(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 10 月 2 日(02.10.2014)

WIPO | PCT

(10) 国際公開番号
WO 2014/156400 A1

(51) 国際特許分類:
G06F 19/22 (2011.01) *G09C 1/00* (2006.01)
G06F 17/30 (2006.01)

(21) 国際出願番号: PCT/JP2014/054145

(22) 国際出願日: 2014 年 2 月 21 日(21.02.2014)

(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 2013-067770 2013 年 3 月 28 日(28.03.2013) JP

(71) 出願人: 三菱スペース・ソフトウエア株式会社 (MITSUBISHI SPACE SOFTWARE CO., LTD.) [JP/JP]; 〒1056132 東京都港区浜松町２丁目４番１号 Tokyo (JP).

(72) 発明者: 谷嶋 成樹 (TANISHIMA, Shigeki); 〒6610001 兵庫県尼崎市塚口本町五丁目４番３６号 三菱スペース・ソフトウエア株式会社 関西事業部内 Hyogo (JP). 松田 規 (MATSUDA, Nori); 〒1008310 東京都千代田区丸の内二丁目７番３号 三菱電機株式会社内 Tokyo (JP).

(74) 代理人: 溝井 章司, 外 (MIZOI, Shoji et al.); 〒2470056 神奈川県鎌倉市大船二丁目１７番１０号 ＮＴＡ大船ビル３階 溝井国際特許事務所 Kanagawa (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, 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).

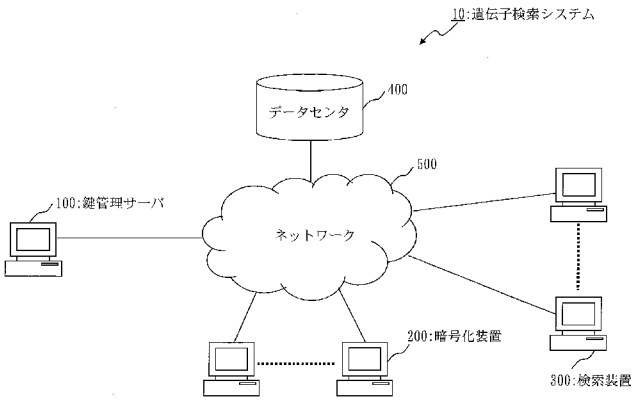
添付公開書類:
— 国際調査報告（条約第 21 条(3)）

WO 2014/156400 A1

(54) **Title**: GENETIC INFORMATION STORAGE DEVICE, GENETIC INFORMATION SEARCH DEVICE, GENETIC INFORMATION STORAGE PROGRAM, GENETIC INFORMATION SEARCH PROGRAM, GENETIC INFORMATION STORAGE METHOD, GENETIC INFORMATION SEARCH METHOD, AND GENETIC INFORMATION SEARCH SYSTEM

(54) 発明の名称 : 遺伝子情報記憶装置、遺伝子情報検索装置、遺伝子情報記憶プログラム、遺伝子情報検索プログラム、遺伝子情報記憶方法、遺伝子情報検索方法及び遺伝子情報検索システム

[図1]



10... GENE SEARCH SYSTEM
100... KEY MANAGEMENT SERVER
200... ENCRYPTION DEVICE
300... SEARCH DEVICE
400... DATA CENTER
500... NETWORK

(57) **Abstract**: An objective of the present invention is to allow searching genetic information in an encrypted state. An encryption device (200) encrypts a gene of interest which is genetic information which is stored in a storage device and generates an encrypted gene, compares a reference gene which is prescribed genetic information with the gene of interest and generates difference information, and generates an encrypted tag which is encrypted with the generated difference information embedded therein. A data center (400) links the encrypted gene and the encrypted tag and stores the same in the storage device. With the difference information as a search keyword, a search device (300) generates a search query which is encrypted with the difference information embedded therein, and transmits the generated search query to the data center (400). The data center (400) identifies the encrypted tag which includes the difference information which is designated in the search query, extracts the linked encrypted gene, and transmits the same to the search device (300).

(57) 要約:

[続葉有]

暗号化した状態で遺伝子情報の検索を可能とすることを目的とする。暗号化装置（２００）は、記憶装置に記憶させる遺伝子情報である対象遺伝子を暗号化して暗号化遺伝子を生成するとともに、所定の遺伝子情報である基準遺伝子と、対象遺伝子とを比較して差分情報を生成し、生成した差分情報を埋め込んで暗号化した暗号化タグを生成する。データセンタ（４００）は、暗号化遺伝子と暗号化タグとを関連付けて記憶装置に記憶する。検索装置（３００）は、差分情報を検索キーワードとして、差分情報を埋め込んで暗号化した検索クエリを生成し、生成した検索クエリをデータセンタ（４００）へ送信する。データセンタ（４００）は、検索クエリに指定された差分情報を含む暗号化タグを特定し、関連付けられた暗号化遺伝子を抽出して検索装置（３００）へ送信する。

明　細　書

発明の名称 ：

遺伝子情報記憶装置、遺伝子情報検索装置、遺伝子情報記憶プログラム、遺伝子情報検索プログラム、遺伝子情報記憶方法、遺伝子情報検索方法及び遺伝子情報検索システム

技術分野

[0001] この発明は、ＤＮＡ配列を解析した結果として得られるゲノムや遺伝子の解析情報である遺伝子情報を暗号化して記憶装置に記憶する記憶技術に関する。また、この発明は、前記記憶技術により記憶された遺伝子情報を、暗号化された状態で検索する技術に関する。

背景技術

[0002] 近年生命工学は著しい発達をしており、ゲノム配列の読み取り能力も発達している。そのため、人間一人分の全ゲノム配列を解読するコストは低下し、多くの人のゲノム配列が解読されるようになってきている。

個人のゲノム配列やその発現情報などの関連情報を含む遺伝子情報は究極の個人情報と言われており、本人以外への情報開示は安全に行われなければならない。ゲノムＤＮＡ及びＲＮＡ配列の読み取り装置から出力され、解析されて意味のある情報になった遺伝子情報は、すみやかに暗号化されて、所有者の同意なしには解読されないようにしなければならない。

[0003] データベースに蓄積されたデータと検索キーワードとを暗号化した状態で、データベースに蓄積されたデータから検索キーワードを含むデータを検索可能とする秘匿検索方式がある（非特許文献１参照）。

この秘匿検索方式では、データベースにデータを蓄積する際、そのデータを検索する場合に用いられると想定される検索キーワードをタグとして抽出しておく。そして、データとタグとをそれぞれ暗号化し、暗号化したデータに暗号化したタグを添付してデータベースに蓄積する。

データベースに蓄積されたデータから検索キーワードを含むデータを検索

する場合、暗号化された検索キーワードが入力される。そして、暗号化された検索キーワードに対応する暗号化されたタグが検索される。検索キーワードに対応するタグが見つかった場合、そのタグが添付されたデータが検索キーワードを含むデータであると特定される。

先行技術文献

非特許文献

[0004] 非特許文献1：Ｂｏｎｅｈ、Ｄｉ Ｃｒｅｓｃｅｎｚｏ、Ｏｓｔｒｏｖｓｋｉ、ａｎｄ Ｐｅｒｓｉａｎｏ“Ｐｕｂｌｉｃ ｋｅｙ ｅｎｃｒｙｐｔｉｏｎ ｗｉｔｈ ｋｅｙｗｏｒｄ ｓｅａｒｃｈ” ＥＵＲＯＣＲＹＰＴ ２００４、ｐｐ５０６－５２２

非特許文献2：Ａｌｌｉｓｏｎ Ｌｅｗｋｏ、Ｔａｔｓｕａｋｉ Ｏｋａｍｏｔｏ、Ａｍｉｔ Ｓａｈａｉ、Ｋａｔｓｕｙｕｋｉ Ｔａｋａｓｈｉｍａ、Ｂｒｅｎｔ Ｗａｔｅｒｓ、“Ｆｕｌｌｙ Ｓｅｃｕｒｅ Ｆｕｎｃｔｉｏｎａｌ Ｅｎｃｒｙｐｔｉｏｎ：Ａｔｔｒｉｂｕｔｅ－Ｂａｓｅｄ Ｅｎｃｒｙｐｔｉｏｎ ａｎｄ（Ｈｉｅｒａｒｃｈｉｃａｌ）Ｉｎｎｅｒ Ｐｒｏｄｕｃｔ Ｅｎｃｒｙｐｔｉｏｎ”、ＥＵＲＯＣＲＹＰＴ２０１０、Ｌｅｃｔｕｒｅ Ｎｏｔｅｓ Ｉｎ Ｃｏｍｐｕｔｅｒ Ｓｃｉｅｎｃｅ、２０１０、Ｖｏｌｕｍｅ ６１１０／２０１０.

発明の概要

発明が解決しようとする課題

[0005] これまでは、暗号化された遺伝子情報を検索する場合、検索が実行される時に、一時的に遺伝子情報が復号化され、検索されていた。

秘匿検索方式を用いた場合であっても、暗号化された遺伝子情報に平文で生成された検索インデックスが付与され、この検索インデックスを用いて検索されていた。また、１つの遺伝子情報に対して膨大な数の検索インデックスが付与されており、検索に膨大な時間がかかっていた。

したがって、遺伝子情報を蓄積しているコンピュータの特権ユーザであれ

ば、復号化されている時間または検索インデックスに自由にアクセスすることができる。そのため、遺伝子情報が完全に第三者に対して隠蔽された状態とは言えない。

この発明は、遺伝子情報を第三者に対して隠蔽した状態で遺伝子情報の検索を可能とすることを目的とする。

課題を解決するための手段

[0006] この発明に係る遺伝子情報記憶装置は、

遺伝子情報を記憶装置に記憶させる遺伝子情報記憶装置であり、

所定の遺伝子情報である基準遺伝子を取得する基準遺伝子取得部と、

前記記憶装置に記憶させる遺伝子情報である対象遺伝子を入力する遺伝子入力部と、

前記基準遺伝子取得部が取得した基準遺伝子と、前記遺伝子入力部が入力した対象遺伝子とを比較して差分情報を生成する差分生成部と、

前記対象遺伝子を暗号化して暗号化遺伝子を生成するデータ暗号化部と、

前記差分生成部が生成した差分情報を埋め込んだ暗号化タグを生成する暗号化タグ生成部と、

前記データ暗号化部が生成した暗号化遺伝子と、前記暗号化タグ生成部が生成した暗号化タグとを関連付けて、前記記憶装置に記憶させるデータ記憶部と

を備えることを特徴とする。

[0007] 前記差分情報には、複数の種別の情報が含まれており、

前記遺伝子情報記憶装置は、さらに、

前記差分情報に含まれる所定の種別については、その種別が取り得る値を複数のブロックに分け、前記差分生成部が生成した差分情報におけるその種別の値を、その値が属するブロックを識別する識別情報に置き換える差分情報置換部

を備え、

前記暗号化タグ生成部は、前記差分情報置換部が置き換えた差分情報を暗

号化して暗号化タグを生成する

ことを特徴とする。

[0008] 前記差分情報置換部は、各ブロックに属する値の一部が他のブロックにも属するように、前記種別が取り得る値を複数のブロックに分け、前記差分生成部が生成した差分情報におけるその種別の値を、その値が属する各ブロックを識別する識別情報に置き換える

ことを特徴とする。

[0009] 前記暗号化タグ生成部は、暗号化データに設定された暗号属性と、秘密鍵に設定された鍵属性とが対応していない場合に、前記暗号化データを前記秘密鍵で復号できない暗号化方式に基づき、前記暗号化遺伝子を検索可能なユーザの属性情報と差分情報とを前記暗号属性として設定して、乱数値を暗号化して前記暗号化タグを生成する

ことを特徴とする。

[0010] 前記暗号化方式に基づき、前記暗号化遺伝子を復号可能なユーザの属性情報を前記暗号属性として設定して、前記対象遺伝子を暗号化して前記暗号化遺伝子を生成する

ことを特徴とする。

[0011] この発明に係る遺伝子情報検索装置は、

データ管理装置が管理する記憶装置に記憶された遺伝子情報を検索する遺伝子情報検索装置であり、

検索したい遺伝子情報と所定の遺伝子情報である基準遺伝子との差分情報を入力する差分情報入力部と、

前記差分情報入力部が入力した差分情報を埋め込んだ検索クエリを生成する検索クエリ生成部と、

前記検索クエリ生成部が生成した検索クエリを前記データ管理装置へ送信して、前記差分情報を含む遺伝子情報を取得する遺伝子情報取得部と

を備えることを特徴とする。

[0012] 前記差分情報には、複数の種別の情報が含まれており、

前記遺伝子情報検索装置は、さらに、

前記差分情報に含まれる所定の種別については、その種別が取り得る値を複数のブロックに分け、差分情報におけるその種別の値を、その値が属するブロックを識別する識別情報に置き換える差分情報置換部

を備え、

前記検索クエリ生成部は、前記差分情報置換部が置き換えた差分情報を埋め込んだ検索クエリを生成する

ことを特徴とする。

[0013] 前記差分情報置換部は、各ブロックに属する値の一部が他のブロックにも属するように、前記種別が取り得る値を複数のブロックに分け、差分情報におけるその種別の値を、その値が属する各ブロックを識別する識別情報に置き換える

ことを特徴とする。

[0014] 前記遺伝子情報検索装置は、さらに、

暗号化データに設定された暗号属性と、秘密鍵に設定された鍵属性とが対応していない場合に、前記暗号化データを前記秘密鍵で復号できない暗号化方式における前記秘密鍵であって、ユーザの属性情報が前記鍵属性として設定された秘密鍵を管理するユーザ秘密鍵管理部

を備え、

前記検索クエリ生成部は、前記ユーザ秘密鍵管理部が管理する秘密鍵の鍵属性として、前記差分情報を追加して前記検索クエリを生成する

ことを特徴とする。

[0015] 前記遺伝子情報取得部は、前記暗号化方式により、復号可能なユーザの属性情報を前記暗号属性として設定して、前記遺伝子情報を暗号化した暗号化遺伝子を、前記差分情報を含む遺伝子情報として取得し、

前記遺伝子情報検索装置は、さらに、

前記ユーザ秘密鍵管理部が管理する秘密鍵により、前記暗号化遺伝子を復号する復号部を備えることを特徴とする。

[0016] この発明に係る遺伝子情報記憶プログラムは、

遺伝子情報を記憶装置に記憶させる遺伝子情報記憶プログラムであり、

所定の遺伝子情報である基準遺伝子を取得する基準遺伝子取得処理と、

前記記憶装置に記憶させる遺伝子情報である対象遺伝子を入力する遺伝子入力処理と、

前記基準遺伝子取得処理で取得した基準遺伝子と、前記遺伝子入力処理で入力した対象遺伝子とを比較して差分情報を生成する差分生成処理と、

前記対象遺伝子を暗号化して暗号化遺伝子を生成するデータ暗号化処理と、

前記差分生成処理で生成した差分情報を埋め込んだ暗号化タグを生成する暗号化タグ生成処理と、

前記データ暗号化処理で生成した暗号化遺伝子と、前記暗号化タグ生成処理で生成した暗号化タグとを関連付けて、前記記憶装置に記憶させるデータ記憶処理と

をコンピュータに実行させることを特徴とする。

[0017] 前記差分情報には、複数の種別の情報が含まれており、

前記遺伝子情報記憶プログラムは、さらに、

前記差分情報に含まれる所定の種別については、その種別が取り得る値を複数のブロックに分け、前記差分生成処理で生成した差分情報におけるその種別の値を、その値が属するブロックを識別する識別情報に置き換える差分情報置換処理

をコンピュータに実行させ、

前記暗号化タグ生成処理では、前記差分情報置換処理で置き換えた差分情報を暗号化して暗号化タグを生成する

ことを特徴とする。

[0018] 前記差分情報置換処理では、各ブロックに属する値の一部が他のブロックにも属するように、前記種別が取り得る値を複数のブロックに分け、前記差分生成処理で生成した差分情報におけるその種別の値を、その値が属する各

ブロックを識別する識別情報に置き換える
ことを特徴とする。

[0019] 前記暗号化タグ生成処理では、暗号化データに設定された暗号属性と、秘密鍵に設定された鍵属性とが対応していない場合に、前記暗号化データを前記秘密鍵で復号できない暗号化方式に基づき、前記暗号化遺伝子を検索可能なユーザの属性情報と差分情報とを前記暗号属性として設定して、乱数値を暗号化して前記暗号化タグを生成する
ことを特徴とする。

[0020] 前記暗号化方式に基づき、前記暗号化遺伝子を復号可能なユーザの属性情報を前記暗号属性として設定して、前記対象遺伝子を暗号化して前記暗号化遺伝子を生成する
ことを特徴とする。

[0021] この発明に係る遺伝子情報検索プログラムは、
データ管理装置が管理する記憶装置に記憶された遺伝子情報を検索する遺伝子情報検索プログラムであり、
検索したい遺伝子情報と所定の遺伝子情報である基準遺伝子との差分情報を入力する差分情報入力処理と、
前記差分情報入力処理で入力した差分情報を埋め込んだ検索クエリを生成する検索クエリ生成処理と、
前記検索クエリ生成処理で生成した検索クエリを前記データ管理装置へ送信して、前記差分情報を含む遺伝子情報を取得する遺伝子情報取得処理と
をコンピュータに実行させることを特徴とする。

[0022] 前記差分情報には、複数の種別の情報が含まれており、
前記遺伝子情報検索プログラムは、さらに、
前記差分情報に含まれる所定の種別については、その種別が取り得る値を複数のブロックに分け、差分情報におけるその種別の値を、その値が属するブロックを識別する識別情報に置き換える差分情報置換処理
をコンピュータに実行させ、

前記検索クエリ生成処理では、前記差分情報置換処理で置き換えた差分情報を埋め込んだ検索クエリを生成する

ことを特徴とする。

[0023] 前記差分情報置換処理では、各ブロックに属する値の一部が他のブロックにも属するように、前記種別が取り得る値を複数のブロックに分け、差分情報におけるその種別の値を、その値が属する各ブロックを識別する識別情報に置き換える

ことを特徴とする。

[0024] 前記遺伝子情報検索プログラムは、さらに、

暗号化データに設定された暗号属性と、秘密鍵に設定された鍵属性とが対応していない場合に、前記暗号化データを前記秘密鍵で復号できない暗号化方式における前記秘密鍵であって、ユーザの属性情報が前記鍵属性として設定された秘密鍵を管理するユーザ秘密鍵管理処理

をコンピュータに実行させ、

前記検索クエリ生成処理では、前記ユーザ秘密鍵管理処理で管理する秘密鍵の鍵属性として、前記差分情報を追加して前記検索クエリを生成する

ことを特徴とする。

[0025] 前記遺伝子情報取得処理では、前記暗号化方式により、復号可能なユーザの属性情報を前記暗号属性として設定して、前記遺伝子情報を暗号化した暗号化遺伝子を、前記差分情報を含む遺伝子情報として取得し、

前記遺伝子情報検索プログラムは、さらに、

前記ユーザ秘密鍵管理処理で管理する秘密鍵により、前記暗号化遺伝子を復号する復号処理

をコンピュータに実行させることを特徴とする。

[0026] この発明に係る遺伝子情報記憶方法は、

遺伝子情報を記憶装置に記憶させる遺伝子情報記憶方法であり、

処理装置が、所定の遺伝子情報である基準遺伝子を取得する基準遺伝子取得工程と、

入力装置が、前記記憶装置に記憶させる遺伝子情報である対象遺伝子を入力する遺伝子入力工程と、

処理装置が、前記基準遺伝子取得工程で取得した基準遺伝子と、前記遺伝子入力工程で入力した対象遺伝子とを比較して差分情報を生成する差分生成工程と、

処理装置が、前記対象遺伝子を暗号化して暗号化遺伝子を生成するデータ暗号化工程と、

処理装置が、前記差分生成工程で生成した差分情報を埋め込んだ暗号化タグを生成する暗号化タグ生成工程と、

処理装置が、前記データ暗号化工程で生成した暗号化遺伝子と、前記暗号化タグ生成工程で生成した暗号化タグとを関連付けて、前記記憶装置に記憶させるデータ記憶工程と

を備えることを特徴とする。

[0027] この発明に係る遺伝子情報検索方法は、

データ管理装置が管理する記憶装置に記憶された遺伝子情報を検索する遺伝子情報検索方法であり、

入力装置が、検索したい遺伝子情報と所定の遺伝子情報である基準遺伝子との差分情報を入力する差分情報入力工程と、

処理装置が、前記差分情報入力工程で入力した差分情報を埋め込んだ検索クエリを生成する検索クエリ生成工程と、

処理装置が、前記検索クエリ生成工程で生成した検索クエリを前記データ管理装置へ送信して、前記差分情報を含む遺伝子情報を取得する遺伝子情報取得工程と

を備えることを特徴とする。

[0028] この発明に係る遺伝子情報検索システムは、

データ管理装置が管理する記憶装置に遺伝子情報を記憶させる遺伝子情報記憶装置と、前記遺伝子情報記憶装置が記憶させた遺伝子情報から検索キーワードを含む遺伝子情報を検索する遺伝子情報検索装置とを備える遺伝子情

報検索システムであり、

前記遺伝子情報記憶装置は、

所定の遺伝子情報である基準遺伝子を取得する基準遺伝子取得部と、

前記記憶装置に記憶させる遺伝子情報である対象遺伝子を入力する遺伝子入力部と、

前記基準遺伝子取得部が取得した基準遺伝子と、前記遺伝子入力部が入力した対象遺伝子とを比較して差分情報を生成する差分生成部と、

前記対象遺伝子を暗号化して暗号化遺伝子を生成するデータ暗号化部と、

前記差分生成部が生成した差分情報を埋め込んだ暗号化タグを生成する暗号化タグ生成部と、

前記データ暗号化部が生成した暗号化遺伝子と、前記暗号化タグ生成部が生成した暗号化タグとを関連付けて、前記記憶装置に記憶させるデータ記憶部と

を備え、

前記遺伝子情報検索装置は、

検索したい遺伝子情報と前記基準遺伝子との差分情報を前記検索キーワードとして入力する差分情報入力部と、

前記差分情報入力部が入力した差分情報を埋め込んだ検索クエリを生成する検索クエリ生成部と、

前記検索クエリ生成部が生成した検索クエリを前記データ管理装置へ送信して、前記差分情報を含む遺伝子情報を取得する遺伝子情報取得部と

備えることを特徴とする。

発明の効果

[0029] この発明によれば、データベースに蓄積された遺伝子情報及び検索インデックスであるタグと、検索キーワードとして用いられる遺伝子情報とのどちらも暗号化した状態で、検索キーワードを含む遺伝子情報をデータベースから抽出することができる。したがって、遺伝子情報が完全に第三者に対して隠蔽された状態となる。

