DNA-IDを用いたDNA個人情報管理システムの提案

板倉征男[†] 辻井重男[††]

DNA情報は人体の設計図ともいわれ，絶対的なヒューマンID機能を持つとともに，究極の医療情報として貴重な個人の情報である．情報量のエントロピーが多大であり，かつ個人のプライバシ情報であるゆえその活用は両刃の剣となり，一歩誤って適用されると人間性の尊厳にかかわる社会問題となる．本論文では，個人情報として慎重に取り扱い，安全な管理を行うべきDNA情報，特に遺伝子領域のような個人のプライバシに直接かかわるDNA個人情報を対象にして，DNA-IDを用いたセキュアな管理方式について述べる．また具体的なDNA個人情報管理システムについて提案する．

DNA Personal Information System for Secure Management Using DNA-ID

Yukio Itakura[†] and Shigeo Tsujii[††]

DNA information, called a design for human body, has absolute personal ID functions and is valuable personal security data for ultimate medical information view. DNA information has much entropy and private data. Therefore, this causes social problems being worth of reverence for human nature if the use of DNA information is wrong. This paper demonstrates a method of secure management system that can treat cautiously for personal data to be secure control, especially for personal DNA information based on protection of privacy such as genetic field. In addition, the paper proposes for personal DNA information management system concretely.

1. はじめに

20世紀における最大級の発明および発見といわれるコンピュータと遺伝子構造解析は，やがてバイオインフォマテックス（生命情報科学）という研究領域を生み出し，21世紀に継承されてさらなる進展をとげようとしている．

人間の細胞核には4種類の塩基が約30億並んだ，いわゆるDNA情報がたたみ込まれており，人体の設計図の機能を果たしている．

この塩基配列の解明は完了しつつある．その中で設計図として意味のある部分（遺伝子）は約10万カ所あるといわれるが[1)]，その機能解析は一部が解明されているだけで本格的な遺伝子配列解析，構造および機能の特定は，ポストゲノムの主要課題として21世紀に引き継がれた．

今後は，ホモロジー検索等のデータマイニングを中心とするあらゆるコンピュータ技術が動員され，課題の解明がさらに進展することとなろう[2)]．

一方，これらを含めた人間のDNA情報は，個人の究極の人体設計図なのであるから，個人のプライバシという観点から見れば，その取扱いについては解明のエネルギーと同様な知力を注いだしくみ作りが必要である．病気の根本的治療や創薬のための貴重な情報であるが，一歩情報の管理を誤れば取り返しのつかない社会的道徳的混乱を招くとも限らない．

このためにはまず人間の尊厳性の原点に立脚した総合的な「DNA情報に関する倫理法と取扱指針」の策定があり，次にこれらをきちんと実施するための社会システムの構築と適正な運用手段が必要である．

従来バイオインフォマテックスの領域における情報工学の役割は，もっぱら遺伝子構造解析に注力されてきたが，今後は個人のDNA情報の適正で合理的な管理について情報セキュリティ技術の果たすべき役割が新たに生じるものと考える．

