLASSIFICATION OF SUBJECT MATTER

G01N27/447(2006.01)i, C12M1/00(2006.01)i, C12Q1/68(2006.01)i

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)

G01N27/447, C12M1/00, C12Q1/68

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	1996-2007
Kokai Jitsuyo Shinan Koho	1971-2007	Toroku Jitsuyo Shinan Koho	1994-2007

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

JST7580 (JDream2), JSTPlus (JDream2)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	Kentaro KASAI, "Hanzai Sosa ni Okeru DNA-gata Kantei", Gene & medicine, Kabushiki Kaisha	1-4, 8-16, 20-24
A	Medical Do Hakko, 10 November, 1999 (10.11.99) Hakko, Vol.3, No.4, pages 726 to 730	5-7, 17-19
Y	Yukio ITAKURA, "DNA Ninsho Hoshiki to Sono Kadai", Chemical Engineering, Kabushiki Kaisha	1-4, 8-16, 20-24
A	Kagaku Kogyosha Hakko, 01 July, 2005 (01.07.05) Hakko, Vol.50, No.7, pages 500 to 507	5-7, 17-19
Y	JP 2003-150720 A (Dainippon Printing Co., Ltd.), 23 May, 2003 (23.05.03), Claims (Family: none)	11-12, 23-24

☒ 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
04 July, 2007 (04.07.07)

Date of mailing of the international search report
17 July, 2007 (17.07.07)

Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2007/058055

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 2001-338135 A (Hidekazu KIHARA), 07 December, 2001 (07.12.01), Claims (Family: none)	11-12, 23-24
Y	JP 2001-168855 A (Sony Corp.), 22 June, 2001 (22.06.01), Claims (Family: none)	11-12, 23-24
A	JP 2003-245098 A (Hitachi, Ltd.), 02 September, 2003 (02.09.03), Full text & US 2003/0175759 A1	1-24

A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int.Cl. G01N27/447 (2006.01)i, C12M1/00 (2006.01)i, C12Q1/68 (2006.01)i

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. G01N27/447, C12M1/00, C12Q1/68

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	1971-2007年
日本国実用新案登録公報	1996-2007年
日本国登録実用新案公報	1994-2007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)

JST7580(JDream2), JSTPlus(JDream2)

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 請求の範囲の番号
Y A	笠井 賢太郎, 「犯罪捜査におけるDNA型鑑定」, 遺伝子医学, 株式会社 メディカル ドゥ 発行, 1999年11月10日 発行, Vol. 3, No. 4, P. 726-730	1-4,8-16,20-24 5-7,17-19
Y A	板倉 征男, 「DNA認証方式とその課題」, ケミカル・エンジニア リング, 株式会社 化学工業社 発行, 2005年7月1日 発行, 第50巻, 第7号, P. 500-507	1-4,8-16,20-24 5-7,17-19

☒ 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

04.07.2007

国際調査報告の発送日

17.07.2007

国際調査機関の名称及びあて先

日本国特許庁 (ISA/J P)
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)

河野 隆一郎

2 J

3708

電話番号 03-3581-1101 内線 3252

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 請求の範囲の番号
Y	JP 2003-150720 A (大日本印刷株式会社), 2003. 05. 23 , 【特許請求の範囲】 (ファミリーなし)	11-12, 23-24
Y	JP 2001-338135 A (木原 秀和), 2001. 12. 07, 【特許請求の範囲】 (ファミリーなし)	11-12, 23-24
Y	JP 2001-168855 A (ソニー株式会社), 2001. 06. 22 , 【特許請求の範囲】 (ファミリーなし)	11-12, 23-24
A	JP 2003-245098 A (株式会社日立製作所), 2003. 09. 02, 全文 & US 2003/0175759 A1	1-24