特に、この発明によれば、基準遺伝子との差分情報を検索インデックス及び検索キーワードとする。そのため、検索インデックスの数が少なくて済み、検索を高速に行うことができる。

図面の簡単な説明

[0030] [図1]遺伝子検索システム１０の構成図。

[図2]鍵管理サーバ１００の構成図。

[図3]暗号化装置２００の構成図。

[図4]検索装置３００の構成図。

[図5]データセンタ４００の構成図。

[図6]階層的内積述語暗号を用いた暗号方式の説明図。

[図7]階層的内積述語暗号を用いた秘匿検索方式の説明図。

[図8]初期設定処理の流れを示すフローチャート。

[図9]エンコードしたＳＮＶ情報の構成図。

[図10]ＳＮＶ情報における位置情報のブロック化の説明図。

[図11]ＳＮＶ情報における信頼度のブロック化の説明図。

[図12]エンコードしたＳＶ情報の構成図。

[図13]ＳＶ情報におけるＣＮＶゲインのブロック化の説明図。

[図14]エンコードしたＮＣ情報の構成図。

[図15]タグＩＤの階層構造の説明図。

[図16]復号者ＩＤの階層構造の説明図。

[図17]ユーザＩＤ情報データベースの説明図。

[図18]ユーザ秘密鍵発行処理の流れを示すフローチャート。

[図19]患者ゲノム配列の暗号化処理の流れを示すフローチャート。

[図20]タグ付き暗号化データの説明図。

[図21]暗号化データの記憶例を示す図。

[図22]患者の電子カルテの暗号化処理の流れを示すフローチャート。

[図23]検索処理の流れを示すフローチャート。

[図24]アクセス権限管理テーブルの説明図。

[図25]鍵管理サーバ１００、暗号化装置２００、検索装置３００、データセンタ４００のハードウェア構成の一例を示す図。

発明を実施するための形態

[0031] 実施の形態１．

図１は、遺伝子検索システム１０の構成図である。

遺伝子検索システム１０は、鍵管理サーバ１００、複数の暗号化装置２００、複数の検索装置３００、データセンタ４００（データ管理装置）を備える。鍵管理サーバ１００、暗号化装置２００、検索装置３００、データセンタ４００は、ネットワーク５００を介して接続される。

鍵管理サーバ１００は、暗号化用ユーザ秘密鍵や秘匿検索用ユーザ秘密鍵などのユーザ秘密鍵を生成し、暗号化装置２００や検索装置３００へ配布するサーバである。なお、暗号化用ユーザ秘密鍵は暗号化データの復号に用いる鍵であり、秘匿検索用ユーザ秘密鍵は秘匿検索に用いる鍵である。

暗号化装置２００は、データセンタ４００に保管する情報を暗号化するための端末である。暗号化装置２００は、主に、病院の医師や、ゲノム解読センタの社員や患者などのユーザによって利用される端末である。

検索装置３００は、データセンタ４００に保管された情報を検索して取得するために用いられる端末である。検索装置３００は、主に、製薬会社などの研究者や、病院の医師などのユーザによって利用される。

データセンタ４００は、患者から集めたゲノム情報や、患者の病歴が記載された電子カルテなどを保管するサーバである。データセンタ４００は、患者や医師や研究者などのユーザからの要求によってゲノム情報や電子カルテなどを検索・閲覧するサービスを提供する。

ネットワーク５００は、例えばインターネットのような公衆回線網である。

[0032] 図２は、鍵管理サーバ１００の構成図である。

鍵管理サーバ１００は、マスター鍵生成部１１０、鍵記憶部１２０、ユーザ秘密鍵生成部１３０、データ送受信部１４０、ユーザＩＤ記憶部１５０を

備える。

マスター鍵生成部１１０は、処理装置により、秘匿検索を利用するユーザ全員が共通で利用する公開パラメータを生成するとともに、ユーザ秘密鍵を生成する元になるマスター鍵を生成する。

鍵記憶部１２０は、マスター鍵生成部１１０が生成したマスター鍵や公開パラメータを記憶装置に記憶する。

ユーザ秘密鍵生成部１３０は、処理装置により、ユーザにユニークに割り当てられたユーザＩＤを用いて、マスター鍵からユーザ秘密鍵を生成する。

データ送受信部１４０は、公開パラメータを暗号化装置２００、検索装置３００、データセンタ４００へネットワーク５００を介して送信する。また、データ送受信部１４０は、ユーザ秘密鍵を検索装置３００へネットワーク５００を介して送信する。また、データ送受信部１４０は、ユーザの要求に応じて、ユーザＩＤを暗号化装置２００、検索装置３００、データセンタ４００のユーザに対して送信する。

ユーザＩＤ記憶部１５０は、各ユーザのユーザＩＤを記憶装置に記憶する。ユーザＩＤは、ユーザの氏名、所属、ログインＩＤ、メールアドレスなどの属性情報である。ユーザＩＤ記憶部１５０は、現時点での属性情報だけでなく、過去の属性情報も履歴として記憶しておいてもよい。

[0033] 図３は、暗号化装置２００の構成図である。

暗号化装置２００は、基準遺伝子取得部２１０、対象遺伝子入力部２２０、公開パラメータ記憶部２３０、差分情報生成部２４０、差分情報エンコード部２５０、データ暗号化部２６０、暗号化タグ生成部２７０、タグ付き暗号化データ生成部２８０を備える。

基準遺伝子取得部２１０は、一般に公開されている所定のゲノム配列を基準ゲノム配列（基準遺伝子）として取得する。

対象遺伝子入力部２２０は、データセンタ４００に保管する患者ゲノム配列（対象遺伝子）を取得する。また、対象遺伝子入力部２２０は、患者ゲノム配列とともに、その患者ゲノム配列の患者を示す患者ＩＤを取得する。

公開パラメータ記憶部２３０は、鍵管理サーバ１００が生成した公開パラメータを受信し、記憶装置に記憶する。

差分情報生成部２４０は、処理装置により、患者ゲノム配列と基準ゲノム配列とを比較して、複数の差分情報を生成する。

差分情報エンコード部２５０は、処理装置により、差分情報生成部２４０が生成した各差分情報を、暗号化したまま検索を行うのに適した形に符号化してエンコード差分情報を生成する。暗号化したまま検索を行うのに適した形については、後述する。

データ暗号化部２６０は、処理装置により、対象遺伝子入力部２２０が入力した患者ゲノム配列を暗号化して、暗号化データ（暗号化遺伝子）を生成する。

暗号化タグ生成部２７０は、処理装置により、差分情報エンコード部２５０が生成したエンコード差分情報を暗号化して、暗号化タグを生成する。

タグ付き暗号化データ生成部２８０は、処理装置により、データ暗号化部２６０が生成した暗号化データと、暗号化タグ生成部２７０が生成した複数の暗号化タグと、患者ＩＤとを結合してタグ付き暗号化データを生成する。タグ付き暗号化データ生成部２８０は、生成したタグ付き暗号化データの保管をデータセンタ４００に依頼する。

[0034] 図４は、検索装置３００の構成図である。

検索装置３００は、差分情報入力部３１０、ユーザ秘密鍵記憶部３２０、差分情報エンコード部３３０、検索クエリ生成部３４０、遺伝子情報取得部３５０、データ復号部３６０を備える。

差分情報入力部３１０は、入力装置により、基準ゲノム配列との差分情報を検索キーワードとして含む検索要求を入力する。

ユーザ秘密鍵記憶部３２０は、鍵管理サーバ１００がユーザに個別に発行したユーザ秘密鍵と、公開パラメータとを記憶装置に記憶する。

差分情報エンコード部３３０は、差分情報エンコード部２５０と同じ機能である。差分情報エンコード部３３０は、処理装置により、差分情報入力部

３１０が入力した検索要求に含まれる差分情報を暗号化したまま検索を行うのに適した形に符号化してエンコード差分情報を生成する。

検索クエリ生成部３４０は、処理装置により、ユーザ秘密鍵記憶部３２０が記憶したユーザ秘密鍵及び公開パラメータと、差分情報エンコード部３３０が生成したエンコード差分情報から、検索クエリを生成する。

遺伝子情報取得部３５０は、検索クエリ生成部３４０が生成した検索クエリをデータセンタ４００へネットワーク５００を介して送信する。そして、遺伝子情報取得部３５０は、検索要求に含まれる差分情報を含む（あるいは、類似する差分情報を含む）患者ゲノム配列が暗号化された暗号化データをデータセンタ４００からネットワーク５００を介して受信する。また、遺伝子情報取得部３５０は、暗号化データとともに、患者ＩＤを受信する。

データ復号部３６０は、処理装置により、データセンタ４００から受信した暗号化データを、ユーザ秘密鍵記憶部３２０が記憶したユーザ秘密鍵を用いて復号し、患者ゲノム配列を取得する。

[0035] 図５は、データセンタ４００の構成図である。

データセンタ４００は、保管要求処理部４１０、暗号化データ記憶部４２０、暗号化タグ記憶部４３０、検索要求処理部４４０、公開パラメータ記憶部４５０、アクセス権限記憶部４６０を備える。

保管要求処理部４１０は、暗号化装置２００からタグ付き暗号化データを受信する。保管要求処理部４１０は、受信したタグ付き暗号化データを解析し、暗号化データと複数の暗号化タグと患者ＩＤとに分解する。保管要求処理部４１０は、分解した暗号化データと各暗号化タグとに共通の管理番号を割り振り、暗号化データを患者ＩＤ及び管理番号とともに暗号化データ記憶部４２０へ送信し、各暗号化タグを患者ＩＤ及び管理番号とともに暗号化タグ記憶部４３０へ送信する。

暗号化データ記憶部４２０は、保管要求処理部４１０から受信した暗号化データを、患者ＩＤ及び管理番号と関連付けて記憶装置に記憶する。

暗号化タグ記憶部４３０は、保管要求処理部４１０から受信した暗号化タ

グを、患者ＩＤ及び管理番号と関連付けて記憶装置に記憶する。

検索要求処理部４４０は、検索装置３００から検索クエリを受信する。検索要求処理部４４０は、処理装置により、受信した検索クエリと、暗号化タグ記憶部４３０が記憶した暗号化タグとについて比較処理を行う。この比較処理によって、暗号化タグに含まれる患者ゲノム配列の差分情報が、検索クエリに含まれる患者ゲノム配列の差分情報（検索要求）で指定された条件に合致するかどうかが判定される。その後、検索要求処理部４４０は、検索にヒットした暗号化タグと関連付けられた暗号化データを暗号化データ記憶部４２０から取得し、検索装置３００へ返送する。なお、暗号化タグと関連付けられた暗号化データとは、暗号化タグと同一の管理番号が付された暗号化データである。

公開パラメータ記憶部４５０は、鍵管理サーバ１００が生成した公開パラメータを受信し、記憶装置に記憶する。

アクセス権限記憶部４６０は、患者が誰に対して患者ゲノム配列の開示を許可しているかを管理する。

[0036] 遺伝子検索システム１０で利用する暗号方式について説明する。

遺伝子検索システム１０では、非特許文献２等に記載された階層的内積述語暗号と呼ばれる暗号方式と、同じく階層的内積述語暗号を用いた暗号化したままキーワード検索が可能な秘匿検索方式とを用いる。

[0037] 図６は、階層的内積述語暗号を用いた暗号方式の説明図である。

この暗号方式は、マスター鍵生成アルゴリズム、秘密鍵生成アルゴリズム、委譲鍵生成アルゴリズム、暗号化アルゴリズム、復号アルゴリズムによって構成される。

[0038] 最初に、マスター鍵生成アルゴリズムを用いて、暗号化用マスター鍵と暗号化用公開パラメータとが生成される。暗号化用マスター鍵は、復号ユーザの暗号化用ユーザ秘密鍵を生成するために使われる秘密鍵である。暗号化用公開パラメータは、暗号化する際に用いられる公開情報であり、暗号化するユーザに対して広く配布される。

なお、この処理の際には、条件式の構成を事前に決定しておき、その条件式をパラメータとして与える必要がある。非特許文献２等の既存の階層的内積述語暗号の文献では、条件式の構成を与えるのではなく、条件式をベクトルとして表記した場合の次元数を与えるように記載しているが、分かりやすさの観点から、ここでは条件式を与える、としておく。以下同様である。

[0039] 次に、秘密鍵生成アルゴリズムを用いて、暗号化用条件式から復号ユーザに対して配布する暗号化用ユーザ秘密鍵が発行される。暗号化用条件式は、復号ユーザがどの様な属性を持つ文書を復号できるかが決定され、その条件がＡＮＤ／ＯＲの論理演算を用いた条件式として記述されたものである。

[0040] 次に、暗号化アルゴリズムを用いて、データが暗号化される。その際、暗号化データに付加する暗号化用属性を指定して、これを暗号化データに埋め込む。

[0041] 最後に、復号アルゴリズムを用いて、暗号化データが復号される。復号時には、暗号化用ユーザ秘密鍵を指定するが、暗号化用ユーザ秘密鍵に埋め込まれた条件式を満たすような暗号化用属性が付与された暗号化データのみ復号ができる。条件式を満たさないような暗号化データに関しては一切復号できない。

[0042] なお、階層的内積述語暗号には、鍵委譲という特徴も存在する。これは、暗号化用ユーザ秘密鍵を生成する際に条件式を一部だけ設定し、条件式の一部を未指定の状態にする。この未指定の条件式に、追加で条件式を設定する仕組みが鍵委譲である。

具体的には、委譲鍵生成アルゴリズムを用いて、暗号化用ユーザ秘密鍵に対して追加で設定する暗号化用追加条件式を指定し、暗号化用委譲秘密鍵を生成する。この暗号化用委譲秘密鍵は、暗号化用ユーザ秘密鍵と同様に暗号化データの復号に利用することができる。

[0043] 少し言い方を変えると、秘密鍵生成アルゴリズムでは、復号ユーザの属性情報が鍵属性として設定された暗号化用ユーザ秘密鍵が発行される。また、暗号化アルゴリズムでは、復号可能なユーザの属性情報が暗号属性として設

定された暗号化データが生成される。そして、復号アルゴリズムでは、暗号化用ユーザ秘密鍵に設定された鍵属性と、暗号化データに設定された暗号属性とが対応している場合にのみ、暗号化データを暗号化用ユーザ秘密鍵で復号できる。

また、委譲鍵生成アルゴリズムでは、暗号化用ユーザ秘密鍵にさらに属性情報を追加設定して、その暗号化用ユーザ秘密鍵で復号可能な暗号化データの一部のみを復号可能な下位の暗号化用ユーザ秘密鍵を生成する。

[0044] 図７は、階層的内積述語暗号を用いた秘匿検索方式の説明図である。

この秘匿検索方式は、マスター鍵生成アルゴリズム、秘密鍵生成アルゴリズム、委譲鍵生成アルゴリズム、暗号化タグ生成アルゴリズム、一致判定アルゴリズムで構成される。

[0045] 最初に、マスター鍵生成アルゴリズムを用いて、秘匿検索用マスター鍵と秘匿検索用公開パラメータとが生成される。秘匿検索用マスター鍵は、検索ユーザの秘匿検索用ユーザ秘密鍵を生成するために使われる秘密鍵である。秘匿検索用公開パラメータは、検索する際に用いられる公開情報であり、検索するユーザに対して広く配布される。

なお、条件式の記述に関しては、暗号化方式と同様である。

[0046] 次に、秘密鍵生成アルゴリズムを用いて、秘匿検索用条件式から検索ユーザに対して配布する秘匿検索用ユーザ秘密鍵が発行される。秘匿検索用条件式は、検索ユーザがどの様な属性を持つ文書を検索できるかが指定され、さらに、どの様な条件式で検索キーワードを指定できるかと言う枠組みだけが決定され、その条件がＡＮＤ／ＯＲの論理演算を用いた条件式として記述されたものである。検索キーワード自体は、後に委譲鍵生成アルゴリズムで設定できるようにしておく。

[0047] 次に、暗号化タグ生成アルゴリズムを用いて、検索処理に用いられる暗号化タグが生成される。始めに任意の乱数値が生成され、検索を許可するユーザを制限するために属性が決定され、キーワードが決定される。乱数値と属性とキーワードとを入力として暗号化タグ生成アルゴリズムを用いて暗号化

タグが生成される。具体的には、乱数値を暗号化アルゴリズムで暗号化し、暗号化乱数と乱数値の組みを暗号化タグとする。

[0048] 次に、委譲鍵生成アルゴリズムを用いて、検索クエリが生成される。具体的には、秘匿検索用ユーザ秘密鍵のうち、後でキーワードが指定できるように値が未指定となっている部分に対して検索キーワードを埋め込むことにより、検索クエリが生成される。この検索クエリは、暗号化方式の暗号化委譲秘密鍵に対応するものである。

[0049] 最後に、一致判定アルゴリズムを用いて、受領した検索クエリと暗号化タグの両者に含まれるキーワードが同一であるかどうかの一致判定が行われる。具体的には、暗号化タグの中から暗号化乱数を取り出し、暗号化用委譲秘密鍵に相当する検索クエリを用いて復号する。そして、復号結果が暗号化タグに含まれる乱数値と同一であった場合、すなわち検索クエリで正しく復号できた場合、キーワードが同一であったと判断する。なぜなら、検索クエリ（暗号化用委譲秘密鍵に相当）に含まれるキーワードと、暗号化乱数を生成する時に指定したキーワードが同一でなければ、正しく乱数値を復号できないからである。

[0050] 上述した暗号方式と同様に少し言い方を変えると、秘密鍵生成アルゴリズムでは、検索ユーザの属性情報が鍵属性として設定された秘匿検索用ユーザ秘密鍵が発行される。また、暗号化アルゴリズムでは、検索可能なユーザの属性情報と、キーワードとが暗号属性として設定された暗号化データが生成される。委譲鍵生成アルゴリズムでは、秘匿検索用ユーザ秘密鍵にさらに検索キーワードを鍵属性として追加設定して、検索クエリを生成する。

そして、復号アルゴリズムでは、検索クエリに設定された鍵属性情報と、暗号化タグに設定された属性情報とが対応している場合にのみ、暗号化タグから乱数値を復号できる。つまり、検索クエリに設定されたユーザの属性情報と、暗号化タグに設定されたユーザの属性情報とが対応しており、検索クエリに設定された検索キーワードと、暗号化タグに設定されたキーワードとが対応している場合にのみ、暗号化タグから乱数値を復号できる。

[0051] 遺伝子検索システム１０の処理について説明する。

遺伝子検索システム１０は、初期設定処理、ユーザ秘密鍵発行処理、暗号化処理、検索処理を実行する。なお、暗号化処理には、患者ゲノム配列の暗号化処理と、患者の電子カルテの暗号化処理とが含まれる。

[0052] 図８は、初期設定処理の流れを示すフローチャートである。

初期設定処理は、鍵管理サーバ１００で実行される処理であり、遺伝子検索システム１０を利用する前に１回実行される処理である。

[0053] （Ｓ１０１：エンコード方法決定処理）

マスター鍵生成部１１０は、処理装置により、患者ゲノム配列の差分情報をエンコードする方法を決定する。このエンコードする方法を決定する際に、どの様に検索されるかという検索式についても考慮する。

主要な差分情報として、ＳＮＶ（Ｓｉｎｇｌｅ Ｎｕｃｌｅｏｔｉｄｅ Ｖａｒｉａｎｔ）や、ＳＶ（Ｓｔｒｕｃｔｕｒａｌ Ｖａｒｉａｎｔｓ）や、ＮＣ（Ｎｏｖｅｌ Ｃｏｎｔｉｇ）がある。そこで、この３つを例として、そのエンコード方法と検索方法について説明する。

[0054] 図９は、エンコードしたＳＮＶ情報の構成図である。

ＳＮＶ情報は、ゲノム配列の中で塩基１個が変化していることを示す。ＳＮＶ情報は、患者ＩＤ、染色体番号、位置情報、置換情報１、置換情報２、信頼度で構成される。

患者ＩＤは、別途管理される電子カルテと対応付けるために割り振られた識別番号である。患者ＩＤは、カルテと対応付けられる番号であればよいため、病院で患者に付与した識別番号だけでなく、カルテ保管時にランダムに番号を振って、その値が設定されてもよい。

染色体番号は、ＳＮＶが検出された染色体の番号である。ヒトゲノムの場合、例えば１，２，・・・，２２，Ｘ，Ｙ，Ｍ等の値が設定される。

位置情報は、ＳＮＶが検出された位置に関する情報である。位置は、塩基配列の何番目かという数値で表され、ヒトゲノムの場合は１から３０億ぐらいの値となる。一般に、位置などの数値情報の検索は範囲を指定して検索さ

れる。しかし、上述した秘匿検索方式では範囲検索が実施困難なため、ブロック化を行うことで高速化を図る。例えば、図１０は、検索時に指定される範囲が一般的に５０００以下に抑えられる場合のエンコード方法を示したものである。位置７７７７にＳＮＶが検出された場合、ブロック１とブロック２とが該当する位置と考え、位置情報には「１」と「２」と言う２つの値が設定される。

置換情報１と置換情報２は、ゲノム情報の変化を示す情報である。これは、例えばＳＮＶとしてＧで表される塩基が挿入されていた場合は「Ｇ追加」、Ａで表される塩基が消えていた場合は「Ａ削除」など表記するように定められ、その値が設定される。置換情報が２つあるのは、例えば、Ｇで表される塩基がＡで表される塩基に変化していた場合、置換情報１にはＧ削除、置換情報２にはＡ追加等として表されるためである。

信頼度は、そのＳＮＶ情報の信頼性を表す情報であり、０から１００までの実数で表される。これも位置情報と同様に、ブロック化して検索の高速化を図る。例えば図１１の様に、信頼度を１０％単位で区切り、その対応する値を設定する。例えば、信頼度が７．４４％の場合、対応するブロックの値０が設定される。

[0055] 検索の場合、上記の患者ＩＤから信頼度までの値のうち、条件として指定したいものを幾つか指定し、さらに未指定の場所はどの値とでも一致とみなす“＊”（ワイルドカード）を指定することで差分情報（検索要求）を生成する。そして、ＳＮＶ情報の各要素について、暗号化タグと検索クエリとの間で全てが一致した場合に検索にヒットする、と言う様なルールを決定する。