本論文ではまずDNA情報と現代暗号技術の接点について述べ，考えられるいくつかのかかわり合いについて提言する．

その1としてDNA情報のうち遺伝子としての意味や，病因にはかかわらないが人によって著しく個人差

† NTTデータテクノロジ株式会社
NTT DATA TECHNOLOGY Corporation
†† 中央大学
Chuo University

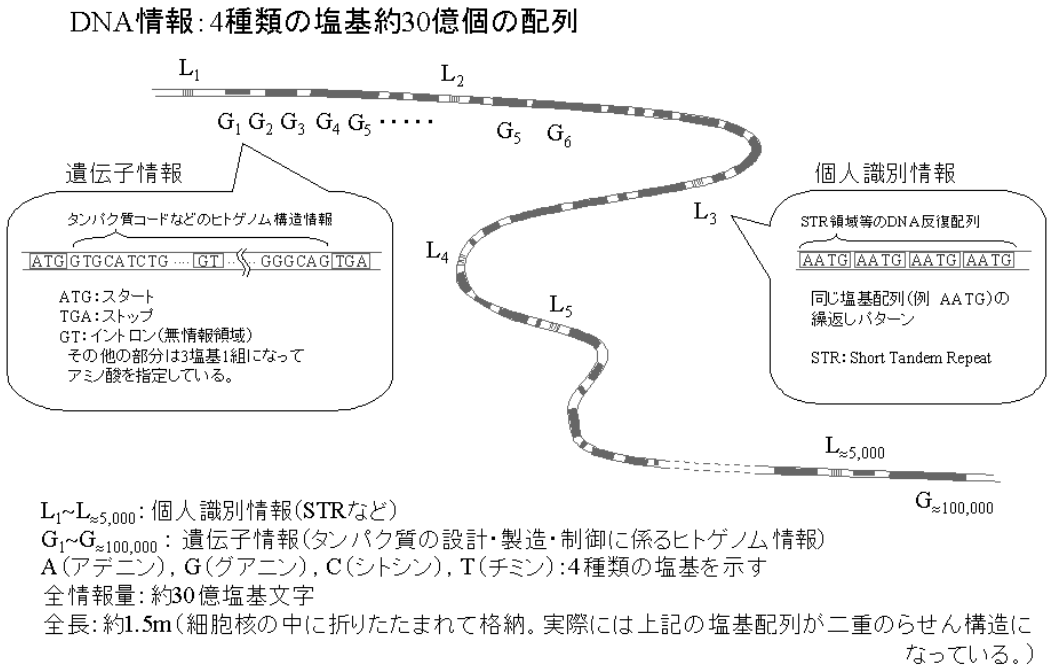


図 1 DNA 情報の模式図
Fig. 1 Structure of DNA information.

のある部分を個人識別用情報として取り上げ，これらが厳密な個人の ID（DNA-ID）として適用できるかという点について提案する．

次に，DNA 情報を秘密鍵に組み込んだ公開鍵暗号方式を提案する．これにより情報の暗号化ばかりでなく，生体情報を使った本人認証方式が実現できる．

その 2 として DNA 情報は，個人の重要なプライバシ情報を含むものであるから，これを適切に管理するしくみに上記で提案する暗号方式を適用することを提案する．

最初に作った公開鍵暗号方式を適用し，これらがあるべき DNA 倫理法規と呼応して適切に運用されれば，個人の DNA 情報は本人の承諾を前提として利用するという原則を遵守できるしくみとなろう．

なお DNA 情報は，人間のみならず動物および植物に共通するしくみであり，その解明が同様に進んでいるが，ここでは人間社会の情報セキュリティを研究対象としているので，人間の DNA 情報に限定して論ずることとする．

2. DNA 情報の概要

2.1 DNA 情報のしくみ

人体はおよそ 50 兆個の細胞で成り立っている．各細胞は 1 個の細胞核と複数のミトコンドリア等を有する．直径およそ 5 ミクロンの細胞核の中には，約 30 億の塩基配列がデジタル情報として格納されている．本論文ではこの塩基配列を DNA 情報と呼ぶ．

塩基配列は 4 つの塩基を要素としている．各々の塩基の頭文字をとって A（アデニン），G（グアニン），C（シトシン），T（チミン）と呼ぶ．DNA 情報は塩基文字による暗号文とも考えられる．これが遺伝子暗号ともいわれるゆえんである．DNA 情報は人体のどの部分の細胞をとっても基本的な配列は同じで，終生不変である．

ミトコンドリアは細胞の中でタンパク質の生成に必要なエネルギーを供給する機能を持つが，その他細胞核中と同様な DNA 情報がある．そのうちで個人によって 1 つの塩基の違いがある部分（SNPs：Single Nucleotide Polymorphisms という）をとらえて個人識別情報に用いることも研究されている[3]．

図 1 は，本論文で取り扱う DNA 情報に着目して描いた模式図である．

2.2 個人識別情報

$L_1, L_2, L_3, \cdots, L_{\approx 5000}$ は，STR（Short Tandem Repeat）と呼ばれる領域で，通常 4 文字の塩基配列が数回〜数十回繰り返し並んでいる領域である．繰返し回数を STR のアリル（Allele）といい，回数の個人差が著しいので個人識別用の情報として適用される．

STR 等 DNA 反復配列のある座位は，人体では数千カ所以上あるが，実際に識別に使われる箇所は 16 カ所が実用化されている[4]．これは STR の各座位に対応する酵素試薬の開発に依存している．

各座位にはその箇所を特定する名称がつけられている．法医学鑑定で一番使われる TH01 という STR 座位については，繰返し塩基配列が AATG となっている．

実際の DNA 情報は，図 1 で示した塩基配列が 2 重のらせん構造となって，情報としても完全に 2 重化されている．さらに，細胞分裂時の染色体を見れば分かるとおり，これらがさらにペアになっており，一方は父親から，他方は母親からの DNA 情報を引き継いでいる．STR のアリルについても 1 座位につき 1 ペアの情報となる．たとえば TH01 の座位についてその人のアリルすなわち 4 塩基文字の繰返し回数を調べると，(11 回，13 回) のように 2 つの数字の情報となる．各々の数値は父親と母親の持つアリルから 1 つずつの情報を受け継いだものである．

なお，STR は図 1 の DNA 情報の模式図にあるように，遺伝子情報とは異なる領域にあるので，人体のたんぱく質の製造に関する情報や，その情報の乱れが原因で起こる病気には関係しない情報である．

STR にはもっぱら塩基配列の反復回数に個人差があることにのみ意味のある情報として，DNA 鑑定等に使われる[5]．

2.3 遺伝子情報

$G_1, G_2, G_3, \cdots, G_{\approx 100,000}$ は，タンパク質のコード領域等に代表されるヒトゲノム構造情報で一般に遺伝子情報とはこの部分の塩基配列を指す．情報量としては，数十塩基から長いのは数千塩基に及ぶものがある．

3 つの塩基が 1 組となってアミノ酸に対応しており，それらが複数の集合情報となってタンパク質の立体構造や制御等，人体の構造設計に関与する情報となっている．なかでも ATG (スタート指令)，TGA (ストップ指令) 等の塩基配列は，制御用の情報となっている．またイントロンといわれるダミーの情報を示す塩基配列があり，長いものでは数百字から成るイントロンが含まれている．G の領域でタンパク質の立体構造と機能まで完全に解明されたものはごく一部にすぎず，21 世紀のバイオインフォマテックスの最重点課題となっている．

またこれらの領域での塩基配列の 1 つの乱れがタンパク質の製造に直接影響し，それが間接的に病気の原因になる．

またその情報は，創薬の研究情報となる．将来は個人の DNA 情報の異常を見つけ，病状にカスタマイズした治療法が現れることになろう．

$G_1, G_2, G_3 \cdots$ は 10 万カ所あるといわれているが，今後の遺伝子機能の解析が進めば，さらに新たな遺伝子の存在が発見されることも考えられる．

DNA 情報の概要を**表 1** に整理して示す．

表 1 DNA 情報の属性分類
Table 1 Attribute classification of DNA information.

区分	領域	機能	情報量	備考
個人識別情報	STR (Short Tandem Repeat) 領域 など	DNA反復配列があり、反復数の個人差が明白な情報	反復回数情報 ×5000ヶ所以上	模式図(図1)対応 L_1〜$L_{\approx 5,000}$
遺伝子情報	ゲノム構造情報領域	人体の設計図(ヒトゲノム構造情報)といわれる領域でタンパク質の設計・製造に係る情報、制御情報、病気治療や創薬など係る情報	数十〜数千塩基配列情報 ×約10万ヶ所以上	G_1〜$G_{\approx 100,000}$

(人体の場合を示す)

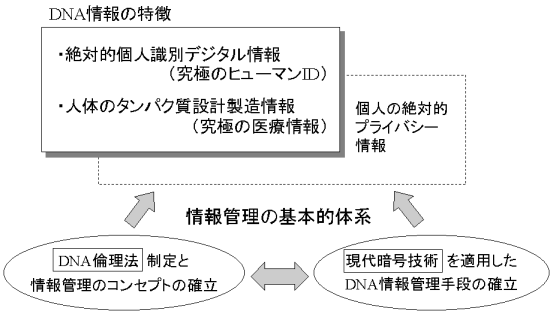


図 2 DNA 情報の特長とあるべき情報管理のしくみ
Fig. 2 Characteristic of DNA information and ideal information management systems.