[0056] また、各要素がどの様な条件で一致するかも決定する。

例えば、患者ＩＤが指定された場合、患者ＩＤが同一のものを一致とみなす。染色体番号も同様に、検索したい染色体番号を指定し、同一のものを一致とみなす。位置情報の場合、例えば７０００～１２０００の範囲を検索したいとしたら、その範囲を含むブロックとしてブロック２を特定し、検索式

には「２」を指定する。そして、ＳＮＶ情報の位置情報として設定されている２つの値のいずれか一方と同一である場合、一致とみなす。置換情報の場合、ＳＮＶ情報に指定した値と同様のルールで検索したい値を設定し、両者が同一であった場合に一致とみなす。信頼度も同様に、同じ値が指定されている場合に一致とみなす。

以上のように、ＳＮＶ情報のエンコード方法と、どの様に検索を実施できるかという方式を決定する。

[0057] 図１２は、エンコードしたＳＶ情報の構成図である。

ＳＮＶ情報は塩基１つが変化していることを示したが、ＳＶ情報は連続する複数の塩基配列が変化していることを示す。ＳＶ情報は、患者ＩＤ、染色体番号、開始位置情報、終了位置情報、変異種別、ＣＮＶゲイン、組換染色体番号、組換染色体開始位置情報、組換染色体終了位置情報、分類ＩＤ（挿入配列）、バージョン（挿入配列）、染色体番号（挿入配列）、開始位置情報（挿入配列）、終了位置情報（挿入配列）で構成される。

患者ＩＤ、染色体番号は、図９に示したＳＮＶと同一であるので、説明は省略する。

開始位置情報は、変化が開始している位置を示す情報である。ＳＮＶ情報の位置情報と同様の手順で位置情報がブロック化されて、ブロックを示す値が設定される。

終了位置情報は、変化が終了している位置を示す情報である。ＳＮＶ情報の位置情報と同様の手順で位置情報がブロック化されて、ブロックを示す値が設定される。

変異種別は、ＳＶの種別を示す情報である。例えば、遺伝子配列の繰り返し回数の追加／削減を示す「ＣＮＶゲイン」や「ＣＮＶロス」、塩基配列の反転を示す「Ｉｎｖｅｒｓｉｏｎ」、塩基配列の大規模挿入を示す「Ｉｎｓｅｒｔｉｏｎ」、塩基配列の大規模欠損を示す「Ｄｅｌｅｔｉｏｎ」、他染色体との組換を示す「組換え」などの値が設定される。

ＣＮＶゲインは、変位種別として、「ＣＮＶゲイン」と「ＣＮＶロス」と

が設定された場合に利用される。ＣＮＶゲインは、変異種別として「ＣＮＶゲイン」が設定された場合は、繰り返し回数が何回増えているか、「ＣＮＶロス」が設定された場合は、繰り返し回数が何回減ったかを示す情報が設定される。この値も整数値が含まれるため、ブロック化して値を設定する。例えば、図１３に示すようにブロック化するというルールが定められ、ＣＮＶゲインが１５であった場合、対応するブロックの番号「３」が設定される。

組換染色体番号は、変異種別として「組換え」が指定された場合に利用される。組換染色体番号は、組み換えられた他染色体の番号が設定される。設定可能な値は、染色体番号と同一である。

組換染色体開始位置情報と組換染色体終了位置情報とは、変異種別として「組換え」が指定された場合に利用される。組換染色体開始位置情報と組換染色体終了位置情報とは、他染色体のどの塩基配列と組み換えが発生したかを示すため、他染色体における組み換えの開始位置と終了位置とを示す情報である。設定可能な値は、開始位置情報と終了位置情報と同一である。

分類ＩＤ（挿入配列）は、変異種別として「Ｉｎｓｅｒｔｉｏｎ」が指定された場合に利用される。分類ＩＤ（挿入配列）は、挿入された塩基配列が、他のどの生物のものかを示す番号が設定される。

バージョン（挿入配列）は、上記の分類ＩＤを決定するのに用いたゲノム・データベースのバージョンを示すための数値である。

染色体番号（挿入配列）、開始位置情報（挿入配列）、終了位置情報（挿入配列）は、他生物のどの部分の塩基配列が挿入されたかを示す情報で、設定される値は染色体番号や開始位置情報や終了位置情報と同一の手法でエンコードされる。但し、ゲノム配列の長さは生物種によって異なるため、分類ＩＤ毎に異なるブロックの区切り方を用意することが好ましい。

[0058] また、検索の場合、上記の患者ＩＤから終了位置情報（挿入配列）までの値のうち、条件として指定したいものを幾つか指定し、さらに未指定の場所はどの値とでも一致とみなす“＊”（ワイルドカード）を指定することで差分情報（検索要求）を生成し、これをもとに検索クエリを生成する。そして

、ＳＶ情報の各要素について、暗号化タグと検索クエリとの間で全てが一致した場合に検索にヒットする、と言う様なルールを決定する。

[0059] また、各要素がどの様な条件で一致するかも決定する。

患者ＩＤ、染色体番号は、ＳＮＶ情報の場合と同一である。開始位置情報、終了位置情報、組換染色体開始位置情報、組換染色体終了位置情報、開始位置情報（挿入配列）、終了位置情報（挿入配列）は、ＳＮＶ情報の位置情報と同一に扱えばよい。変異種別は、どの種類の変異種別を検索したいかを検索クエリに指定し、同一の値が設定されたＳＶ情報を一致とみなす。ＣＮＶゲインは、例えば１０から３０までの間にあるものを検索したい場合、その対応するブロックの番号「３」を指定する。そして、ＳＶ情報に含まれるＣＮＶゲインと同一であった場合、一致とみなす。組換染色体番号、染色体番号（挿入配列）は、染色体番号と同一で扱えばよい。分類ＩＤ、バージョンは、検索クエリに検索したい分類ＩＤとバージョンを入れ、ＳＶ情報と同一の場合に一致とみなす。

以上のように、ＳＶ情報のエンコード方法と、どの様に検索を実施できるかという方式を決定する。

[0060] 図１４は、エンコードしたＮＣ情報の構成図である。

ＮＣ情報は、基準ゲノムにマッピングできなかった塩基を示す。塩基が基準ゲノムにマッピングできなかったのは、多くの場合は、ウイルスに感染しているなどで特殊なゲノムが検知された場合と考えられる。そのため、ＮＣ情報は、患者ＩＤに加え、分類ＩＤ（挿入配列）、バージョン（挿入配列）、染色体番号（挿入配列）、開始位置情報（挿入配列）、終了位置情報（挿入配列）で構成される。

値の設定の方法や、検索の方法はＳＶ情報と同一であるため、説明は省略する。

[0061] （Ｓ１０２：方式決定処理）

マスター鍵生成部１１０は、処理装置により、利用する秘匿検索方式や、データ本体を暗号化する暗号方式を決定する。ここでは、複数の検索キーワ

ードを指定できる秘匿検索方式を必要としているため、上述した階層的内積述語暗号を用いた秘匿検索方式を用いる。同様に、暗号方式も上述した階層的内積述語暗号を用いた暗号方式を用いる。

そして、マスター鍵生成部１１０は、秘匿検索方式の利用方法を決定する。ここでは、タグＩＤの階層構造を決定する。例えば、図１５に示す様に、タグＩＤは３要素で構成され、検索可能なユーザが所属するグループのグループ名を格納するグループ名欄、氏名などを格納するユーザ名欄、患者ゲノム配列の差分情報を格納する差分情報欄で構成する。検索の際は、グループ名、ユーザ名、差分情報の全てが合致と判断された場合にのみ、検索にヒットしたとみなすルールとする。

同様に、マスター鍵生成部１１０は、データ本体を暗号化するための暗号化方式の利用方法を決定する。ここでは、復号者ＩＤの階層構造を決定する。例えば、図１６に示す様に、復号者ＩＤは２要素で構成され、復号可能なユーザが所属するグループのグループ名を格納するグループ名欄、氏名等を格納するユーザ名欄で構成する。復号の際は、グループ名、ユーザ名の全てが合致したものだけが復号できるルールとする。

[0062] （Ｓ１０３：ユーザＩＤ記憶処理）

ユーザＩＤ記憶部１５０は、ユーザＩＤを保管するユーザＩＤ情報データベースを構築する。ユーザＩＤ情報データベースは、ユーザ秘密鍵を生成するために必要な情報と、暗号化装置２００がデータを暗号化する際に、相手のグループ名／ユーザ名を特定するために必要な情報とが記憶されるものである。

例えば、図１７に示すように、ユーザＩＤ情報データベースには、グループ名である会社名、ユーザ名である氏名、所属情報、有効期間等が格納される。また、ユーザＩＤ情報データベースには、最新の状況だけでなく、過去の履歴を全て格納するようにしてもよい。

[0063] （Ｓ１０４：マスター鍵生成処理）

マスター鍵生成部１１０は、処理装置により、秘匿検索方式のマスター鍵

生成アルゴリズムを実行して、秘匿検索用マスター鍵と秘匿検索用公開パラメータとを生成する。同様に、マスター鍵生成部１１０は、処理装置により、暗号化方式のマスター鍵生成アルゴリズムを実行して暗号化用マスター鍵と暗号化用公開パラメータとを生成する。

以降、秘匿検索用マスター鍵と暗号化用マスター鍵をまとめてマスター鍵、秘匿検索用公開パラメータと暗号化用公開パラメータをまとめて公開パラメータと呼ぶ。

[0064] （Ｓ１０５：マスター鍵記憶処理）

鍵記憶部１２０は、マスター鍵生成部１１０が生成したマスター鍵と公開パラメータとを記憶装置に記憶する。

[0065] （Ｓ１０６：公開パラメータ公開処理）

データ送受信部１４０は、鍵記憶部１２０が記憶した公開パラメータを暗号化装置２００、検索装置３００、データセンタ４００に対してネットワーク５００を介して公開する。

なお、公開された公開パラメータは、暗号化装置２００では、公開パラメータ記憶部２３０に記憶され、検索装置３００では、ユーザ秘密鍵記憶部３２０に記憶され、データセンタ４００では、公開パラメータ記憶部４５０に記憶される。

[0066] 以上の手順によって、遺伝子検索システム１０のセットアップが完了する。

なお、Ｓ１０３で生成したユーザＩＤ情報データベースは、システムの運用において、ユーザの人事異動や入社や退社があるたびに内容がメンテナンスされる。

[0067] 図１８は、ユーザ秘密鍵発行処理の流れを示すフローチャートである。

ユーザ秘密鍵発行処理は、主に、鍵管理サーバ１００と検索装置３００とで実行される処理であり、新規ユーザを追加した場合や、ユーザの所属するグループ名が変わった時などに実行される。

[0068] （Ｓ２０１：ユーザＩＤ取得処理）

ユーザ秘密鍵生成部１３０は、ユーザＩＤ記憶部１５０が保持しているユーザＩＤ情報データベースから、ユーザ秘密鍵を発行するユーザのグループ名とユーザ名とを取得する。

[0069] （Ｓ２０２：ユーザ秘密鍵生成処理）

ユーザ秘密鍵生成部１３０は、処理装置により、検索クエリを生成するために用いる秘匿検索用ユーザ秘密鍵と、暗号化データを復号するために用いる暗号化用ユーザ秘密鍵とを生成する。

秘匿検索方式では、秘匿検索用ユーザ秘密鍵を生成する際にタグＩＤ階層構造を指定する必要がある。ここでは、Ｓ２０１で取得したグループ名をグループ名欄に、同じくユーザ名をユーザ名に設定し、差分情報は後で検索するユーザが設定できるように委譲可能な要素として指定する事で、秘匿検索用ユーザ秘密鍵を生成できる。

同様に、暗号化方式で暗号化用ユーザ秘密鍵を生成する際に復号者ＩＤ階層構造を指定する必要がある。ここでは、Ｓ２０１で取得したグループ名をグループ名欄に、同じくユーザ名をユーザ名に指定する事で、暗号化用ユーザ秘密鍵を生成できる。

上記で生成した秘匿検索用ユーザ秘密鍵、暗号化用ユーザ秘密鍵をまとめてユーザ秘密鍵と呼ぶ。

[0070] （Ｓ２０３：ユーザ秘密鍵送信処理）

データ送受信部１４０は、Ｓ２０２で生成したユーザ秘密鍵を、検索装置３００へ送信する。

[0071] （Ｓ２０４：ユーザ秘密鍵受信処理）

ユーザ秘密鍵記憶部３２０は、Ｓ２０３にて送信されたユーザ秘密鍵を受信し、記憶装置に記憶する。

[0072] 図１９は、患者ゲノム配列の暗号化処理の流れを示すフローチャートである。

患者ゲノム配列の暗号化処理は、主に、暗号化装置２００とデータセンタ４００とで実行される処理であり、患者ゲノム配列を暗号化してデータセン

タ４００に保管する時に実行される。

[0073] （Ｓ３０１：差分情報抽出処理）

基準遺伝子取得部２１０は、例えばインターネットで公開されている基準ゲノム配列を取得する。また、対象遺伝子入力部２２０は、入力装置により、患者ゲノム配列を入力する。

差分情報生成部２４０は、処理装置により、患者ゲノム配列と基準ゲノム配列とを比較することで、ＳＮＶやＳＶやＮＣ等の差分情報を生成する。この差分情報を生成する方法としては、ＣｈＩＰ－ｓｅｑ法またはＲＮＡ－ｓｅｑ法またはＭｅＤＩＰ－ｓｅｑ法または変異解析法またはバイサルファイト法などが知られており、それらの一般的な手法を用いる。

[0074] （Ｓ３０２：ユーザ決定処理）

データ暗号化部２６０は、暗号化装置２００を操作するユーザに、暗号化データを復号可能なユーザのグループ名及びユーザ名を入力させる。同様に、暗号化タグ生成部２７０は、暗号化データを検索可能なユーザのグループ名及びユーザ名を入力させる。

ここで入力されるグループ名やユーザ名は一つである必要はなく、復号又は検索可能なユーザが複数名いる場合は複数入力することもできる。なお、ここで利用する秘匿検索方式や暗号化方式は、グループ名やユーザ名として誰でもよい事を意味するワイルドカードを受け取ることも可能である。

[0075] （Ｓ３０３：データ暗号化処理）

データ暗号化部２６０は、処理装置により、Ｓ３０２で入力された復号可能なグループ名及びユーザ名を用いて、Ｓ３０１で入力された患者ゲノム配列を暗号化する。

具体的には、データ暗号化部２６０は、ランダムにセッション鍵を生成し、そのセッション鍵でＡＥＳやＣａｍｅｌｌｉａ（登録商標）等の共通鍵暗号を用いて患者ゲノム配列を暗号化し、暗号化データ本体を生成する。次に、データ暗号化部２６０は、復号者ＩＤ階層構造のグループ名とユーザ名に、Ｓ３０２で入力された復号可能なグループ名及びユーザ名をそれぞれ指定

して、これを暗号化用公開鍵として、Ｓ１０２で決定した暗号化方式を用いてセッション鍵を暗号化し、暗号化セッション鍵を生成する。そして、データ暗号化部２６０は、前述の２個の暗号化結果（暗号化データ本体と暗号化セッション鍵）を組み合わせることで、暗号化データを生成する。

生成した暗号化データのデータ構造を図２０の符号６０３部分に示す。なお、Ｓ３０２にて複数のグループ名とユーザ名を受領している場合、グループ名とユーザ名の各組に対して暗号化セッション鍵を生成する必要がある。

[0076] （Ｓ３０４：差分情報エンコード処理）

差分情報エンコード部２５０は、処理装置により、Ｓ３０１で生成された各差分情報を、Ｓ１０１で決定されたエンコード方法に従ってエンコードして、エンコード差分情報を生成する。また、差分情報エンコード部２５０は、ユーザから患者ＩＤを入力させ、エンコード差分情報に含める。

[0077] （Ｓ３０５：暗号化タグ生成処理）

暗号化タグ生成部２７０は、エンコード差分情報を暗号化して暗号化タグを生成する。

具体的には、暗号化タグ生成部２７０は、処理装置により、タグＩＤ階層構造のグループ名とユーザ名にＳ３０２で入力させた検索可能なグループ名及びユーザ名を指定し、差分情報欄にＳ３０４にてエンコードしたエンコード差分情報を指定して、秘匿検索方式で乱数値を暗号化して暗号化タグを生成する。また、暗号化タグ生成部２７０は、乱数値を平文のまま暗号化タグに含める。

なお、上記処理は差分情報１個に対する処理であるため、この処理は各エンコード差分情報に対して実施される。例えば、ＳＮＶ、ＳＶ、ＮＣの各エンコード差分情報に対して実施される。また、Ｓ３０２にてグループ名及びユーザ名の組が複数の入力された場合、グループ名とユーザ名の各組に対して暗号化タグが生成される。

[0078] （Ｓ３０６：保管依頼処理）

タグ付き暗号化データ生成部２８０は、処理装置により、Ｓ３０３で生成

された暗号化データと、Ｓ３０５で生成された暗号化タグと、Ｓ３０４で入力された患者ＩＤとを結合してタグ付き暗号化データを生成する（図２０の符号６０１）。そして、タグ付き暗号化データ生成部２８０は、生成したタグ付き暗号化データをデータセンタ４００へ送信して、保管を依頼する。

この時、データセンタ４００でタグ付き暗号化データの保管を容易にするため、タグ付き暗号化データ生成部２８０は、Ｓ３０２で入力された復号可能なグループ名及びユーザ名と、検索可能なグループ名及びユーザ名とを共に送信する。図２０で示したタグ付き暗号化データの構成では、復号可能なグループ名及びユーザ名と、検索可能なグループ名及びユーザ名とをタグ付き暗号化データに含めている。

[0079] （Ｓ３０７：暗号化データ保管処理）

保管要求処理部４１０は、処理装置により、暗号化装置２００から受信したタグ付き暗号化データを分解し、暗号化データと複数の暗号化タグと患者ＩＤとを取り出す。そして、保管要求処理部４１０は、暗号化データ記憶部４２０に暗号化データを患者ＩＤとともに保管させる。

なお、暗号化データ記憶部４２０は、タグ付き暗号化データに含まれるグループ名とユーザ名毎に暗号化データを分けて保管し、さらに保管した暗号化データに管理番号を付与して、後で管理番号から暗号化データを一意に特定できる様にする。暗号化データが複数のグループ名やユーザ名と関連付けられている場合、各グループ名とユーザ名に対して暗号化データを関連付けて保管する。複数のグループ名とユーザ名に関連付けられている場合、暗号化データは１個だけ格納し、他は参照情報のみを保管する様にすることでディスク容量の節約が可能である。

図２１は、暗号化データの保管例を示す図である。図２１に示す様に、保管要求処理部４１０は、グループ名が“Ａ製薬会社”でユーザ名が“＊”（ワイルドカード）である暗号化データと患者ＩＤと管理番号とをまとめて管理し、さらにグループ名が“Ｂ病院”でユーザ名が“＊”である暗号化データと患者ＩＤと管理番号とをまとめて管理する。また、Ａ製薬会社とＢ病院

の両方に開示されるデータがあった場合、例えば管理番号０００００１に患者ＩＤと暗号化データ本体を関連付けて保管し、管理番号１００００２には患者ＩＤに加え、暗号化データとして管理番号０００００１を参照する様なポインタを保管する。

[0080] （Ｓ３０８：暗号化タグ保管処理）

保管要求処理部４１０は、Ｓ３０７で取り出した複数の暗号化タグを、対応する暗号化データの管理番号及び患者ＩＤと共に暗号化タグ記憶部４３０に保管させる。暗号化タグ記憶部４３０は、タグ付き暗号化データに含まれるグループ名とユーザ名毎に分けて、暗号化タグと管理番号と患者ＩＤとを保管する。

[0081] 図２２は、患者の電子カルテの暗号化処理の流れを示すフローチャートである。

患者の電子カルテの暗号化処理は、主に、暗号化装置２００とデータセンタ４００とで実行される処理であり、電子カルテを暗号化してデータセンタ４００に保管する時に実行される。

[0082] （Ｓ４０１：ユーザ決定処理）

データ暗号化部２６０は、暗号化装置２００を操作するユーザに、電子カルテを復号可能なユーザのグループ名及びユーザ名を入力させる。

ここで入力されるグループ名やユーザ名は一つである必要はなく、復号可能なユーザが複数名いる場合は複数入力することもできる。

[0083] （Ｓ４０２：データ暗号化処理）

データ暗号化部２６０は、ユーザに患者ＩＤと電子カルテとを入力させる。そして、データ暗号化部２６０は、処理装置により、Ｓ４０１で入力されたグループ名及びユーザ名を用いて、電子カルテを暗号化する。具体的な暗号化方法は、Ｓ３０３で患者ゲノム配列を暗号化した流れと同一のため、詳細は省略する。

[0084] （Ｓ４０３：保管依頼処理）

データ暗号化部２６０は、Ｓ４０２にて生成した暗号化データを、誰の電

子カルテかを示す患者ＩＤと、復号可能なユーザのグループ名及びユーザ名と共にデータセンタ４００へ送信して、保管を依頼する。

[0085] （Ｓ４０４：暗号化データ保管処理）

保管要求処理部４１０は、暗号化装置２００から受領した暗号化データを、患者ＩＤと関連付けて暗号化データ記憶部４２０に保管させる。

[0086] 図２３は、検索処理の流れを示すフローチャートである。

検索処理は、主に、検索装置３００とデータセンタ４００とで実行される処理であり、データセンタ４００に保管された暗号化患者ゲノム配列を取得する時に実行される。

[0087] （Ｓ５０１：差分情報入力処理）

差分情報入力部３１０は、検索装置３００を操作するユーザに、基準ゲノム配列との差分情報を検索キーワードとして含む検索要求を入力させる。

ここで入力される差分情報は、患者ゲノムから抽出した差分情報の様にすべての要素が指定されている必要はなく、例えば染色体番号だけを指定したものや、位置情報だけを指定したものなどでよい。

[0088] （Ｓ５０２：差分情報エンコード処理）

差分情報エンコード部３３０は、処理装置により、Ｓ５０１で入力された差分情報をエンコードして、エンコード差分情報を生成する。この処理は、ステップＳ３０４の処理と同一であるため詳細は省略する。但し、指定されていない要素は“＊”としておく点だけ注意が必要である。

[0089] （Ｓ５０３：検索クエリ生成処理）

検索クエリ生成部３４０は、処理装置により、Ｓ５０２で生成されたエンコード差分情報と、ユーザ秘密鍵記憶部３２０で記憶されているユーザ秘密鍵を用いて、検索クエリを生成する。そして、検索クエリ生成部３４０は、生成した検索クエリをデータセンタ４００へ送信する。