2.4 DNA 情報の特徴とあるべき情報管理のしくみ

DNA 情報は個人の絶対的なプライバシ情報である．適正に活用すれば，絶対的な個人識別用 ID となり，医療情報としても根本的な病因の究明とそれに結び付いた治療や薬事を施す基礎情報となる．

一方適用を誤れば，個人の尊厳にかかわるプライバシの冒とくをきたすことになる．

図 2 にこのような DNA 情報の特長を考慮した，あるべき情報管理の概念を示す．

まず DNA 情報の情報管理のあり方を社会的活用と個人のプライバシ保護という二律相反する視点から十分衆議を尽くし，管理方法に関するコンセプトを確立する必要がある．社会的活用とは医療分野における治療情報，法医学分野における犯罪捜査や行方不明者の確認等の個人識別・鑑定情報を対象とするもので，なんらかの条件で情報を開示して流通させなければ DNA 情報を活かして使えない．

表 2 個人情報を守る現行法
Table 2 Present laws for personal information security.

法律名	条文	守秘義務	守秘情報	備考
国家(地方)公務員法	100(34)条	国家(地方)公務員	職務上知り得た秘密	退職後も守秘義務あり
医療法	72条	医師，歯科医師，助産婦	業務上秘密，個人の秘密(診療録，助産録情報)	罰則:1年以下，50万円以下
刑法	134条	医師，薬剤師，医薬品販売業者，助産婦，弁護士，弁護人，公証人	業務上知り得た秘密	罰則:6ヶ月以下，10万円以下 退職後も守秘義務あり
弁護士法	23条	弁護士	職務上知り得た秘密	秘密保持権利及び義務がある
個人情報保護法	14条	病院，診療所，助産所	不開示情報:診療に関する個人ファイル	個人の生命，身体，財産，その他個人の利益を害する個人ファイルは不開示
		裁判所，検察庁，警察	処分，刑執行に関する個人ファイル	

表 3 遺伝子研究にかかわる倫理規程
Table 3 Ethical codes for analysis of genomes.

	規程	主管	制定時期	主な規程内容
1	遺伝子解析研究に付随する倫理的問題等に対応するための指針	厚生科学審議会先端医療技術評価部会(厚生省)	2000.5	・各研究機関に「倫理審査委員会」を設置し，「個人識別情報管理者」を選任 提供者の同意を義務づけ
2	ヒトゲノム研究に関する基本原則について	科学技術会議生命倫理委員会(科技庁)	2000.6	・基本原則と研究指針 ・各研究機関に「個人情報管理者」を選任 提供者の同意を義務づけ
3	ヒトゲノム・遺伝子解析研究に関する倫理指針	3省関係者会議(文部科学省，厚生労働省，経済産業省)	2001.3	・1．2の指針を一般化し，研究機関以外にも適用するもの 遺伝子の提供は患者の同意が必要 ・インフォームド・コンセント 十分な説明が必要 ・「個人識別情報管理者」を選人，「倫理審査委員会」で研究実施計画を事前承認が必要 ・2001年4月から施行する

一方，個人のプライバシ保護という観点からは，DNA 情報の取扱いに関する基本原則や規定が制定されている[6),7)]．これらの経緯をふまえて総合的な社会コンセプトの醸成をはかり，DNA 倫理法のような法制化が論じられることになると思われる．

表 2 に個人情報を守る現行法を示す．また，**表 3** に 2000 年度から本格化している遺伝子にかかわる倫理規程の制定状況を示す．

次にこれらの社会通念の実現を支援する手段として，図 2 のトライアングルの一翼に示すように，情報セキュリティ技術，具体的には暗号技術を導入し，適確な情報管理システムを構築することを提言したい．

3. DNA 情報と現代暗号技術の連携

3.1 個人識別情報を ID 情報とする現代暗号技術の導入[8)~11)]

複数座位から成る STR（Short Tandem Repeat）のアリル情報（4 塩基配列の繰返し回数）を ID 情報とし，これに個人別の乱数を加えて本人の秘密鍵とする．これよりペアとなる公開鍵を生成して CA（公開鍵認証機関）に登録する．このときの公開鍵は，通常の公開鍵暗号システムのサービスの一環として管理されるので，以後署名，本人認証および暗号通信等の現代暗号技術のサービスが享受できる．

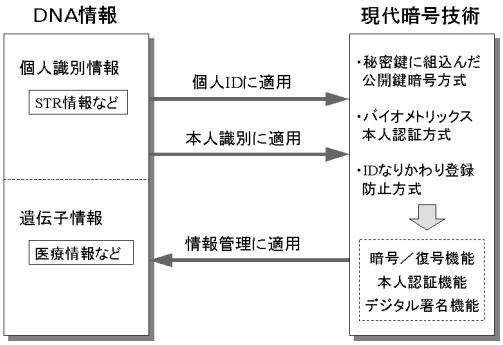


図 3 DNA 情報と現代暗号技術の連携
Fig. 3 Collaboration between DNA information and cryptography technology for security.

後ほど，この方式を使った本人の遺伝子情報を管理する方法を提案する．

また，公開鍵に DNA 情報が組み込まれているので，裁判所等で本人確認を行う場合は，バイオメトリックスによる本人認証を間違いなく行うことができる．

STR アリル情報は，両親から子供へと受け継がれるので，最初に DNA 個人情報採取時に行う本人確認は，両親または親族との STR アリル情報の照合ができれば，パスポート等による補助的確認を割愛して登録を行うことが可能となる．これはなりかわり ID 登録の防止策として有効な方式となる．ただし親子関係そのものが重要なプライバシ情報であるから，運用に際しては倫理的視点や道徳律をふまえた十分なアセスメントが必要である．**図 3** に DNA 情報と現代暗号技術の連携をまとめて示す．

3.2 DNA 情報を秘密鍵に組み込んだ公開鍵暗号方式

(1) 個人識別情報から DNA-ID の生成方法

個人識別情報は，口腔粘膜細胞を綿棒でこすって採取し，複数座位の STR アリル情報を分析機器で読み取る．

i 番目の座位から得られるアリルのペアの数値を $L_i(p_i, q_i)$ とする．ここで p_i，q_i は，$p_i \leq q_i$ となるよう小さい順に並べることとする．

これより，

$$\begin{aligned}\alpha_A &= L_1, L_2, L_3, \cdots, L_M \\ &= p_1, q_1, p_2, q_2, p_3, q_3, \cdots, p_M, q_M \end{aligned} \quad (1)$$

のような数列が得られる．

たとえば $L_1 = (16, 17)$，$L_2 = (7, 8)$，$L_3 = (30, 32), \cdots, L_M = (18, 21)$ のときは $\alpha_A = 1617783032 \cdots 1821$ のようになる．この α_A を ID 情報とする．各 STR アリルの出現頻度は各民族ごと

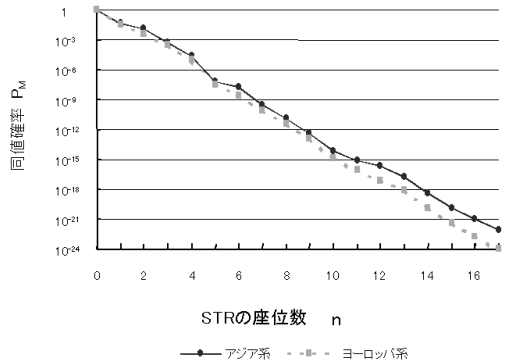


図 4 STR の座位数 n と ID の同値確率 P_M の関係
Fig. 4 Relationship between STR loci combination counts n and identification matching probability P_M.

の統計値が発表されているが[12]，これらを基に任意の 2 人の間で α_A が同値となる確率 P_M を計算すると，STR が 16 座位で各座位間の相関がないとした場合，

$$P_M \approx 10^{-21} \tag{2}$$

となる．これはヨーロッパ系人種の場合であるが，民族間によって計算の根拠となる統計値が異なるので P_M も若干変動する[13]．**図 4** は STR の座位数と P_M の関係を計算した結果を示す（付録参照）．

この α_A を N 人の集団で ID として使った場合，どこかの組合せで最低 1 組の同値の ID が出る確率を P とすると

$$P \approx \frac{N(N-1)}{2} \cdot P_M \tag{3}$$

となる．$N = 10^8$（1 億人）の場合は $P = 10^{-5}$ となり，10 万人に 1 人程度の割合で同じ ID が現れる．

もし集団の ID として必要な同値確率が得られない場合は，STR の座位数を増やすことにより所定の確率以下となるようにする．

（2） 秘密鍵の生成方法

秘密鍵 δ_A は，式 (4) により生成する．

$$\delta_A = \alpha_A + r_A \tag{4}$$

r_A は大きな乱数で，本人の秘密とする．乱数 r_A を加える理由は，α_A が裸のままでは必ずしも秘密が保てないからである．すなわち，他人の人体細胞をなんらかの形で取得できれば他人の α_A を採取できるので，秘密鍵とするにはこれに秘密乱数を加えなければならない．