この時、ユーザ自身のグループ名及びユーザ名も送信する。なお、グループ名やユーザ名の信頼性を検証するため、検索端末を操作するユーザのユーザ認証も行う。

[0090] （Ｓ５０４：暗号化タグ抽出処理）

検索要求処理部４４０は、処理装置により、暗号化タグ記憶部４３０に記憶された全ての暗号化タグの中から、Ｓ５０３で検索クエリとともに送信されたグループ名及びユーザ名で検索が可能な暗号化タグを全て取得する。また、検索要求処理部４４０は、アクセス権限記憶部４６０から該当するグループ名及びユーザ名のユーザがアクセス可能な患者ＩＤのリストを取り出し、その患者ＩＤに対応する暗号化タグのみに取得した暗号化タグを絞込む。

なお、アクセス権限記憶部４６０は、図２４に示すようなアクセス権限管理テーブルを持っており、グループ名及びユーザ名をアクセス者情報として、検索できる患者ゲノム配列に対応する患者ＩＤを特定して、その患者ＩＤに対応する暗号化タグを出力すればよい。

[0091] （Ｓ５０５：一致判定処理）

検索要求処理部４４０は、処理装置により、Ｓ５０４で絞り込んだ暗号化タグに対して秘匿検索方式の一致判定処理を行い、暗号化タグに含まれる差分情報が、Ｓ５０３で送信された検索クエリに含まれる差分情報で指定された条件に合致するか否かを判定する。

秘匿検索方式の一致判定処理は、１個の暗号化タグと、１個の検索クエリとの比較しか実施できない。そのため、一致判定処理をＳ５０４にて取得した全ての暗号化タグに対して実施する。そして、判定処理の結果、一致と判定された暗号化タグに関連付けられた管理番号を特定する。

[0092] （Ｓ５０６：暗号化データ取得処理）

検索要求処理部４４０は、Ｓ５０５で特定された管理番号に対応する暗号化データを、暗号化データ記憶部４２０から全て取得し、対応する患者ＩＤと共に検索装置３００へ送信する。

[0093] （Ｓ５０７：暗号化データ復号処理）

データ復号部３６０は、処理装置により、Ｓ５０６でデータセンタ４００から受信した暗号化データを、ユーザ秘密鍵記憶部３２０に記憶された暗号化用ユーザ秘密鍵を用いて、暗号化方式の復号処理を実行することにより復

号する。データ復号部３６０は、この処理を、受信した全ての暗号化データに対して実施する。

[0094] 以上の手順により、検索装置３００はユーザから検索したい差分情報を受け取り、その差分情報に合致する暗号化データをデータセンタ４００から取得し、それを復号して患者ゲノム配列を閲覧することができる。また、必要に応じて、暗号化データに対応する患者ＩＤをデータセンタ４００へ送信し、対応する電子カルテなどを入手することもできる。

[0095] 以上のように、実施の形態１に係る遺伝子検索システム１０では、患者ゲノムを秘匿検索方式を用いて暗号化してデータセンタ４００に保管し、また検索要求も秘匿検索技術で暗号化してデータセンタ４００に検索を依頼している。そのため、データセンタ４００では患者ゲノムの中身を一切知ることができないにもかかわらず、検索というサービスを提供することができる。

[0096] また、ヒトゲノム情報は３０億塩基からなる非常に膨大なデータである。そのため、ヒトゲノム情報を全てタグとしてしまうと、暗号化によってさらにデータサイズが増加することもあり、ディスク容量やネットワーク容量を圧迫する原因となってしまう。

しかし、実施の形態１に係る遺伝子検索システム１０では、一般公開されている基準ゲノム配列との差分情報に絞ってタグとしたので、ディスク使用量やネットワーク容量を大きく削減することができる。

[0097] また、実施の形態１に係る遺伝子検索システム１０では、ＳＮＶの位置情報や信頼度などの数値情報をブロック化することで、秘匿検索にて実現困難であった範囲検索を、一致検索を用いて実現した。そのため、ゲノム検索で利用する範囲検索にも対応することができる。

特に、図１０に示したＳＮＶの位置情報では、各ブロックが重複するようにした。つまり、ブロック１とブロック２とは、位置５０００から１００００が重複しており、ブロック２とブロック３とでは、位置１００００から１５０００が重複している。これにより、検索時に指定される範囲が５０００以下の場合には、全てのブロックを用いて検索を行い、検索時に指定される

範囲が１００００以下の場合には、奇数のブロックだけで検索を行えばより高速に処理をすることが可能となる。

[0098] また、実施の形態１に係る遺伝子検索システム１０では、染色体番号や位置情報などの複数の検索条件が指定された場合でも、個別に条件の一致判定を行うのではなく、内積述語暗号を用いて全ての検索条件が成立するかどうかを一括で判定する。そのため、部分的に検索にヒットしたことがサーバには分からず、セキュリティが高い。

[0099] また、実施の形態１に係る遺伝子検索システム１０では、暗号化データと共に患者ＩＤを含めるようにした。そのため、検索結果として入手した患者ＩＤから、関連する電子カルテなどの情報を引き出すことができる。したがって、該当する塩基変異が発生している場合、どの様な病気に関連しているかも研究することが可能となる。

[0100] また、実施の形態１に係る遺伝子検索システム１０では、患者ゲノム配列の差分情報を暗号化する際に、タグＩＤ階層構造や復号者ＩＤ階層構造に対して、グループ名やユーザ名を含めるようにした。そのため、検索や復号ができる研究者や医者を制限することができる。例えば、グループ名に「Ａ製薬会社」、ユーザ名にワイルドカード「＊」を指定すれば、その患者の情報はＡ製薬会社の社員に制限することができる。また、グループ名もユーザ名もワイルドカード「＊」にして暗号化することで、そのシステムに登録された医者や研究者であれば、その患者の情報を利用することができるようになる。

[0101] また、実施の形態１に係る遺伝子検索システム１０では、暗号によるアクセス制御とは別に、データセンタ４００がアクセス権限管理テーブルを保管し、この情報に基づくアクセス制御も可能とした。そのため、患者からの要望によって、「ゲノム配列のみ閲覧可能」とか「カルテの閲覧が可能」などを細かく管理することができる。つまり、この情報を元にして、患者ゲノム配列からの検索要求を許可するかどうかを判定するようにしたため、きめ細かなアクセス制御が可能となる。

[0102] また、実施の形態１に係る遺伝子検索システム１０では、検索者のユーザ秘密鍵に対してグループ名やユーザ名を含めるようにした。そのため、そのユーザ秘密鍵から生成された検索クエリに含まれるグループ名やユーザ名を確認することで、認証を行うことも可能である。

[0103] なお、上記説明では、会社単位でアクセス制限を行う場合の例を示した。しかし、このアクセス制御の単位は一例である。例えば、医者や看護婦などの国家資格の条件を入れてもよいし、国家プロジェクトへの参加者かどうかのフラグを立ててもよい。これらのＩＤ階層構造は一例であるため、色々な要素を追加したり、削除したりすることもできる。

[0104] また、上記説明では、国が唯一の鍵管理サーバ１００を管理し、医者の代理でゲノム分析を行ったシーケンサ操作者が暗号化装置２００を利用する場合を想定していた。しかし、利用システムに応じて、システム構成は柔軟に変更することができる。例えば、シーケンサ操作者から検知結果を受け取った医者が暗号化装置２００を操作して患者ゲノム配列を暗号化してもよい。

[0105] また、上記説明では、検索装置３００にユーザ秘密鍵を保管して、検索クエリ生成や暗号化データの復号も実施させるようにした。しかし、より安全性を高めるために、ユーザ秘密鍵の管理を検索装置３００ではなくＩＣカードなどのデバイスを用いて実施してもよい。この場合、ＩＣカード内で安全にユーザ秘密鍵が管理されるため、セキュリティの向上が図れる。

[0106] また、上記説明では、検索端末で利用する秘匿検索用ユーザ秘密鍵は、ＳＮＶ情報やＳＶ情報やＮＣ情報などの検索で共通に使えるような鍵として生成した。しかし、各差分情報は長さが異なるため、共通の秘匿検索用ユーザ秘密鍵を用いるのではなく、ＳＮＶ情報秘匿検索用ユーザ秘密鍵やＳＶ情報秘匿検索用ユーザ秘密鍵やＮＣ情報秘匿検索用ユーザ秘密鍵など、用途に応じて秘匿検索用ユーザ秘密鍵を個別に生成してもよい。この場合、秘匿検索用ユーザ秘密鍵の長さが各情報の長さに最適化したものとなるため、演算時間が高速化される。

[0107] また、上記説明では、グループ名やユーザ名を示す際に「Ａ製薬会社」や

「田中」などの文字列を使って表現した。これは、実施例としての分かりやすさを優先したためで、実際には文字列だけでなく番号などのＩＤを利用してもよい。染色体番号などの他の要素も同様である。

[0108] また、上記説明では、秘匿検索方式や暗号方式として階層的内積述語暗号を用いた場合の例を示した。しかし、同様の機能を持った暗号であれば、階層的内積述語暗号に限る必要はない。また、秘匿検索方式と暗号方式で、異なる方式を利用してもよい。

[0109] また、上記説明では、ユーザＩＤ記憶部１５０において、ユーザの過去の属性情報も管理できるようにした。これは管理上必要な場合にのみ実施すればよく、現時点の属性情報のみ管理するようにしてもよい。

[0110] また、上記説明では、全員が共通で１つの基準ゲノム配列を使用することを仮定して実施するケースを示した。しかし、異なる基準ゲノムを用いて実施することも可能である。この場合、暗号化する際は全ての基準ゲノム配列に対して差分情報を生成し、検索する場合はいずれかひとつの基準ゲノム配列との差分情報を検索クエリとすることも可能であるし、暗号化する際に一つの基準ゲノム配列との差分情報を生成し、検索する場合にすべての基準ゲノム配列との差分情報を検索クエリにすることも可能である。

[0111] また、上記説明では、位置情報や信頼度やＣＮＶゲインなど範囲検索を必要とするものに関しては、ブロック化を行うことで、キーワードの完全一致で判定ができるようにした。しかし、検索したい範囲は用途によって異なることが想定されるため、必ずしも完全一致で検索できるようにする必要はない。例えば、暗号化時はブロック１０に属することを指定しつつも、検索時はブロック１０またはブロック１１の様に複数ブロックを検索範囲として指定するようにしてもよい。

[0112] また、上記説明では、鍵管理サーバ１００が１台しか無いケースを想定した。しかし、秘匿検索方式や暗号化方式として用いた階層的内積述語暗号は、鍵管理サーバ１００を階層化して複数に分散して運用することも可能である。そのため、ここでも、鍵管理サーバ１００を複数に階層化して運用する

ことも可能である。

[0113] また、上記説明では、差分情報としてＳＮＶ情報や、ＳＶ情報や、ＮＣ情報を暗号化する例を示した。しかし、同様の仕組みによって、前記以外の差分情報を暗号化することも可能である。

[0114] 図２５は、鍵管理サーバ１００、暗号化装置２００、検索装置３００、データセンタ４００のハードウェア構成の一例を示す図である。

図２５に示すように、鍵管理サーバ１００、暗号化装置２００、検索装置３００、データセンタ４００は、プログラムを実行するＣＰＵ９１１（Ｃｅｎｔｒａｌ・Ｐｒｏｃｅｓｓｉｎｇ・Ｕｎｉｔ、中央処理装置、処理装置、演算装置、マイクロプロセッサ、マイクロコンピュータ、プロセッサともいう）を備えている。ＣＰＵ９１１は、バス９１２を介してＲＯＭ９１３、ＲＡＭ９１４、ＬＣＤ９０１（Ｌｉｑｕｉｄ Ｃｒｙｓｔａｌ Ｄｉｓｐｌａｙ）、キーボード９０２（Ｋ／Ｂ）、通信ボード９１５、磁気ディスク装置９２０と接続され、これらのハードウェアデバイスを制御する。磁気ディスク装置９２０（固定ディスク装置）の代わりに、光ディスク装置、メモリカード読み書き装置などの記憶装置でもよい。磁気ディスク装置９２０は、所定の固定ディスクインタフェースを介して接続される。

[0115] ＲＯＭ９１３、磁気ディスク装置９２０は、不揮発性メモリの一例である。ＲＡＭ９１４は、揮発性メモリの一例である。ＲＯＭ９１３とＲＡＭ９１４と磁気ディスク装置９２０とは、記憶装置（メモリ）の一例である。また、キーボード９０２、通信ボード９１５は、入力装置の一例である。また、通信ボード９１５は、通信装置の一例である。さらに、ＬＣＤ９０１は、表示装置の一例である。

[0116] 磁気ディスク装置９２０又はＲＯＭ９１３などには、オペレーティングシステム９２１（ＯＳ）、ウィンドウシステム９２２、プログラム群９２３、ファイル群９２４が記憶されている。プログラム群９２３のプログラムは、ＣＰＵ９１１、オペレーティングシステム９２１、ウィンドウシステム９２２により実行される。

[0117] プログラム群９２３には、上記の説明において「マスター鍵生成部１１０」、「ユーザ秘密鍵生成部１３０」、「データ送受信部１４０」、「基準遺伝子取得部２１０」、「対象遺伝子入力部２２０」、「差分情報生成部２４０」、「差分情報エンコード部２５０」、「データ暗号化部２６０」、「暗号化タグ生成部２７０」、「タグ付き暗号化データ生成部２８０」、「差分情報入力部３１０」、「差分情報エンコード部３３０」、「検索クエリ生成部３４０」、「遺伝子情報取得部３５０」、「データ復号部３６０」、「保管要求処理部４１０」、「検索要求処理部４４０」等として説明した機能を実行するソフトウェアやプログラムやその他のプログラムが記憶されている。プログラムは、ＣＰＵ９１１により読み出され実行される。

ファイル群９２４には、上記の説明において「鍵記憶部１２０」、「ユーザＩＤ記憶部１５０」、「公開パラメータ記憶部２３０」、「ユーザ秘密鍵記憶部３２０」、「暗号化データ記憶部４２０」、「暗号化タグ記憶部４３０」、「公開パラメータ記憶部４５０」、「アクセス権限記憶部４６０」等に記憶される情報やデータや信号値や変数値やパラメータが、「ファイル」や「データベース」の各項目として記憶される。「ファイル」や「データベース」は、ディスクやメモリなどの記憶媒体に記憶される。ディスクやメモリなどの記憶媒体に記憶された情報やデータや信号値や変数値やパラメータは、読み書き回路を介してＣＰＵ９１１によりメインメモリやキャッシュメモリに読み出され、抽出・検索・参照・比較・演算・計算・処理・出力・印刷・表示などのＣＰＵ９１１の動作に用いられる。抽出・検索・参照・比較・演算・計算・処理・出力・印刷・表示のＣＰＵ９１１の動作の間、情報やデータや信号値や変数値やパラメータは、メインメモリやキャッシュメモリやバッファメモリに一時的に記憶される。

[0118] また、上記の説明におけるフローチャートの矢印の部分は主としてデータや信号の入出力を示し、データや信号値は、ＲＡＭ９１４のメモリ、その他光ディスク等の記憶媒体やＩＣチップに記録される。また、データや信号は、バス９１２や信号線やケーブルその他の伝送媒体や電波によりオンライン

伝送される。

また、上記の説明において「～部」として説明するものは、「～回路」、「～装置」、「～機器」、「～手段」、「～機能」であってもよく、また、「～ステップ」、「～手順」、「～処理」であってもよい。また、「～装置」として説明するものは、「～回路」、「～機器」、「～手段」、「～機能」であってもよく、また、「～ステップ」、「～手順」、「～処理」であってもよい。さらに、「～処理」として説明するものは「～ステップ」であっても構わない。すなわち、「～部」として説明するものは、ＲＯＭ９１３に記憶されたファームウェアで実現されていても構わない。或いは、ソフトウェアのみ、或いは、素子・デバイス・基板・配線などのハードウェアのみ、或いは、ソフトウェアとハードウェアとの組み合わせ、さらには、ファームウェアとの組み合わせで実施されても構わない。ファームウェアとソフトウェアは、プログラムとして、ＲＯＭ９１３等の記録媒体に記憶される。プログラムはＣＰＵ９１１により読み出され、ＣＰＵ９１１により実行される。すなわち、プログラムは、上記で述べた「～部」としてコンピュータ等を機能させるものである。あるいは、上記で述べた「～部」の手順や方法をコンピュータ等に実行させるものである。

符号の説明

[0119] １０　遺伝子検索システム、１００　鍵管理サーバ、１１０　マスター鍵生成部、１２０　鍵記憶部、１３０　ユーザ秘密鍵生成部、１４０　データ送受信部、１５０　ユーザＩＤ記憶部、２００　暗号化装置、２１０　基準遺伝子取得部、２２０　対象遺伝子入力部、２３０　公開パラメータ記憶部、２４０　差分情報生成部、２５０　差分情報エンコード部、２６０　データ暗号化部、２７０　暗号化タグ生成部、２８０　タグ付き暗号化データ生成部、３００　検索装置、３１０　差分情報入力部、３２０　ユーザ秘密鍵記憶部、３３０　差分情報エンコード部、３４０　検索クエリ生成部、３５０　遺伝子情報取得部、３６０　データ復号部、４００　データセンタ、４１０　保管要求処理部、４２０　暗号化データ記憶部、４３０　暗号化タグ

記憶部、４４０　検索要求処理部、４５０　公開パラメータ記憶部、４６０　アクセス権限記憶部、５００　ネットワーク。

請求の範囲

[請求項1] 遺伝子情報を記憶装置に記憶させる遺伝子情報記憶装置であり、

所定の遺伝子情報である基準遺伝子を取得する基準遺伝子取得部と、

前記記憶装置に記憶させる遺伝子情報である対象遺伝子を入力する遺伝子入力部と、

前記基準遺伝子取得部が取得した基準遺伝子と、前記遺伝子入力部が入力した対象遺伝子とを比較して差分情報を生成する差分生成部と、

前記対象遺伝子を暗号化して暗号化遺伝子を生成するデータ暗号化部と、

前記差分生成部が生成した差分情報を埋め込んだ暗号化タグを生成する暗号化タグ生成部と、

前記データ暗号化部が生成した暗号化遺伝子と、前記暗号化タグ生成部が生成した暗号化タグとを関連付けて、前記記憶装置に記憶させるデータ記憶部と

を備えることを特徴とする遺伝子情報記憶装置。

[請求項2] 前記差分情報には、複数の種別の情報が含まれており、

前記遺伝子情報記憶装置は、さらに、

前記差分情報に含まれる所定の種別については、その種別が取り得る値を複数のブロックに分け、前記差分生成部が生成した差分情報におけるその種別の値を、その値が属するブロックを識別する識別情報に置き換える差分情報置換部

を備え、

前記暗号化タグ生成部は、前記差分情報置換部が置き換えた差分情報を暗号化して暗号化タグを生成する

ことを特徴とする請求項１に記載の遺伝子情報記憶装置。

[請求項3] 前記差分情報置換部は、各ブロックに属する値の一部が他のブロッ

クにも属するように、前記種別が取り得る値を複数のブロックに分け、前記差分生成部が生成した差分情報におけるその種別の値を、その値が属する各ブロックを識別する識別情報に置き換える

ことを特徴とする請求項２に記載の遺伝子情報記憶装置。

[請求項4] 前記暗号化タグ生成部は、暗号化データに設定された暗号属性と、秘密鍵に設定された鍵属性とが対応していない場合に、前記暗号化データを前記秘密鍵で復号できない暗号化方式に基づき、前記暗号化遺伝子を検索可能なユーザの属性情報と差分情報とを前記暗号属性として設定して、乱数値を暗号化して前記暗号化タグを生成する

ことを特徴とする請求項１から３までのいずれかに記載の遺伝子情報記憶装置。

[請求項5] 前記暗号化方式に基づき、前記暗号化遺伝子を復号可能なユーザの属性情報を前記暗号属性として設定して、前記対象遺伝子を暗号化して前記暗号化遺伝子を生成する

ことを特徴とする請求項４に記載の遺伝子情報記憶装置。

[請求項6] データ管理装置が管理する記憶装置に記憶された遺伝子情報を検索する遺伝子情報検索装置であり、

検索したい遺伝子情報と所定の遺伝子情報である基準遺伝子との差分情報を入力する差分情報入力部と、

前記差分情報入力部が入力した差分情報を埋め込んだ検索クエリを生成する検索クエリ生成部と、

前記検索クエリ生成部が生成した検索クエリを前記データ管理装置へ送信して、前記差分情報を含む遺伝子情報を取得する遺伝子情報取得部と

を備えることを特徴とする遺伝子情報検索装置。

[請求項7] 前記差分情報には、複数の種別の情報が含まれており、

前記遺伝子情報検索装置は、さらに、

前記差分情報に含まれる所定の種別については、その種別が取り得

る値を複数のブロックに分け、差分情報におけるその種別の値を、その値が属するブロックを識別する識別情報に置き換える差分情報置換部

を備え、

前記検索クエリ生成部は、前記差分情報置換部が置き換えた差分情報を埋め込んだ検索クエリを生成する

ことを特徴とする請求項６に記載の遺伝子情報検索装置。

[請求項8] 前記差分情報置換部は、各ブロックに属する値の一部が他のブロックにも属するように、前記種別が取り得る値を複数のブロックに分け、差分情報におけるその種別の値を、その値が属する各ブロックを識別する識別情報に置き換える

ことを特徴とする請求項７に記載の遺伝子情報検索装置。

[請求項9] 前記遺伝子情報検索装置は、さらに、

暗号化データに設定された暗号属性と、秘密鍵に設定された鍵属性とが対応していない場合に、前記暗号化データを前記秘密鍵で復号できない暗号化方式における前記秘密鍵であって、ユーザの属性情報が前記鍵属性として設定された秘密鍵を管理するユーザ秘密鍵管理部

を備え、

前記検索クエリ生成部は、前記ユーザ秘密鍵管理部が管理する秘密鍵の鍵属性として、前記差分情報を追加して前記検索クエリを生成する

ことを特徴とする請求項６から８までのいずれかに記載の遺伝子情報検索装置。

[請求項10] 前記遺伝子情報取得部は、前記暗号化方式により、復号可能なユーザの属性情報を前記暗号属性として設定して、前記遺伝子情報を暗号化した暗号化遺伝子を、前記差分情報を含む遺伝子情報として取得し、