また，$\delta_A = \alpha_A{}' + r_A{}'$ となるような $\alpha_A{}'$ を持つ他人が現れたとき，これと本人の δ_A が識別できるよう $g^{r_A} \bmod p$ を公開しておく．離散対数の性質から $g^{r_A} \bmod p$ を開示しても r_A を計算することは困難であるから，r_A の秘密は保たれる．また上記の $\alpha_A{}'$ が現れても，g^{r_A}（公開）と $g^{r_A}{}'$ から計算される公開鍵が本人登録鍵と違う値となるので，実際の本人との識別が可能である．

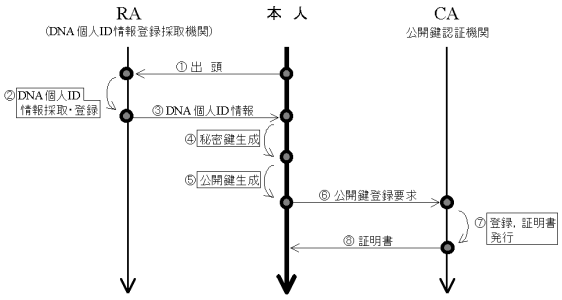


図 5 DNA 個人 ID 情報の採取と公開鍵の登録
Fig. 5 Acquisition and registration of DNA personal ID information.

（3） 公開鍵の生成方法

ここでは前項の流れから暗号方式として離散対数系の ElGamal 暗号を適用することとし，公開鍵はその方式に従って式 (5) により生成する．

$$Y_A = g^{\delta_A} \bmod p = g^{\alpha_A + r_A} \bmod p \tag{5}$$

ここで p は大きな素数，g は乗法群 Z_p* での原始元[14] で，ともにシステム値として公開される．

公開鍵 Y_A および（2）で述べた理由で開示する g^{r_A} とともに CA（公開鍵認証機関）に登録し，公開する．CA は他からの問合せ手続きに対して Y_A 等の公開情報と本人の対応を CA の証明書として発行し，返答する．

公開鍵登録までの本人，RA（DNA 個人 ID 採取登録機関），CA との間のプロトコルを**図 5** に示す．

これまでに生成した秘密鍵と公開鍵は，ElGamal 暗号の鍵として情報の暗号化および復号化に使われる．後述する DNA 情報の管理にこの暗号方式を適用する．

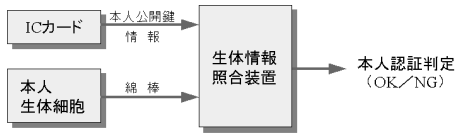
鍵の情報は本人の氏名とともに IC カードに記録し，保管する．これは登録済みの実印の機能を果たすことになる．安全上秘密鍵の情報は，この IC カードには入れず別のカードに記録して保管することが望ましい．

3.3 DNA 情報を用いた本人認証方式

指紋と同様に DNA 個人識別情報を使ったバイオメトリックス本人認証を行うことができる．

公開鍵情報とその他の公開情報等を記録した IC カードの持主が本人であることの正否は，**図 6** のような生体情報照合装置を用いて判定される．入力は IC カード情報と本人の生体細胞を綿棒で採取した情報である．

採取した生体情報から公開鍵 Y_A が生成できれば，本人である[10]．



●生体情報照合装置の機能

①本人の生体細胞をDNA解析し、α_Aを生成する。

② α_Aから $g^{\delta_A} \bmod p = g^{\alpha_A + r_A} \bmod p = g^{\alpha_A} \cdot g^{r_A} \bmod p$を計算し、ICカードに記録されている$Y_A$情報と一致すれば本人（ICカードの持主）と判別する。

注）Y_A, p, g, g^{r_A} は公開されており、本人氏名と共にICカードに記録されている。

図 6 DNA 情報を用いた本人認証
Fig. 6 Personal authentication by DNA information.

図 7 DNA 個人情報管理のコンセプト
Fig. 7 Concept of DNA personal information management.

ここでなりすまし攻撃の可能性について述べる．他人の生体情報を盗んでなりすまし登録に成功しても，将来分析がリアルタイムでできる端末装置ができれば本人生体情報と違うことが分かり，なりすましは不可能である．

またすでに登録してある人の秘密鍵 δ_A $(= \alpha_A + r_A)$ を盗み，α_A とは違う自分の生体情報 $\alpha_A{}'$ から $\delta_A = \alpha_A{}' + r_A{}'$ なる $r_A{}'$ を計算して本人になりすますことが考えられる．この場合正当な公開鍵 Y_A $(= g^{\delta_A} \bmod p)$ を登録する際，$g^{r_A} \bmod p$（r_A は個人の秘密乱数）を同時に登録し公開しておくので，$\alpha_A{}'$ から本人の公開鍵を生成できるか検証すれば，なりすましか否かは判別できる．

つまり，なりすまし者が自分の生体から $\alpha_A{}'$ を採取して端末装置に入れて，オンサイトチェックする際，装置が公開してある g^{r_A} を取り寄せて Y_A を $g^{\alpha_A{}'} \cdot g^{r_A} \bmod p = g^{\alpha_A{}' + r_A} \bmod p$ から計算しようとしても登録済みの正しい Y_A

$$Y_A \ (= g^{\alpha_A + r_A} \bmod)$$

とは一致させることができないからである．

なお，一卵性双生児の場合は α_A が完全に同値となるため，δ_A または r_A が盗まれるとなりすましが可能となる．この場合は別の対策を講じる必要がある．

4. DNA-ID を用いた DNA 個人情報管理方式

4.1 DNA 個人情報の管理の概要

DNA 情報のうちの遺伝子情報のような部分は，本人の重要なプライバシ情報となる．ここではこのような情報を DNA 個人情報と呼ぶことにする．DNA 個人情報のセキュアな管理を行う．基本的な考え方は，表 3 に示した遺伝子にかかわる倫理規程類に共通する「遺伝子の提供は本人の同意が必要」というコンセプトをサポートするしくみ作りであり，この主旨に沿う社会システムを，暗号技術を十分適用して構築しようというものである．

このため前章で述べた公開鍵暗号方式で，本人の秘密鍵がなければ採取後の登録も閲覧もできないようなしくみを作ることができる．

採取した DNA 情報に本人のデジタル署名を添付することを義務づけ，署名のない DNA を取り扱うことは禁令とする．ちょうど透かしのないお札，すなわち“ニセ札”を使うことが禁じられているのと同様なルールとなる．

DNA 情報は毛根付きの毛髪 1 本で他人の情報が採取できるので，このようなルール作りが必要と思われる．

すなわち他人の DNA 情報を自分のものであるかのように偽って登録したり，自分の DNA 情報を他人のもののように偽って登録したりしても，署名確認や現場で DNA を採取してすでに DB（データベース）に書き込まれている DNA 情報と比較チェックをすれば，容易に不正を発見できる．ただし，DB にはこのような判定のために遺伝子情報だけでなく，個人識別用情報（DNA-ID のもととなる情報）も記録しておくこととする．

図 7 は DNA 個人情報管理のコンセプトをシステムの構築の前提として整理したものである．

なお，現行規定は学術研究を前提としたものであるため，DNA 個人情報の社会的活用まで考慮した管理規範とはいえない．

そこで図 7 に示す DNA 個人情報管理の考え方では，2 項以降の「社会システムとして位置づけた場合の情報アクセス，移動・流通時の本人の同意やデジタル署名の義務化」等の考え方を加えた．

4.2 DNA 個人情報管理システムの機能

DNA 個人情報管理システムは，情報の採取と記録，他人からのアクセスおよび移動・流通の機能を持つ．ここでいう DNA 個人情報は，全ゲノム情報ではなくその人個人に特有な情報を想定している．暗号技術が