前記遺伝子情報検索装置は、さらに、

前記ユーザ秘密鍵管理部が管理する秘密鍵により、前記暗号化遺伝子を復号する復号部を備えることを特徴とする請求項９に記載の遺伝子情報検索装置。

[請求項11] 遺伝子情報を記憶装置に記憶させる遺伝子情報記憶プログラムであり、

所定の遺伝子情報である基準遺伝子を取得する基準遺伝子取得処理と、

前記記憶装置に記憶させる遺伝子情報である対象遺伝子を入力する遺伝子入力処理と、

前記基準遺伝子取得処理で取得した基準遺伝子と、前記遺伝子入力処理で入力した対象遺伝子とを比較して差分情報を生成する差分生成処理と、

前記対象遺伝子を暗号化して暗号化遺伝子を生成するデータ暗号化処理と、

前記差分生成処理で生成した差分情報を埋め込んだ暗号化タグを生成する暗号化タグ生成処理と、

前記データ暗号化処理で生成した暗号化遺伝子と、前記暗号化タグ生成処理で生成した暗号化タグとを関連付けて、前記記憶装置に記憶させるデータ記憶処理と

をコンピュータに実行させることを特徴とする遺伝子情報記憶プログラム。

[請求項12] 前記差分情報には、複数の種別の情報が含まれており、

前記遺伝子情報記憶プログラムは、さらに、

前記差分情報に含まれる所定の種別については、その種別が取り得る値を複数のブロックに分け、前記差分生成処理で生成した差分情報におけるその種別の値を、その値が属するブロックを識別する識別情報に置き換える差分情報置換処理

をコンピュータに実行させ、

前記暗号化タグ生成処理では、前記差分情報置換処理で置き換えた差分情報を暗号化し
て暗号化タグを生成する
ことを特徴とする請求項１１に記載の遺伝子情報記憶プログラム。

[請求項13] 前記差分情報置換処理では、各ブロックに属する値の一部が他のブロックにも属するように、前記種別が取り得る値を複数のブロックに分け、前記差分生成処理で生成した差分情報におけるその種別の値を、その値が属する各ブロックを識別する識別情報に置き換える
ことを特徴とする請求項１２に記載の遺伝子情報記憶プログラム。

[請求項14] 前記暗号化タグ生成処理では、暗号化データに設定された暗号属性と、秘密鍵に設定された鍵属性とが対応していない場合に、前記暗号化データを前記秘密鍵で復号できない暗号化方式に基づき、前記暗号化遺伝子を検索可能なユーザの属性情報と差分情報とを前記暗号属性として設定して、乱数値を暗号化して前記暗号化タグを生成する
ことを特徴とする請求項１１から１３までのいずれかに記載の遺伝子情報記憶プログラム。

[請求項15] 前記暗号化方式に基づき、前記暗号化遺伝子を復号可能なユーザの属性情報を前記暗号属性として設定して、前記対象遺伝子を暗号化して前記暗号化遺伝子を生成する
ことを特徴とする請求項１４に記載の遺伝子情報記憶プログラム。

[請求項16] データ管理装置が管理する記憶装置に記憶された遺伝子情報を検索する遺伝子情報検索プログラムであり、

検索したい遺伝子情報と所定の遺伝子情報である基準遺伝子との差分情報を入力する差分情報入力処理と、

前記差分情報入力処理で入力した差分情報を埋め込んだ検索クエリを生成する検索クエリ生成処理と、

前記検索クエリ生成処理で生成した検索クエリを前記データ管理装置へ送信して、前記差分情報を含む遺伝子情報を取得する遺伝子情報

取得処理と

をコンピュータに実行させることを特徴とする遺伝子情報検索プログラム。

[請求項17] 前記差分情報には、複数の種別の情報が含まれており、

前記遺伝子情報検索プログラムは、さらに、

前記差分情報に含まれる所定の種別については、その種別が取り得る値を複数のブロックに分け、差分情報におけるその種別の値を、その値が属するブロックを識別する識別情報に置き換える差分情報置換処理

をコンピュータに実行させ、

前記検索クエリ生成処理では、前記差分情報置換処理で置き換えた差分情報を埋め込んだ検索クエリを生成する

ことを特徴とする請求項１６に記載の遺伝子情報検索プログラム。

[請求項18] 前記差分情報置換処理では、各ブロックに属する値の一部が他のブロックにも属するように、前記種別が取り得る値を複数のブロックに分け、差分情報におけるその種別の値を、その値が属する各ブロックを識別する識別情報に置き換える

ことを特徴とする請求項１７に記載の遺伝子情報検索プログラム。

[請求項19] 前記遺伝子情報検索プログラムは、さらに、

暗号化データに設定された暗号属性と、秘密鍵に設定された鍵属性とが対応していない場合に、前記暗号化データを前記秘密鍵で復号できない暗号化方式における前記秘密鍵であって、ユーザの属性情報が前記鍵属性として設定された秘密鍵を管理するユーザ秘密鍵管理処理をコンピュータに実行させ、

前記検索クエリ生成処理では、前記ユーザ秘密鍵管理処理で管理する秘密鍵の鍵属性として、前記差分情報を追加して前記検索クエリを生成する

ことを特徴とする請求項１６から１８までのいずれかに記載の遺伝子

情報検索プログラム。

[請求項20] 前記遺伝子情報取得処理では、前記暗号化方式により、復号可能なユーザの属性情報を前記暗号属性として設定して、前記遺伝子情報を暗号化した暗号化遺伝子を、前記差分情報を含む遺伝子情報として取得し、

前記遺伝子情報検索プログラムは、さらに、

前記ユーザ秘密鍵管理処理で管理する秘密鍵により、前記暗号化遺伝子を復号する復号処理

をコンピュータに実行させることを特徴とする請求項１９に記載の遺伝子情報検索プログラム。

[請求項21] 遺伝子情報を記憶装置に記憶させる遺伝子情報記憶方法であり、

処理装置が、所定の遺伝子情報である基準遺伝子を取得する基準遺伝子取得工程と、

入力装置が、前記記憶装置に記憶させる遺伝子情報である対象遺伝子を入力する遺伝子入力工程と、

処理装置が、前記基準遺伝子取得工程で取得した基準遺伝子と、前記遺伝子入力工程で入力した対象遺伝子とを比較して差分情報を生成する差分生成工程と、

処理装置が、前記対象遺伝子を暗号化して暗号化遺伝子を生成するデータ暗号化工程と、

処理装置が、前記差分生成工程で生成した差分情報を埋め込んだ暗号化タグを生成する暗号化タグ生成工程と、

処理装置が、前記データ暗号化工程で生成した暗号化遺伝子と、前記暗号化タグ生成工程で生成した暗号化タグとを関連付けて、前記記憶装置に記憶させるデータ記憶工程と

を備えることを特徴とする遺伝子情報記憶方法。

[請求項22] データ管理装置が管理する記憶装置に記憶された遺伝子情報を検索する遺伝子情報検索方法であり、

入力装置が、検索したい遺伝子情報と所定の遺伝子情報である基準遺伝子との差分情報を入力する差分情報入力工程と、

処理装置が、前記差分情報入力工程で入力した差分情報を埋め込んだ検索クエリを生成する検索クエリ生成工程と、

処理装置が、前記検索クエリ生成工程で生成した検索クエリを前記データ管理装置へ送信して、前記差分情報を含む遺伝子情報を取得する遺伝子情報取得工程と

を備えることを特徴とする遺伝子情報検索方法。

[請求項23] データ管理装置が管理する記憶装置に遺伝子情報を記憶させる遺伝子情報記憶装置と、前記遺伝子情報記憶装置が記憶させた遺伝子情報から検索キーワードを含む遺伝子情報を検索する遺伝子情報検索装置とを備える遺伝子情報検索システムであり、

前記遺伝子情報記憶装置は、

所定の遺伝子情報である基準遺伝子を取得する基準遺伝子取得部と、

前記記憶装置に記憶させる遺伝子情報である対象遺伝子を入力する遺伝子入力部と、

前記基準遺伝子取得部が取得した基準遺伝子と、前記遺伝子入力部が入力した対象遺伝子とを比較して差分情報を生成する差分生成部と、

前記対象遺伝子を暗号化して暗号化遺伝子を生成するデータ暗号化部と、

前記差分生成部が生成した差分情報を埋め込んだ暗号化タグを生成する暗号化タグ生成部と、

前記データ暗号化部が生成した暗号化遺伝子と、前記暗号化タグ生成部が生成した暗号化タグとを関連付けて、前記記憶装置に記憶させるデータ記憶部と

を備え、

前記遺伝子情報検索装置は、

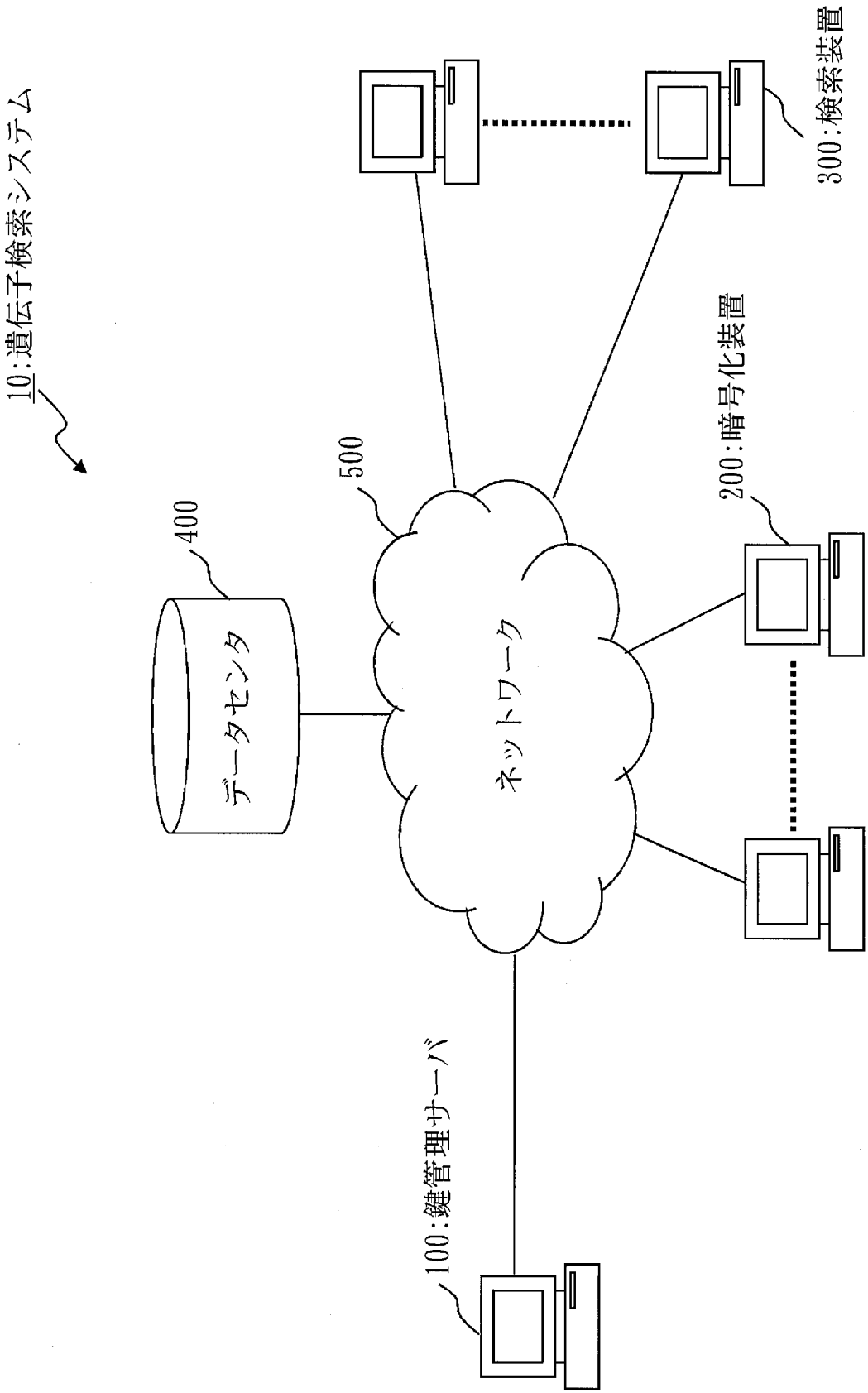
検索したい遺伝子情報と前記基準遺伝子との差分情報を前記検索キーワードとして入力する差分情報入力部と、

前記差分情報入力部が入力した差分情報を埋め込んだ検索クエリを生成する検索クエリ生成部と、

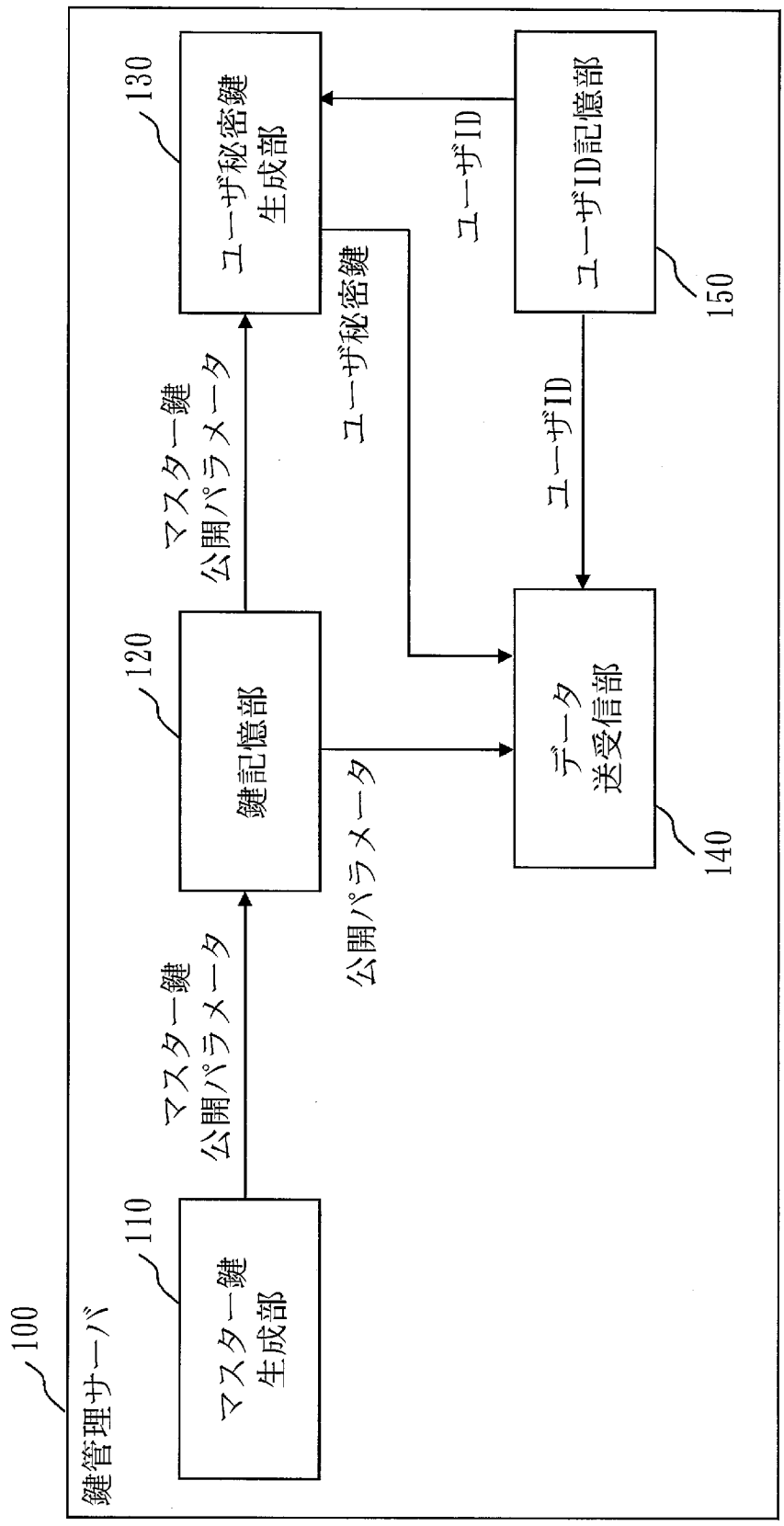
前記検索クエリ生成部が生成した検索クエリを前記データ管理装置へ送信して、前記差分情報を含む遺伝子情報を取得する遺伝子情報取得部と

備えることを特徴とする遺伝子情報検索システム。

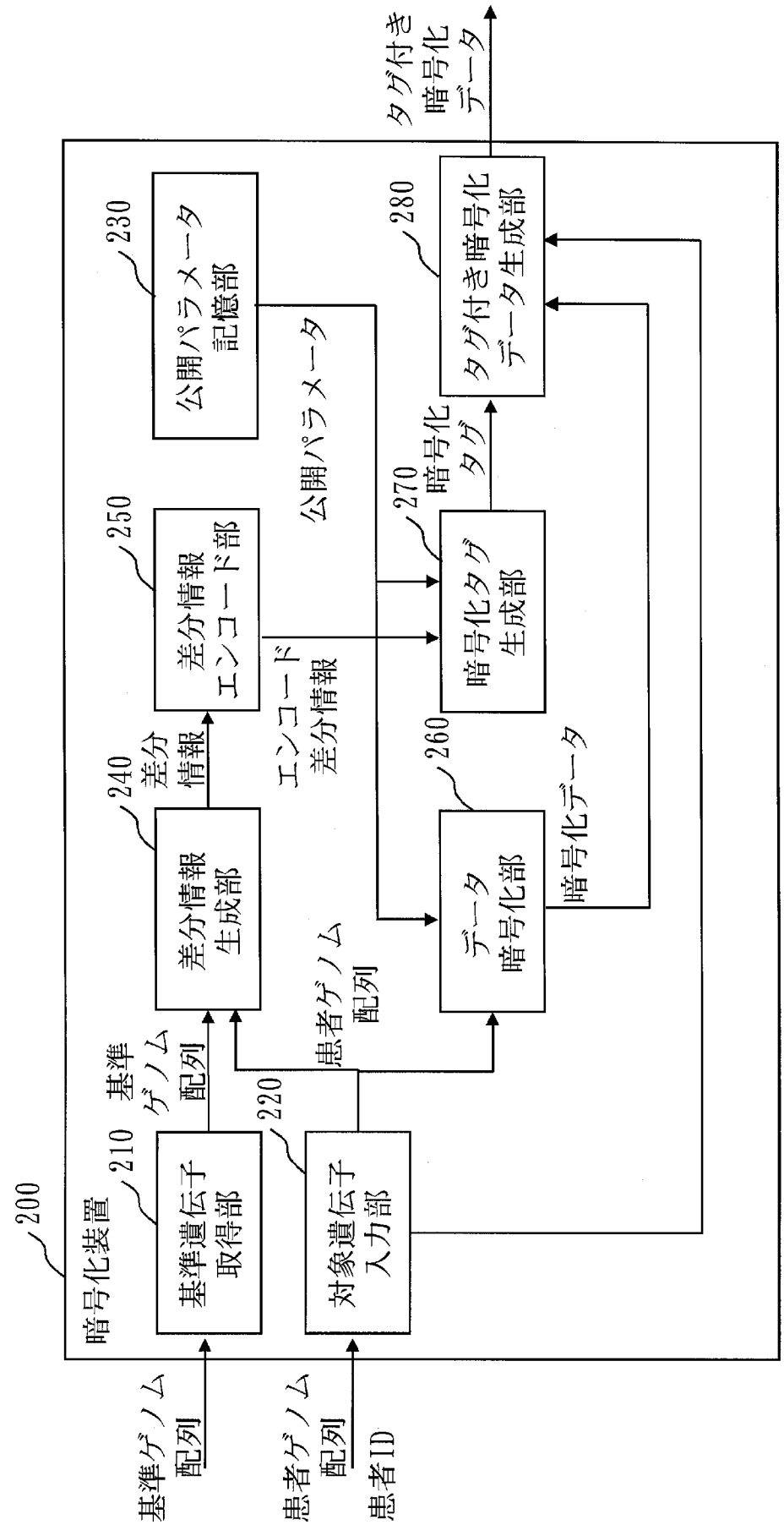
[図1]


10:遺伝子検索システム
300:検索装置
500
400
200:暗号化装置
データセンタ
ネットワーク
100:鍵管理サーバ

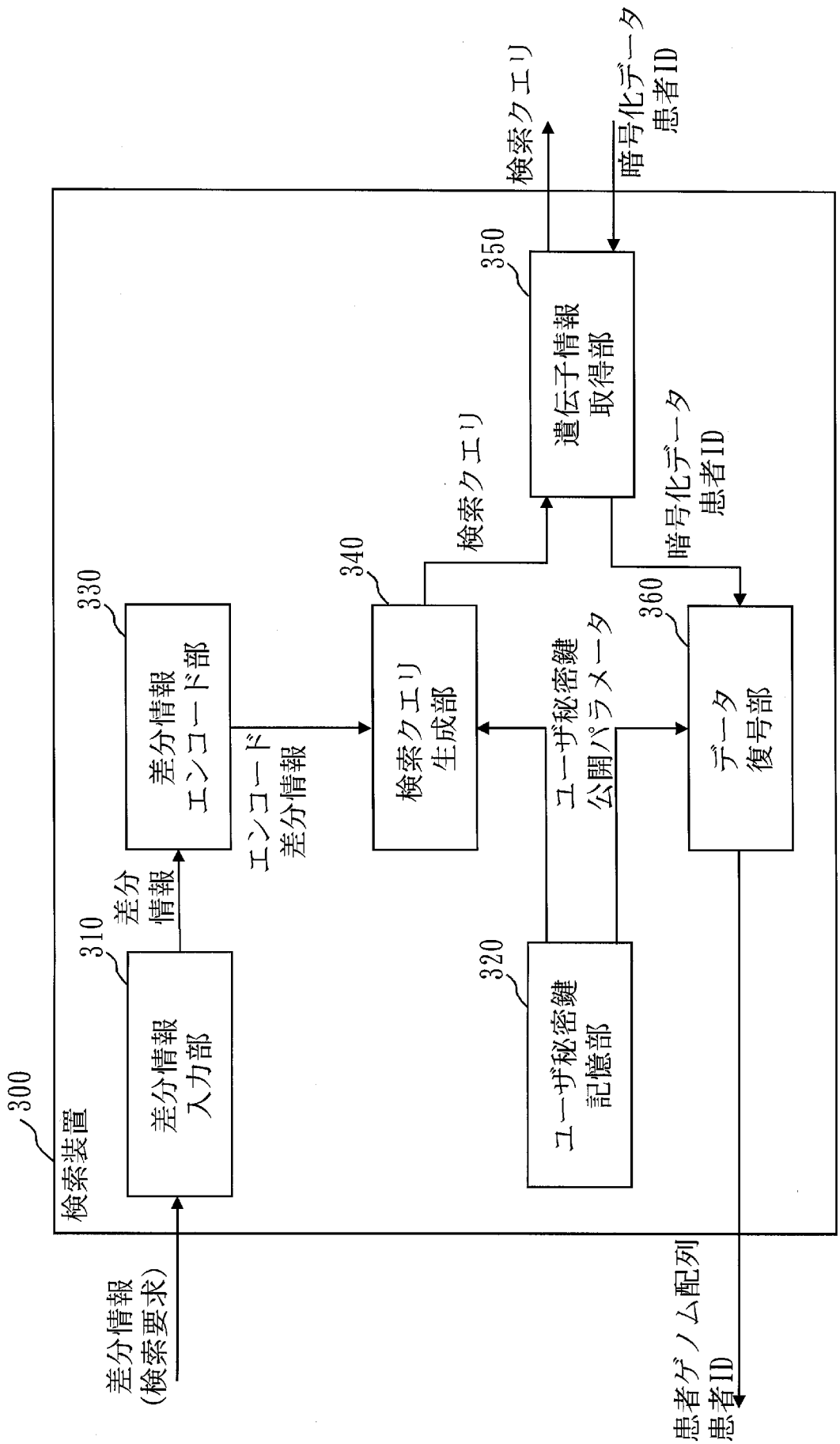
[図2]


鍵管理サーバ
100
マスター鍵
生成部
110
鍵記憶部
120
ユーザ秘密鍵
生成部
130
データ
送受信部
140
ユーザID記憶部
150
マスター鍵
公開パラメータ
マスター鍵
公開パラメータ
公開パラメータ
ユーザ秘密鍵
ユーザID
ユーザID

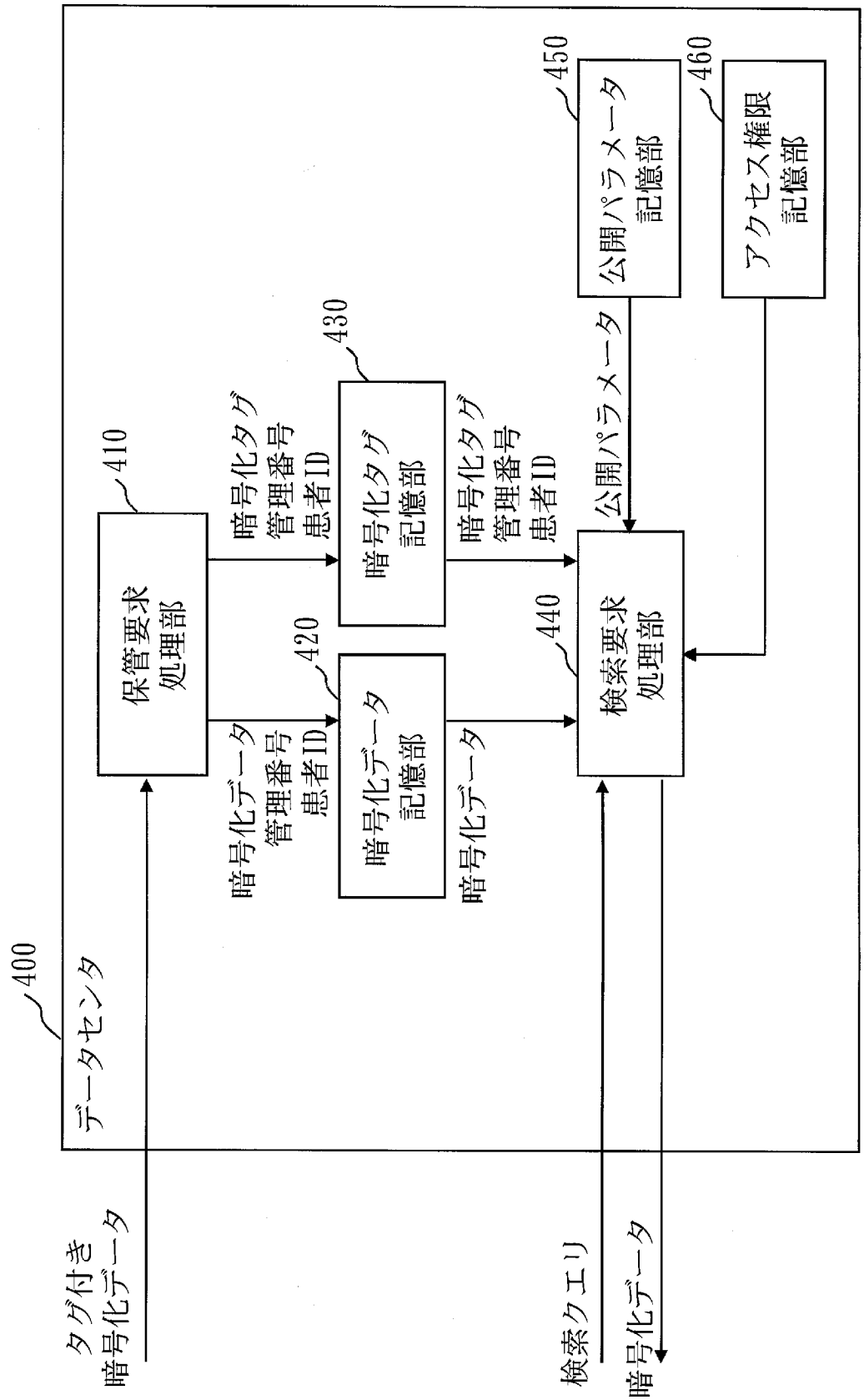
[図3]


200
暗号化装置
210
基準遺伝子
取得部
220
対象遺伝子
入力部
230
公開パラメータ
記憶部
240
差分情報
生成部
250
差分情報
エンコード部
260
データ
暗号化部
270
暗号化タグ
生成部
280
タグ付き暗号化
データ生成部
基準ゲノム
配列
患者ゲノム
配列
患者ID
基準
ゲノム
配列
患者ゲノム
配列
差分
情報
エンコード
差分情報
公開パラメータ
暗号化
タグ
暗号化データ
タグ付き
暗号化
データ

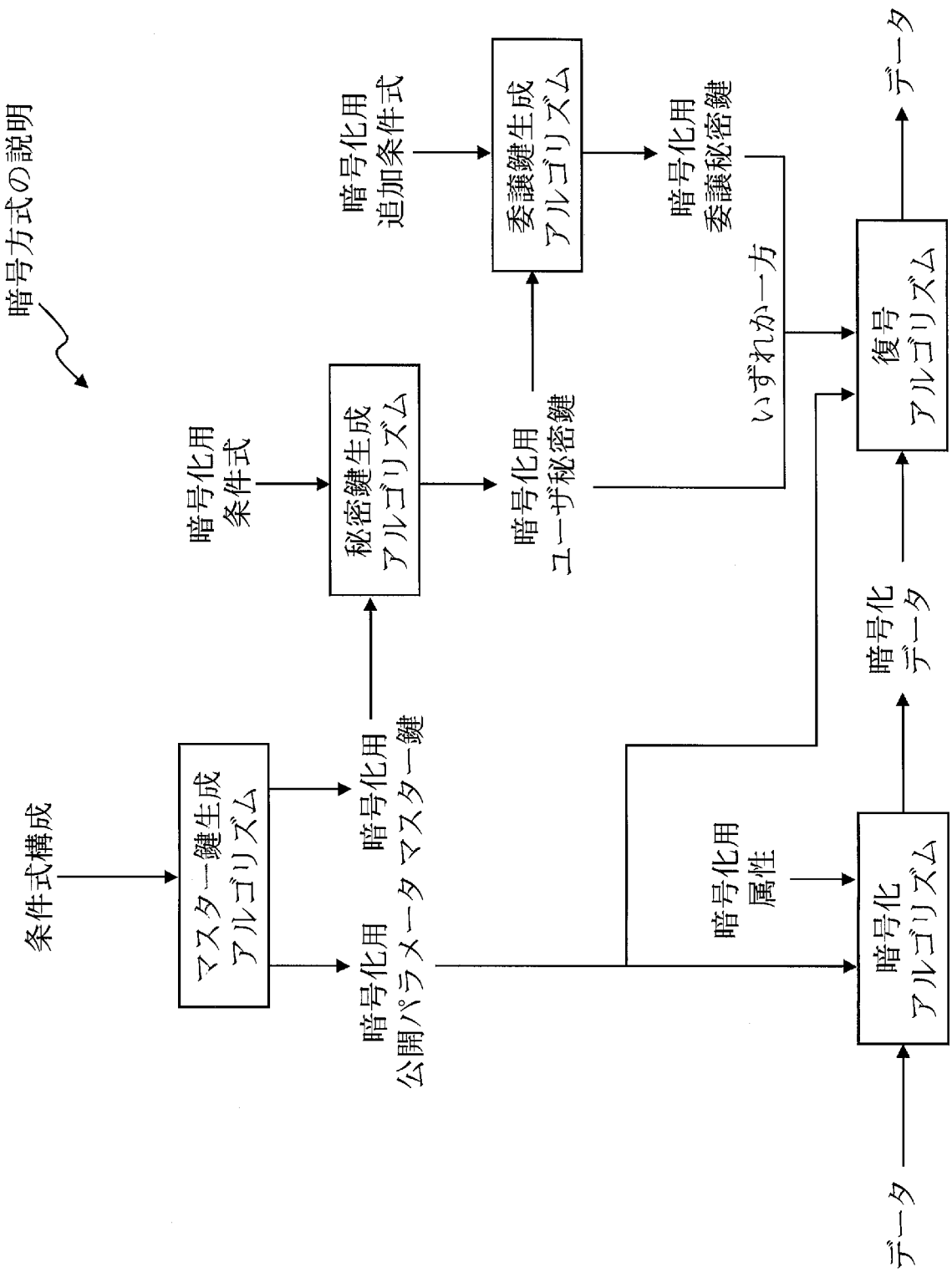
[図4]


検索装置
300
差分情報入力部
310
差分情報
差分情報エンコード部
330
エンコード差分情報
検索クエリ生成部
340
検索クエリ
遺伝子情報取得部
350
検索クエリ
暗号化データ患者ID
ユーザ秘密鍵記憶部
320
ユーザ秘密鍵公開パラメータ
データ復号部
360
暗号化データ患者ID
差分情報(検索要求)
患者ゲノム配列患者ID

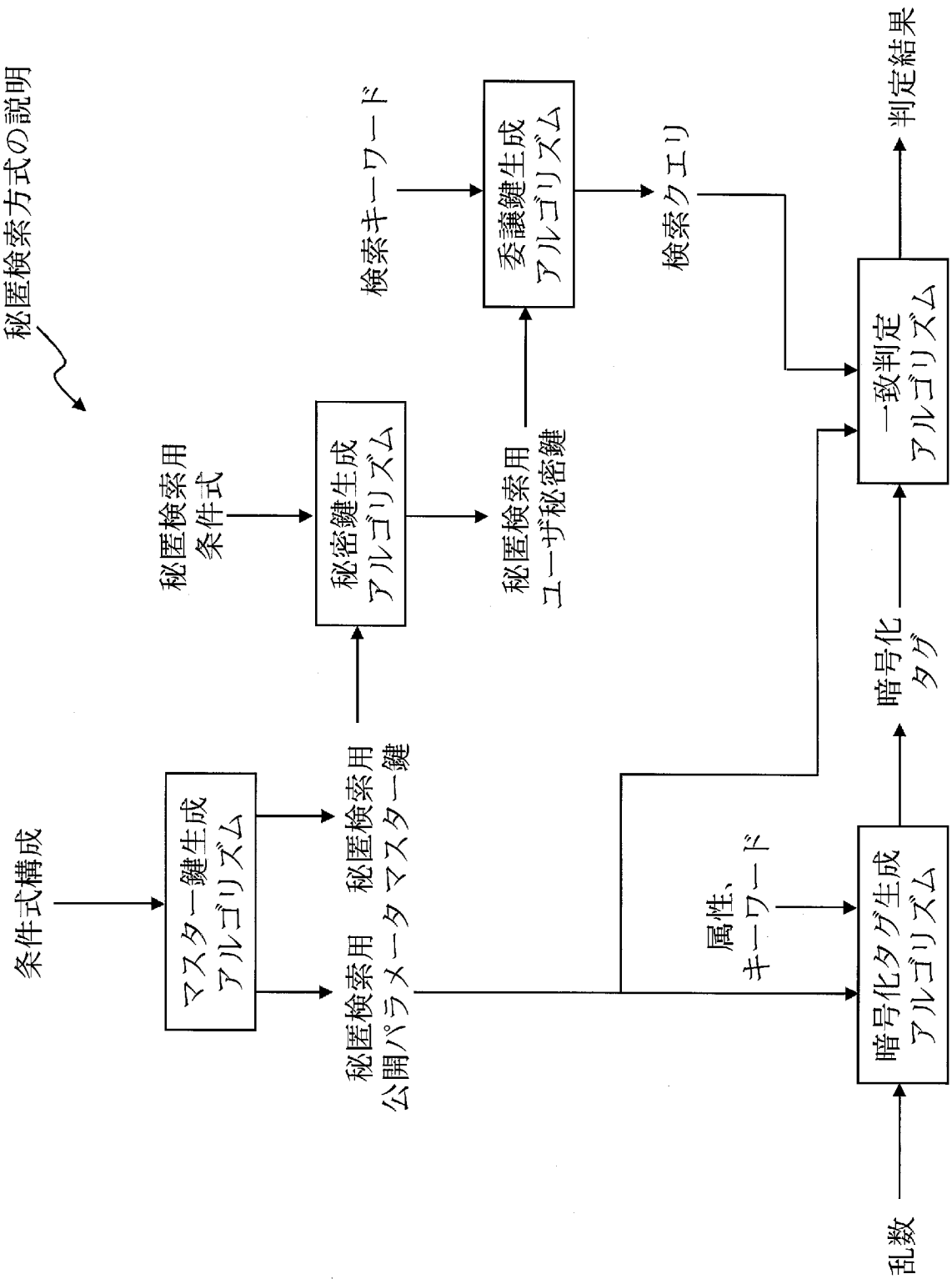
[図5]


400
データセンタ
タグ付き
暗号化データ
410
保管要求
処理部
暗号化データ
管理番号
患者ID
420
暗号化データ
記憶部
暗号化タグ
管理番号
患者ID
430
暗号化タグ
記憶部
暗号化タグ
管理番号
患者ID
440
検索要求
処理部
公開パラメータ
450
公開パラメータ
記憶部
460
アクセス権限
記憶部
検索クエリ
暗号化データ

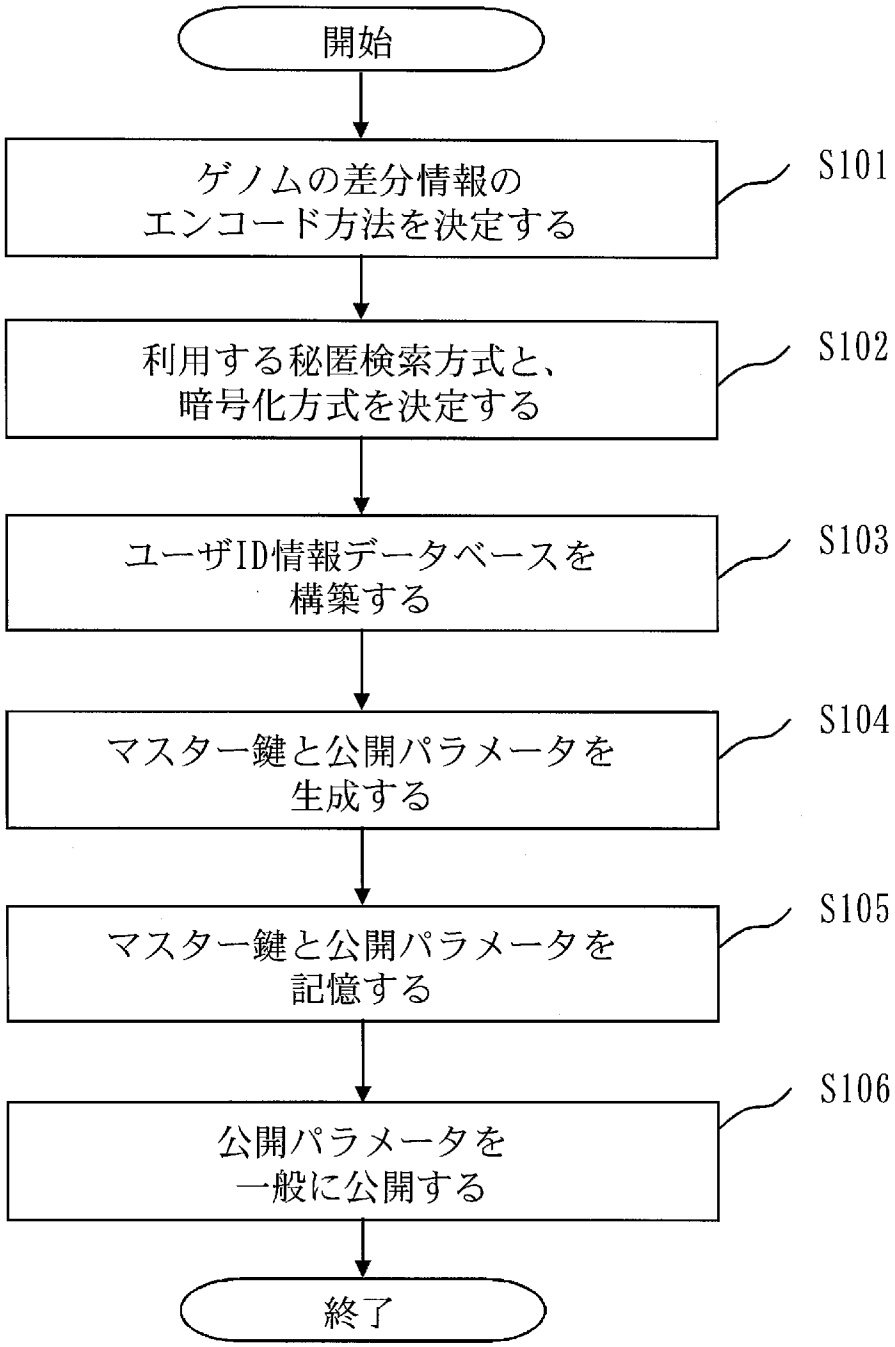
[図6]


暗号方式の説明
条件式構成
マスター鍵生成アルゴリズム
暗号化用公開パラメータ
暗号化用マスター鍵
暗号化用条件式
秘密鍵生成アルゴリズム
暗号化用ユーザ秘密鍵
暗号化用追加条件式
委譲鍵生成アルゴリズム
暗号化用委譲秘密鍵
いずれか一方
データ
暗号化用属性
暗号化アルゴリズム
暗号化データ
復号アルゴリズム
データ

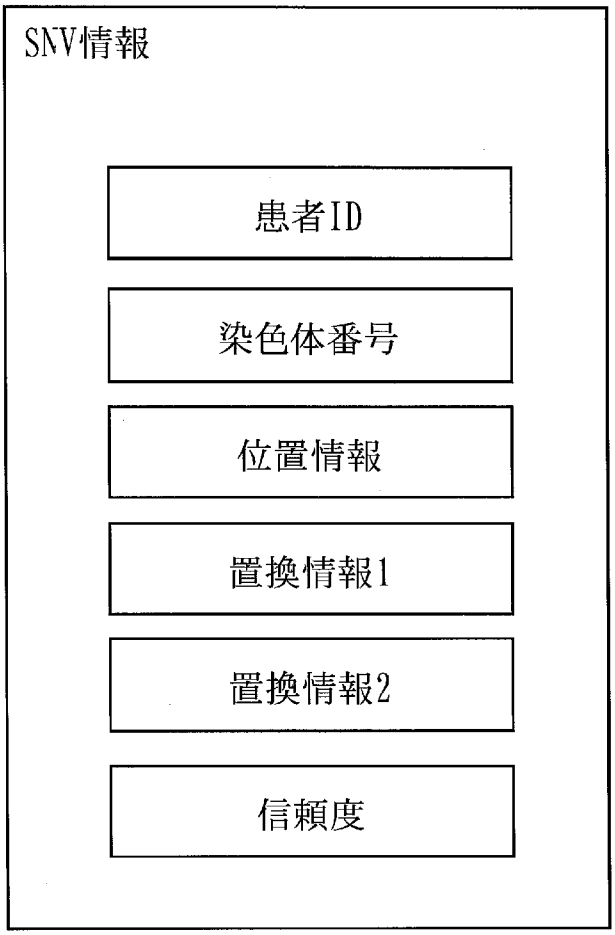
[図7]


秘匿検索方式の説明
条件式構成
マスター鍵生成アルゴリズム
秘匿検索用公開パラメータ
秘匿検索用マスター鍵
秘匿検索用条件式
秘密鍵生成アルゴリズム
秘匿検索用ユーザ秘密鍵
検索キーワード
委譲鍵生成アルゴリズム
検索クエリ
属性、キーワード
暗号化タグ生成アルゴリズム
暗号化タグ
一致判定アルゴリズム
判定結果
乱数

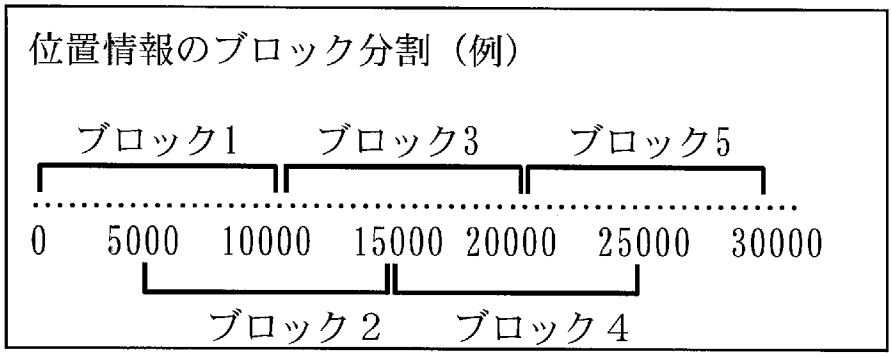
[図8]


開始
ゲノムの差分情報の
エンコード方法を決定する
S101
利用する秘匿検索方式と、
暗号化方式を決定する
S102
ユーザID情報データベースを
構築する
S103
マスター鍵と公開パラメータを
生成する
S104
マスター鍵と公開パラメータを
記憶する
S105
公開パラメータを
一般に公開する
S106
終了

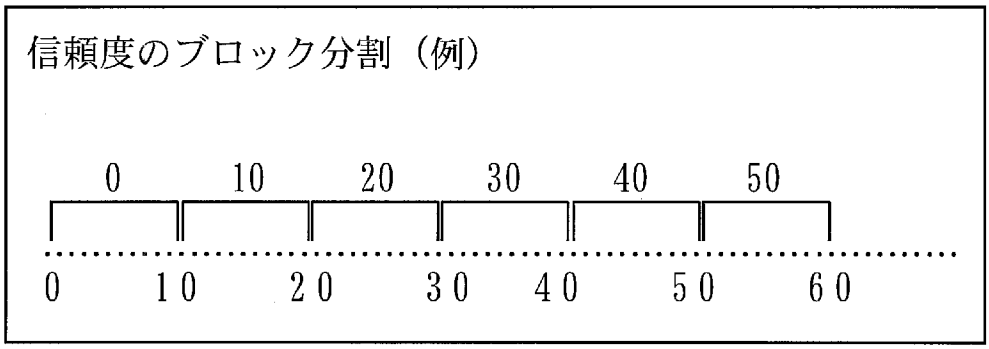
[図9]


SNV情報
患者ID
染色体番号
位置情報
置換情報1
置換情報2
信頼度

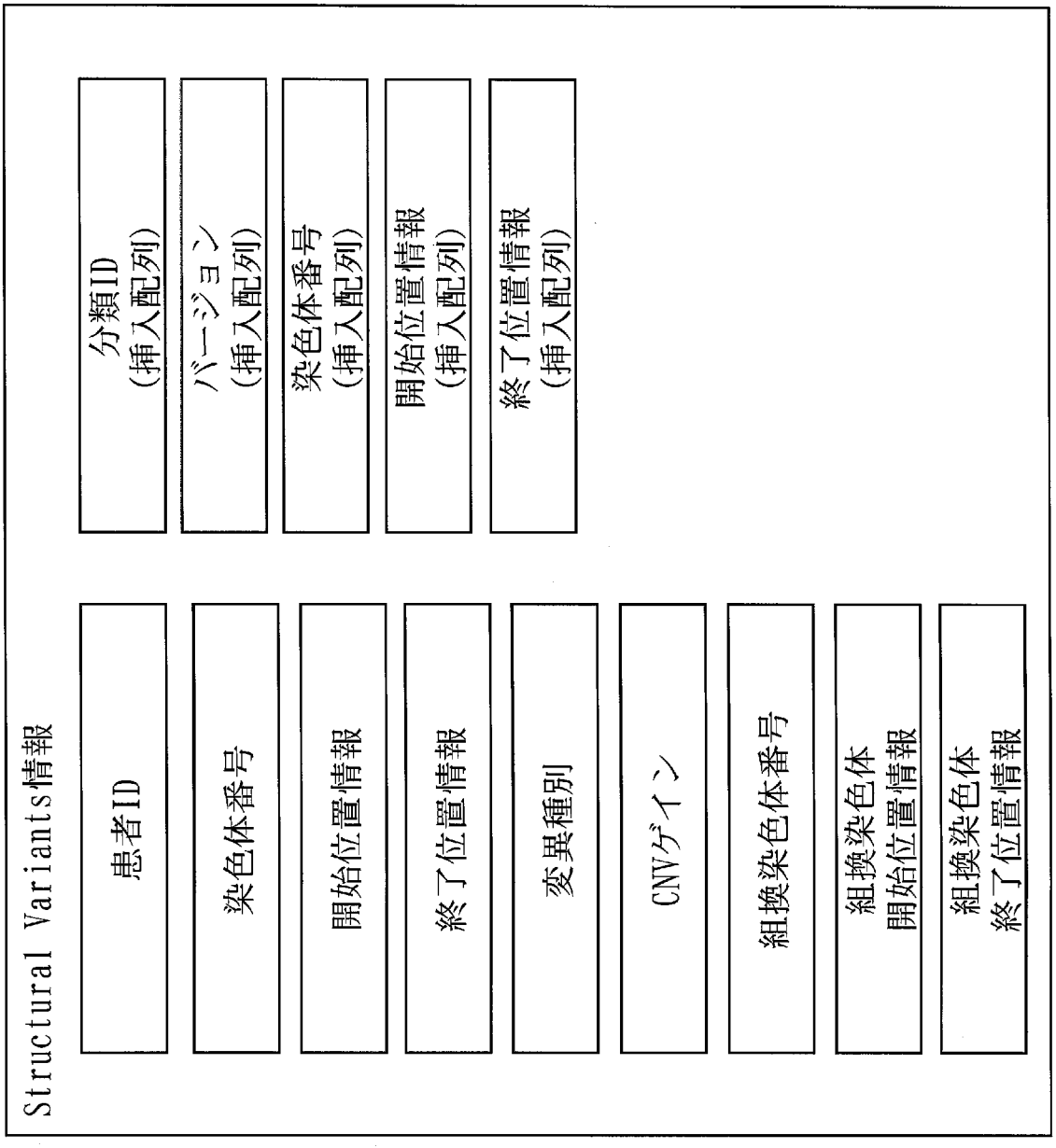
[図10]


位置情報のブロック分割（例）
ブロック1
ブロック3
ブロック5
0
5000
10000
15000
20000
25000
30000
ブロック２
ブロック４

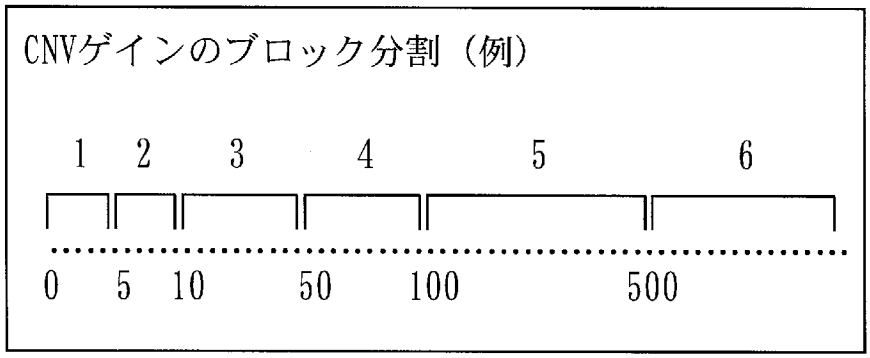
[図11]


信頼度のブロック分割（例）
0
10
20
30
40
50
0
10
20
30
40
50
60

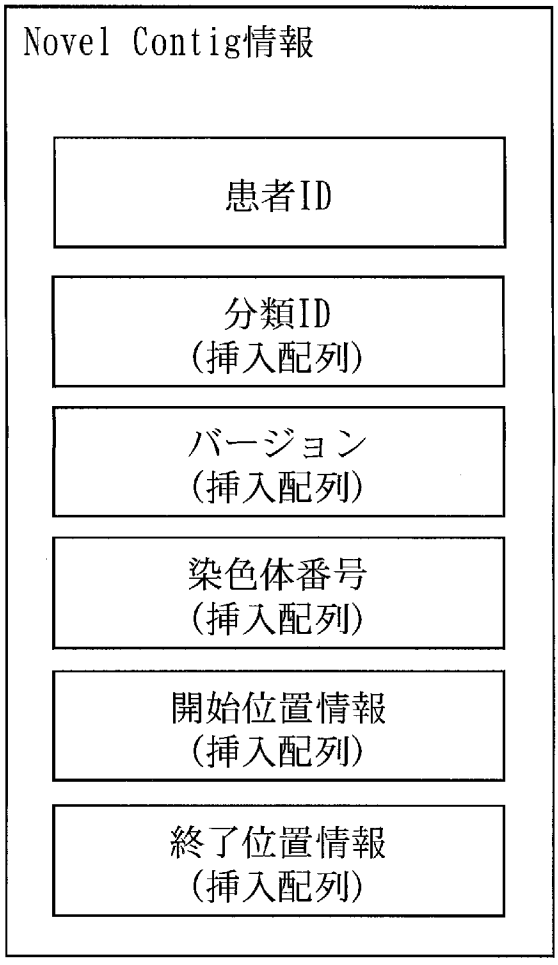
[図12]


Structural Variants情報
患者ID
染色体番号
開始位置情報
終了位置情報
変異種別
CNVゲイン
組換染色体番号
組換染色体
開始位置情報
組換染色体
終了位置情報
分類ID
（挿入配列）
バージョン
（挿入配列）
染色体番号
（挿入配列）
開始位置情報
（挿入配列）
終了位置情報
（挿入配列）

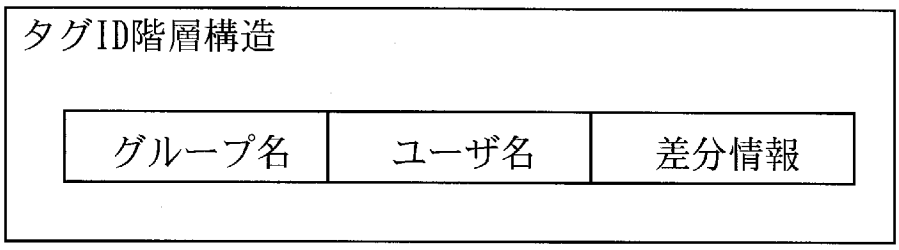
[図13]


CNVゲインのブロック分割（例）
1
2
3
4
5
6
0
5
10
50
100
500

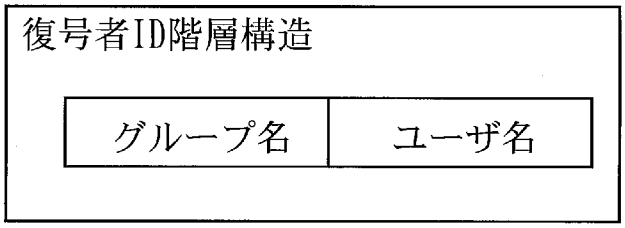
[図14]


Novel Contig情報
患者ID
分類ID
(挿入配列)
バージョン
(挿入配列)
染色体番号
(挿入配列)
開始位置情報
(挿入配列)
終了位置情報
(挿入配列)

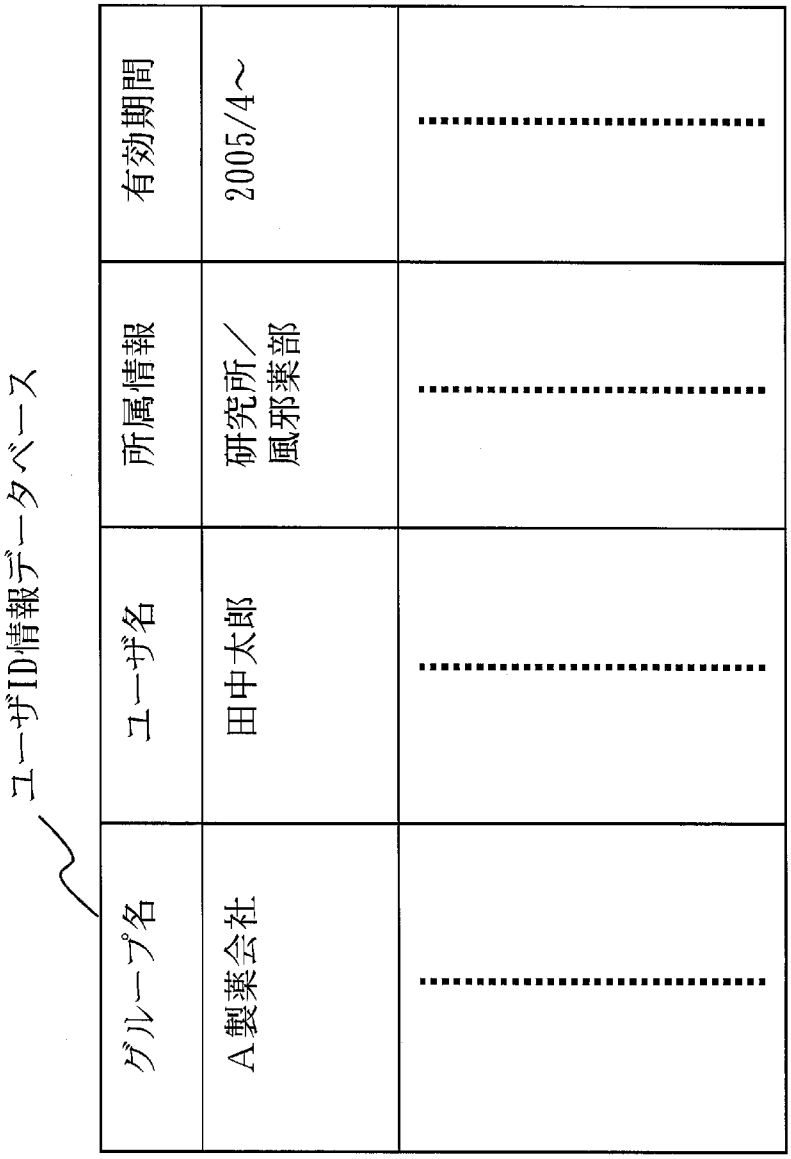
[図15]


タグID階層構造
グループ名
ユーザ名
差分情報

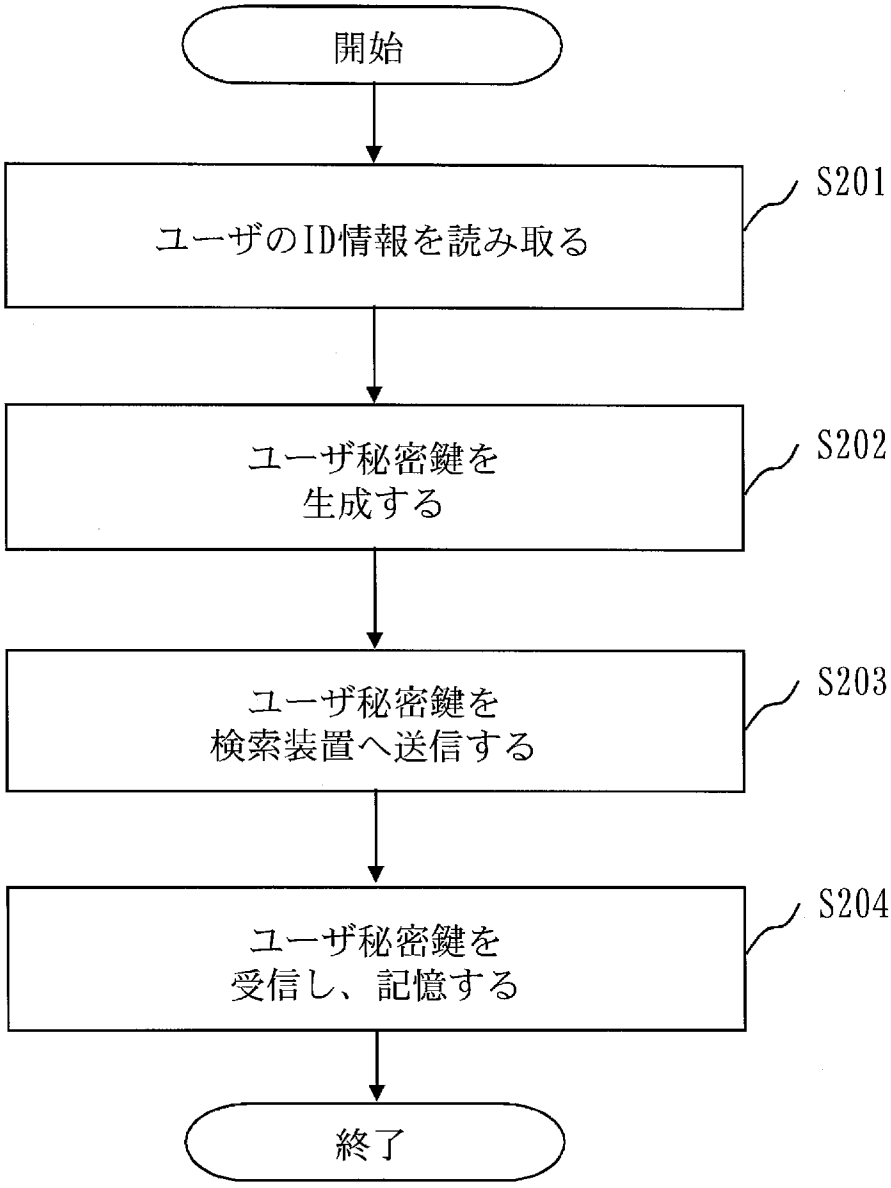
[図16]


復号者ID階層構造
グループ名
ユーザ名

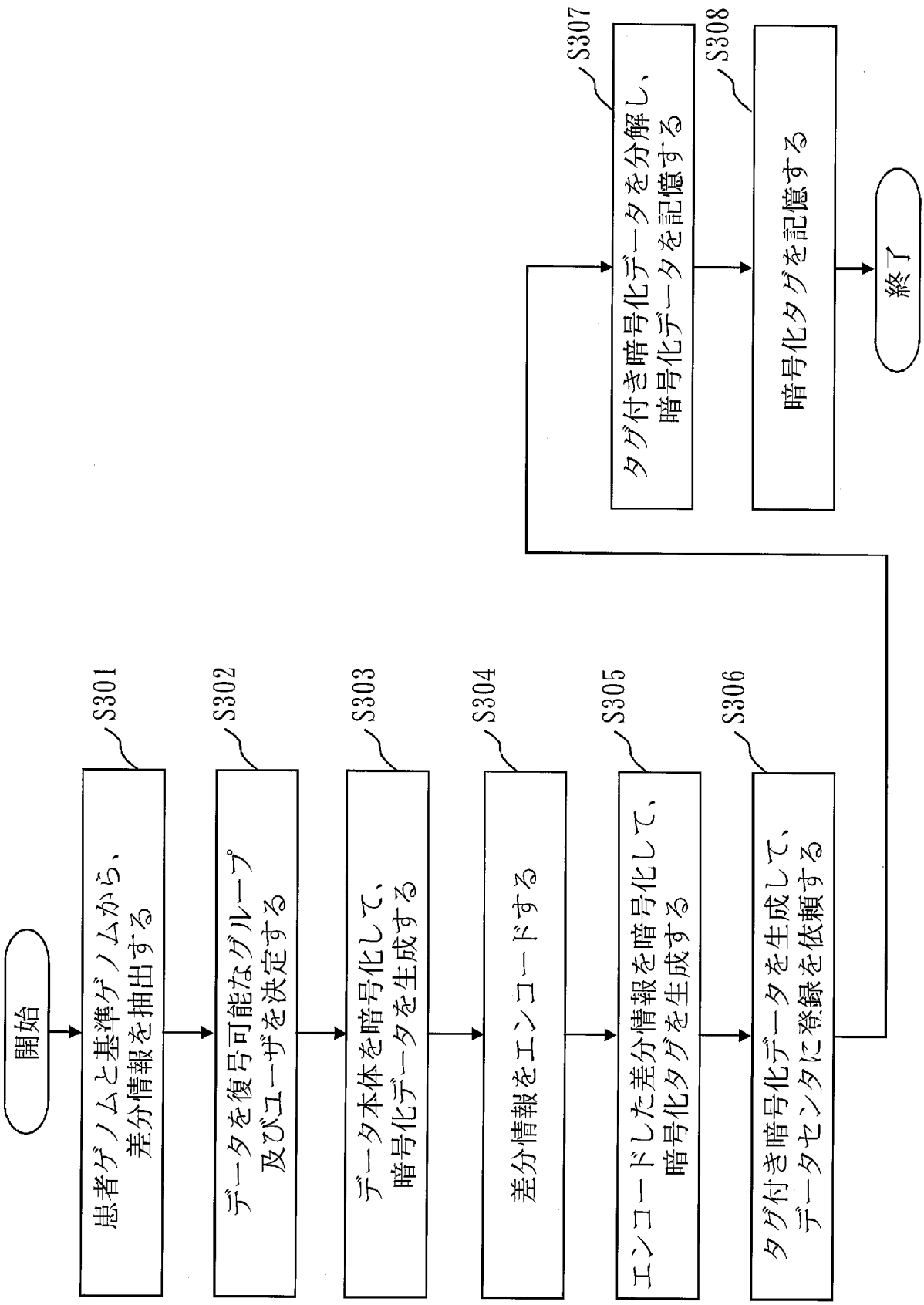
[図17]


ユーザID情報データベース
グループ名
ユーザ名
所属情報
有効期間
A製薬会社
田中太郎
研究所／
風邪薬部
2005/4～

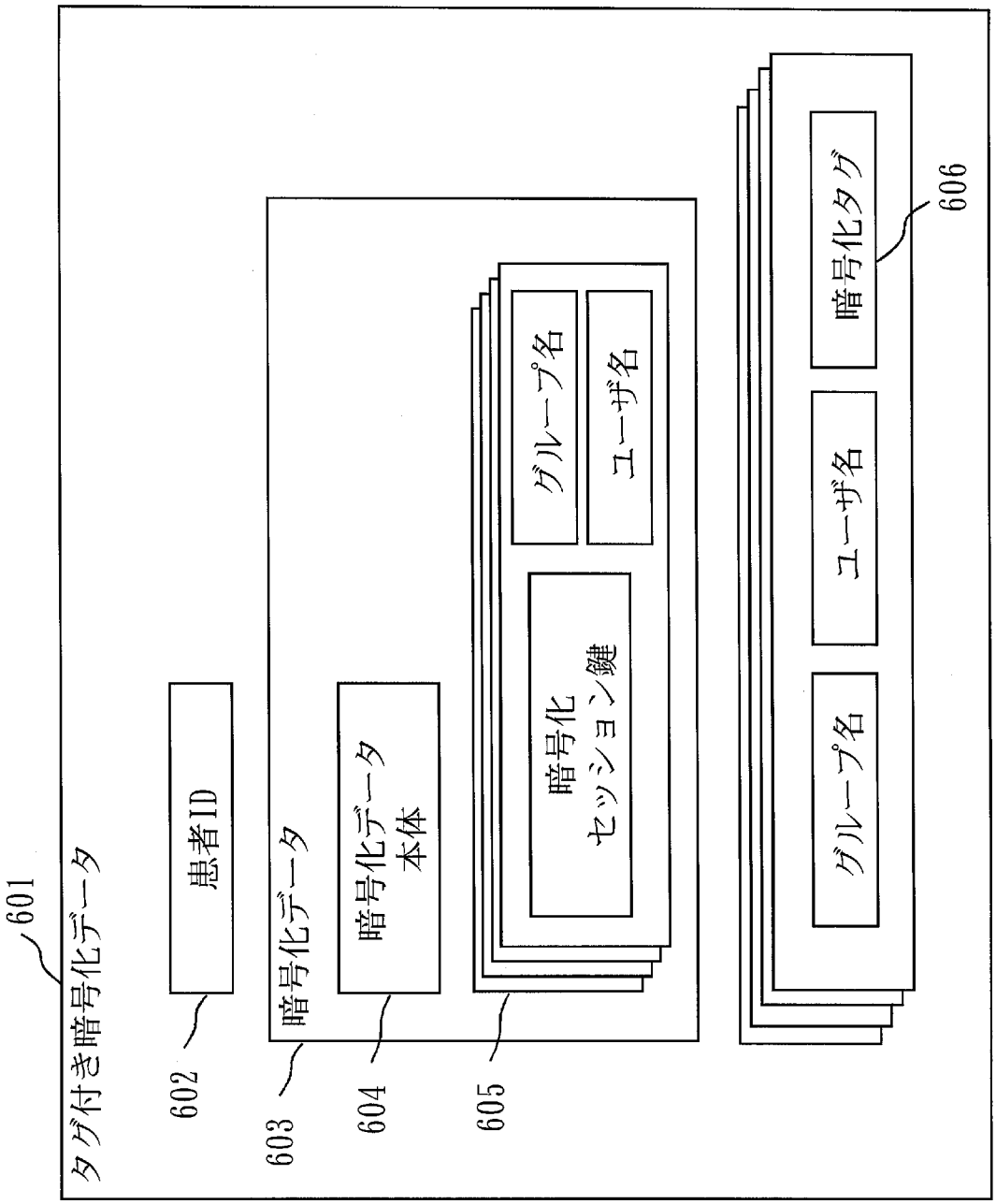
[図18]


開始
ユーザのID情報を読み取る
S201
ユーザ秘密鍵を
生成する
S202
ユーザ秘密鍵を
検索装置へ送信する
S203
ユーザ秘密鍵を
受信し、記憶する
S204
終了

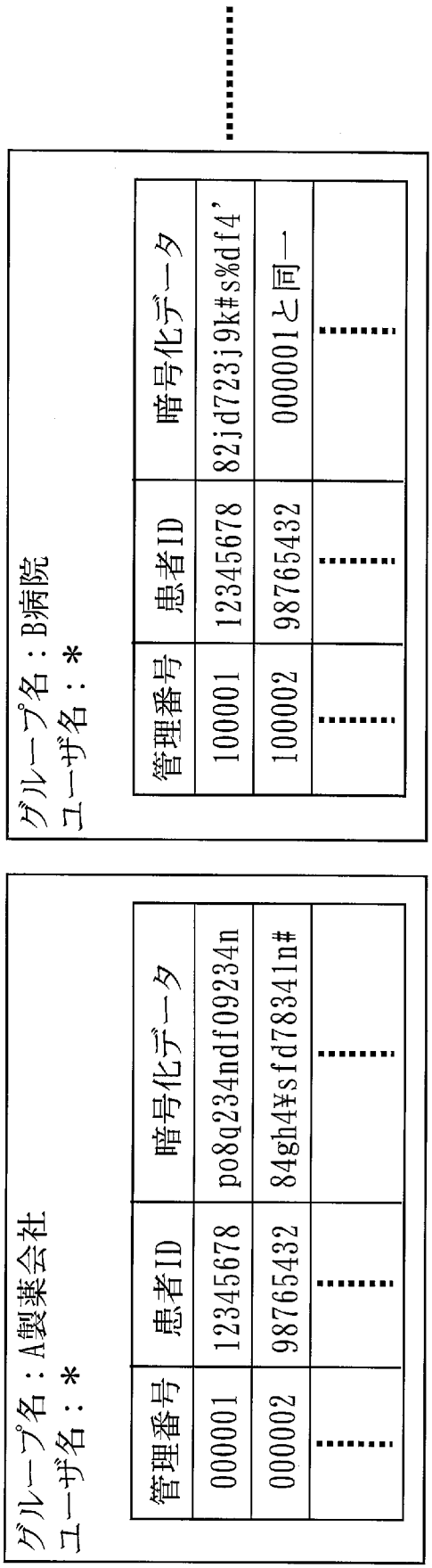
[図19]


開始
S301
患者ゲノムと基準ゲノムから、差分情報を抽出する
S302
データを復号可能なグループ及びユーザを決定する
S303
データ本体を暗号化して、暗号化データを生成する
S304
差分情報をエンコードする
S305
エンコードした差分情報を暗号化して、暗号化タグを生成する
S306
タグ付き暗号化データを生成して、データセンタに登録を依頼する
S307
タグ付き暗号化データを分解し、暗号化データを記憶する
S308
暗号化タグを記憶する
終了

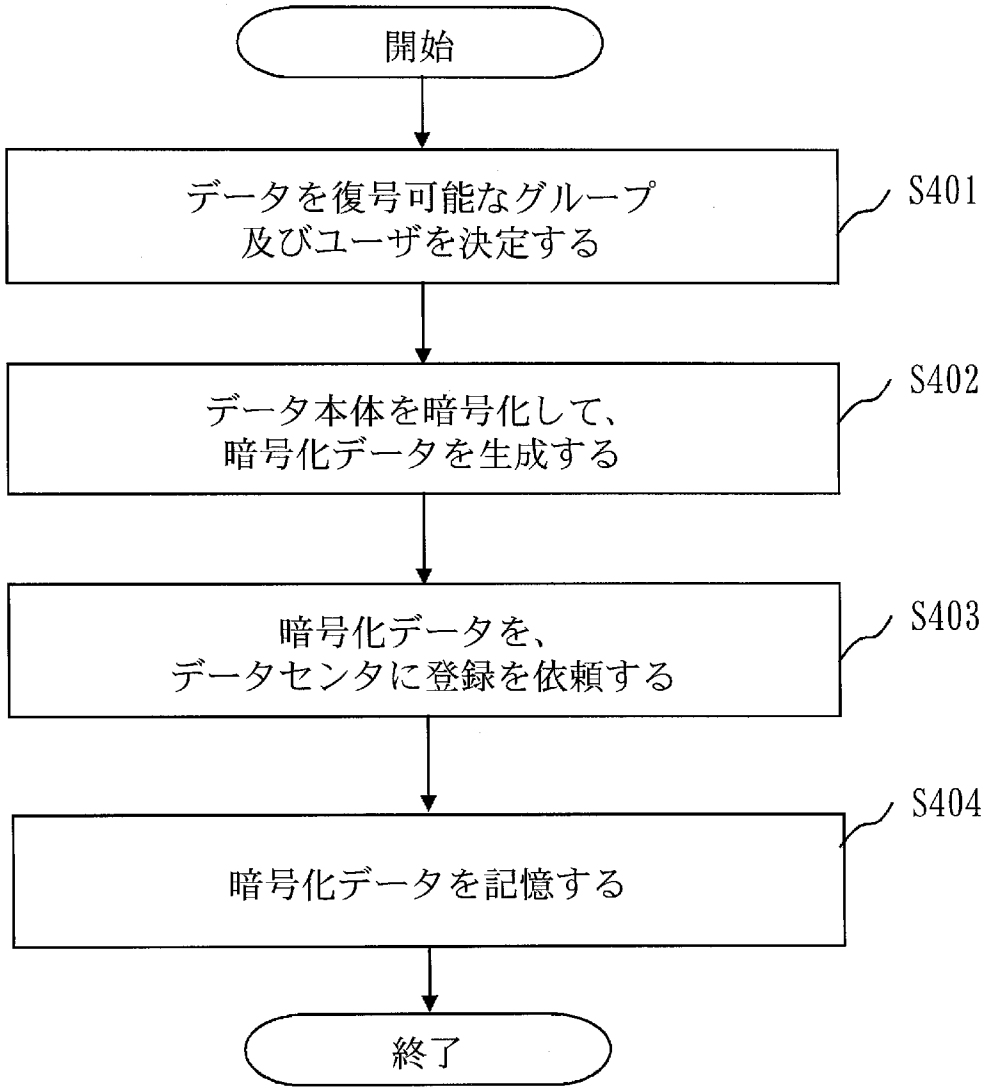
[図20]


601
タグ付き暗号化データ
患者ID
602
暗号化データ
603
暗号化データ
本体
604
暗号化
セッションキー
605
グループ名
ユーザ名
606
暗号化タグ
ユーザ名
グループ名

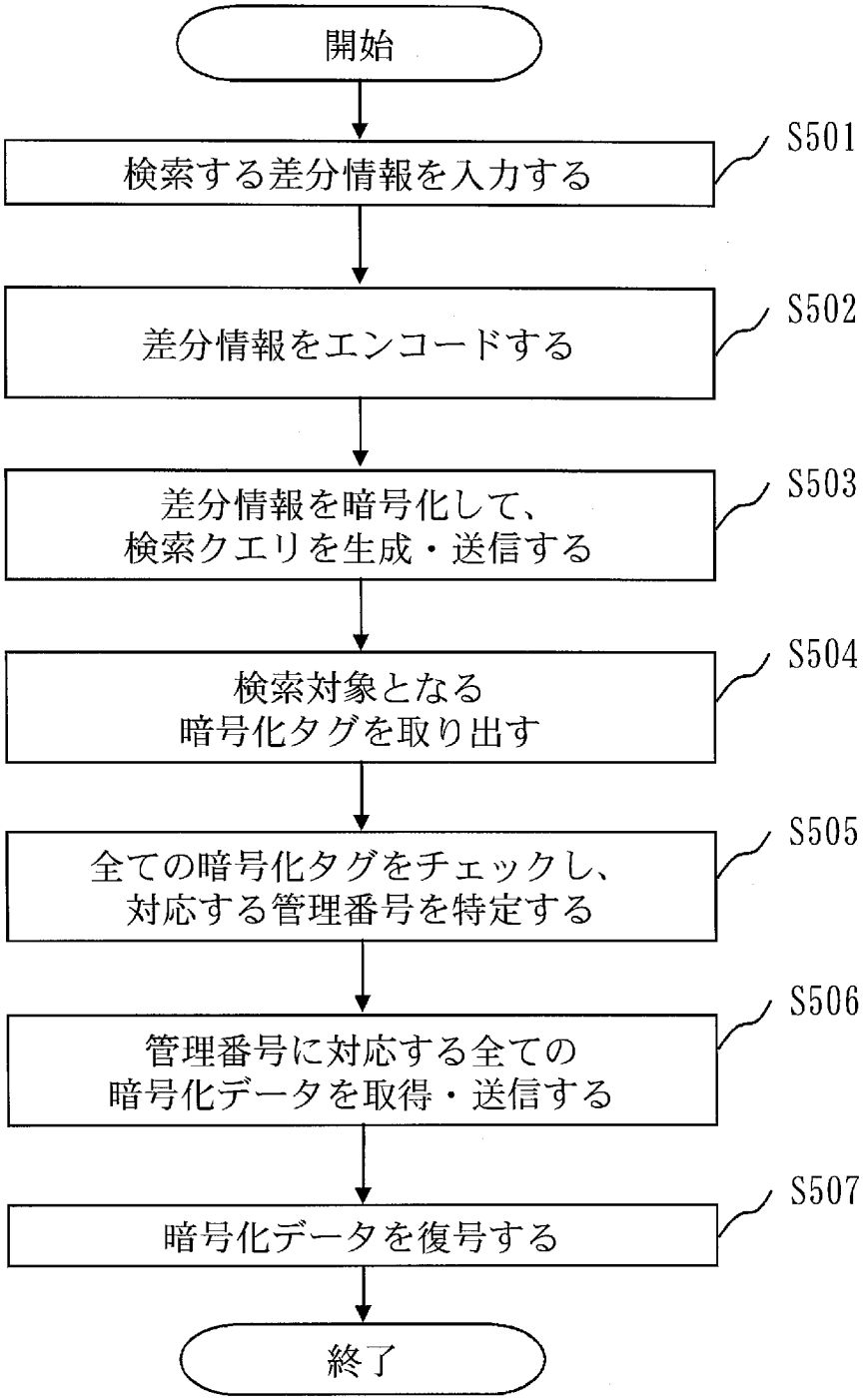
[図21]


グループ名：A製薬会社
ユーザ名：*
管理番号
患者ID
暗号化データ
000001
12345678
po8q234ndf09234n
000002
98765432
84gh4¥sfd78341n#
グループ名：B病院
ユーザ名：*
管理番号
患者ID
暗号化データ
100001
12345678
82jd723j9k#s%df4'
100002
98765432
000001と同一

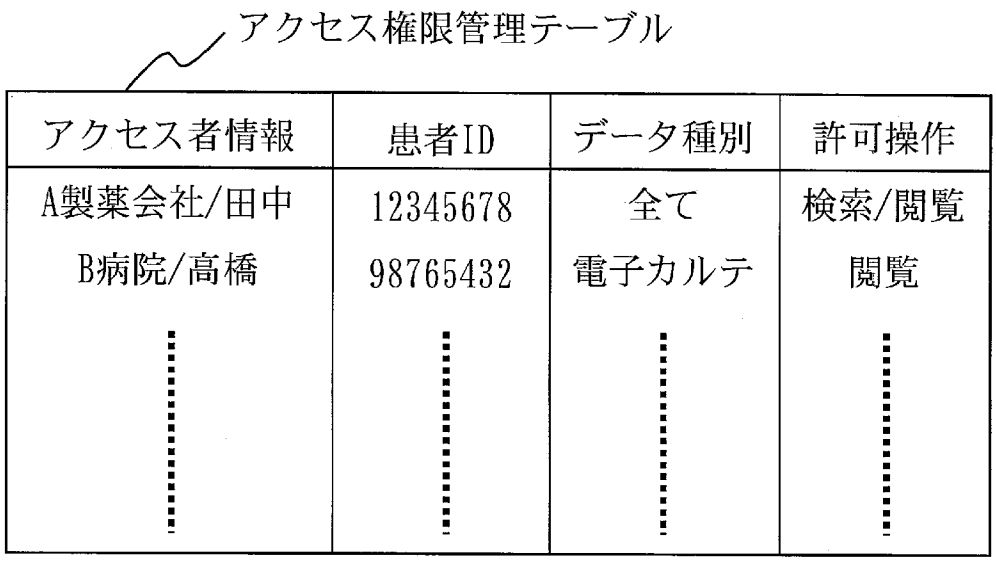
[図22]


開始
データを復号可能なグループ
及びユーザを決定する
S401
データ本体を暗号化して、
暗号化データを生成する
S402
暗号化データを、
データセンタに登録を依頼する
S403
暗号化データを記憶する
S404
終了

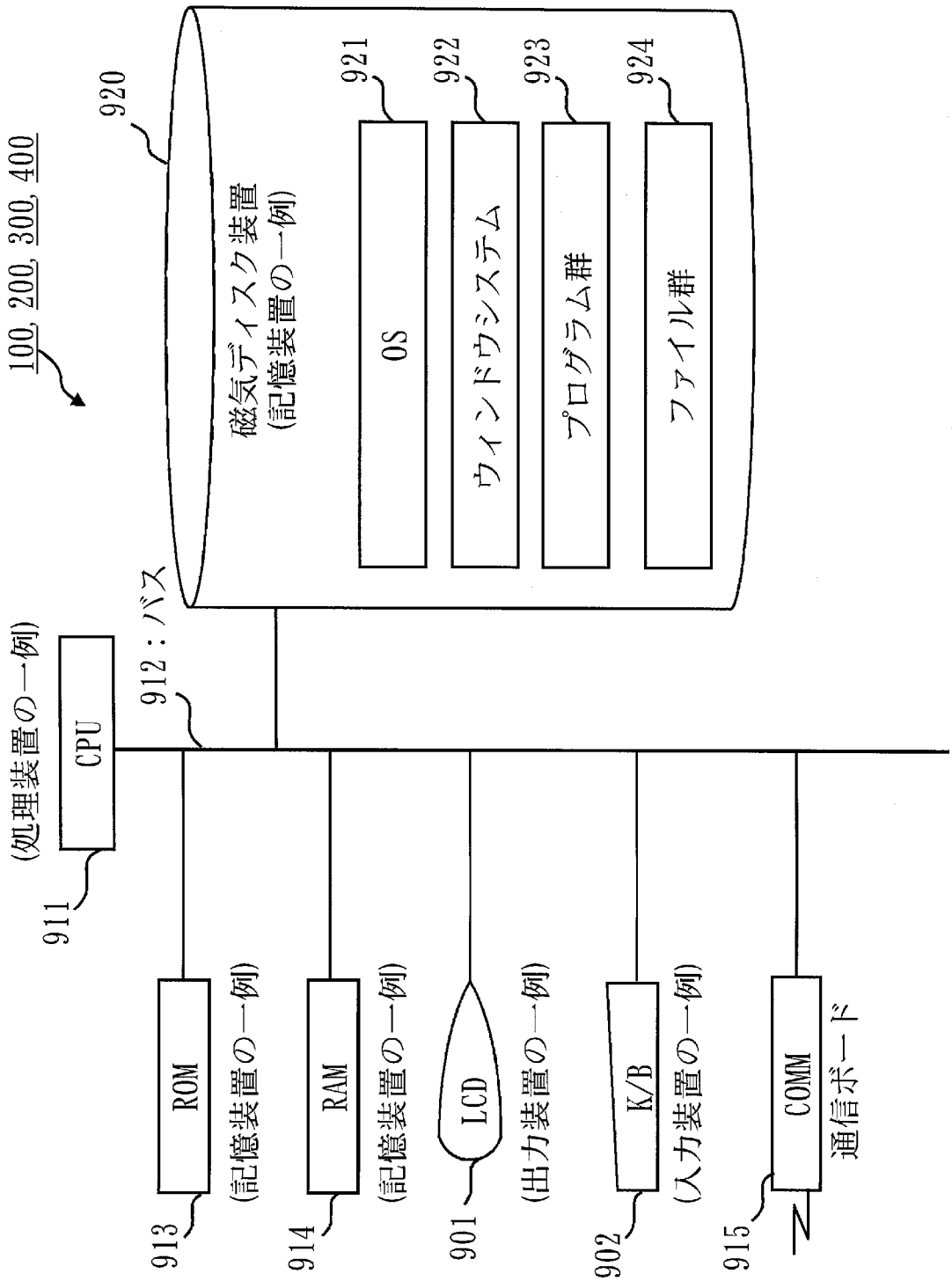
[図23]


開始
検索する差分情報を入力する
S501
差分情報をエンコードする
S502
差分情報を暗号化して、
検索クエリを生成・送信する
S503
検索対象となる
暗号化タグを取り出す
S504
全ての暗号化タグをチェックし、
対応する管理番号を特定する
S505
管理番号に対応する全ての
暗号化データを取得・送信する
S506
暗号化データを復号する
S507
終了

[図24]


アクセス権限管理テーブル
アクセス者情報
患者ID
データ種別
許可操作
A製薬会社/田中
12345678
全て
検索/閲覧
B病院/高橋
98765432
電子カルテ
閲覧

[図25]


100, 200, 300, 400
920
磁気ディスク装置
(記憶装置の一例)
921
OS
922
ウィンドウシステム
923
プログラム群
924
ファイル群
912：バス
(処理装置の一例)
CPU
911
ROM
(記憶装置の一例)
913
RAM
(記憶装置の一例)
914
LCD
(出力装置の一例)
901
K/B
(入力装置の一例)
902
COMM
通信ボード
915

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2014/054145

A. CLASSIFICATION OF SUBJECT MATTER
G06F19/22(2011.01)i, *G06F17/30*(2006.01)i, *G09C1/00*(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)
G06F19/22, G06F17/30, G09C1/00

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.
Y A	JP 2012-73693 A (Mitsubishi Space Software Co., Ltd.), 12 April 2012 (12.04.2012), paragraphs [0020] to [0042] (Family: none)	1,6,11,16, 21-23 2-5,7-10, 12-15,17-20
Y	JP 2003-177992 A (Seiko Epson Corp.), 27 June 2003 (27.06.2003), paragraphs [0067] to [0087] & US 2003/0108021 A1 & CN 1426173 A	1,6,11,16, 21-23
A	Shogo SHIMIZU et al., "DAS Model ni Okeru Anzen na Ruiji Mojiretsu Kensaku Hoshiki no Teian", Journal of the DBSJ, 27 June 2008 (27.06.2008), vol.7, no.1, pages 13 to 18	1-23

☒ 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 May, 2014 (09.05.14)	20 May, 2014 (20.05.14)
Name and mailing address of the ISA/ Japanese Patent Office Facsimile No.	Authorized officer Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2014/054145

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,A	Kazushiro MINEGISHI et al., "Data Outsourcing ni Okeru Privacy Hogo SNP Kensaku System no Jisso", Dai 5 Kai Forum on Data Engineering and Information Management (Dai 11 Kai The Database Society of Japan Nenji Taikai) [online], 31 May 2013 (31.05.2013)	1-23

国際調査報告

国際出願番号　ＰＣＴ／ＪＰ２０１４／０５４１４５

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, G06F17/30(2006.01)i, G09C1/00(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.　G06F19/22, G06F17/30, G09C1/00

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
Y	JP 2012-73693 A（三菱スペース・ソフトウエア株式会社）2012.04.12,　【００２０】－【００４２】（ファミリーなし）	1,6,11,16,21-23
A		2-5,7-10,12-15,17-20
Y	JP 2003-177992 A（セイコーエプソン株式会社）2003.06.27,　【００６７】－【００８７】 & US 2003/0108021 A1 & CN 1426173 A	1,6,11,16,21-23

☑　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
０９．０５．２０１４	２０．０５．２０１４

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　３６５８
松野　広一
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
Ａ	清水　将吾　外１名，ＤＡＳモデルにおける安全な類似文字列検索方式の提案，Ｊｏｕｒｎａｌ　ｏｆ　ｔｈｅ　ＤＢＳＪ，2008.06.27，Ｖｏｌ．７　Ｎｏ．１，ｐｐ．１３－１８	1-23
Ｐ，Ａ	嶺岸　和城　外３名，データアウトソーシングにおけるプライバシー保護ＳＮＰ検索システムの実装，第５回データ工学と情報マネジメントに関するフォーラム　（第１１回日本データベース学会年次大会）　［ｏｎｌｉｎｅ］，2013.05.31	1-23