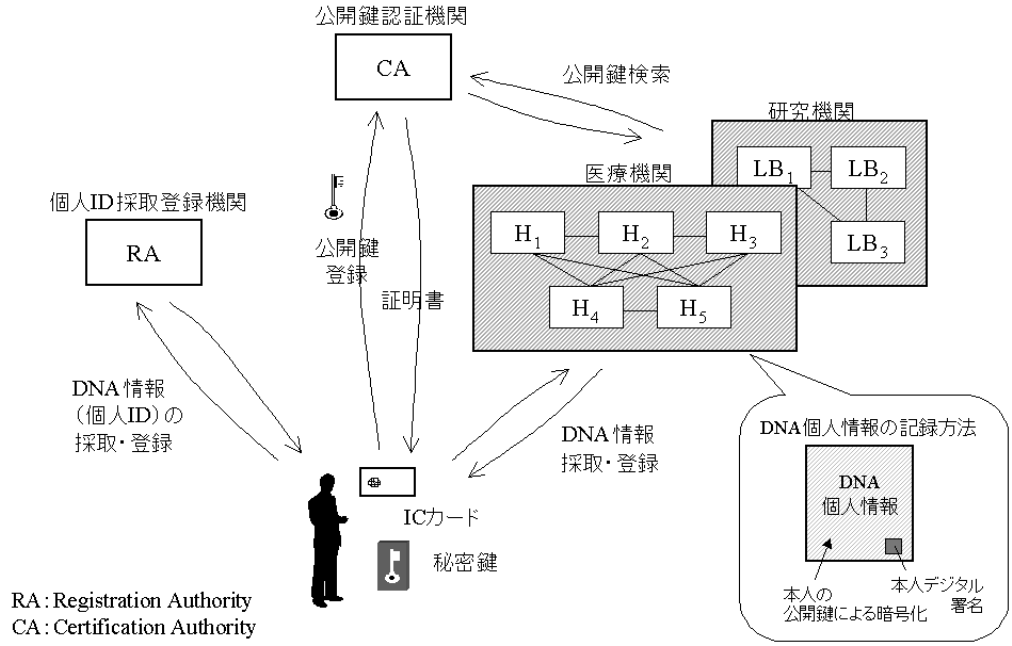


図 8 DNA 個人情報管理システムの概念図
Fig. 8 Conceptual design of the DNA personal information management system.

登場するのは，以下の機能を発揮する場合である（図 3 参照）．

暗号化/復号化

DNA 個人情報を DB に記録するとき，もとの情報の暗号化を行う．

本人の公開鍵で暗号化する．復号化は本人の同意すなわち本人の秘密鍵の直接的，間接的提供を得て可能となる．

このほか，他人アクセスおよび流通の機能を使う場合も，公開鍵暗号方式により暗号化通信が行われる．

デジタル署名

DNA 個人情報の採取時および記録時の本人同意の証として，もとの DNA 個人情報に本人のデジタル署名を付与する．すなわち本人の秘密鍵の提供を得て可能となる．

デジタル署名のない DNA 個人情報は意味を持たないし，そのような DNA 個人情報を取り扱うことは透かしのないニセ札を使うと同様，違法とすることを想定している．

4.3 DNA 個人情報管理システムの構成

図 8 は実際の DNA 個人情報管理システムの構成を示す概念図である．RA および CA とのやりとりで公開鍵を登録することは前述したが，ここでは医療機関や研究機関がそれらの基本機能をネットワーク経由で活用して DNA 個人情報の管理を行うしくみを提案している．

表 4 DNA 個人情報管理システムの機能
Table 4 Functions of DNA personal information management system.

DNA個人情報の取扱い	システムの情報管理機能	関連する暗号技術
●採取 ●登録 ●保管	①採取時の本人承諾	本人の秘密鍵で署名
	②DB登録保管時の本人承諾	本人の秘密鍵でデジタル署名
	③同登録時の暗号化	本人の公開鍵で暗号化
	④同読出時の復号化	本人の秘密鍵で復号化
●他人からのアクセス	⑤DBアクセスの本人承認	本人の秘密鍵で署名
	⑥同読出時の復号化	本人の秘密鍵で復号化
●流通 ●移動	⑦DB間の情報移動または流通時の本人承諾	本人の秘密鍵で署名
	⑧同移動または流通後の復号化	本人の秘密鍵で復号化

医療および研究機関は，各機関において独自の DB を構築しているが，それらが次第にネットワークでつながり，連動化すると考えられる．もちろんそれには DB の格納やアクセス方法の標準化が基本である．

各医療機関は倫理規程に基づき，各機関の責任で DNA 個人情報の管理を行うが，本人の同意が必要なポイントでは**表 4** に示すように暗号技術を使ったシステムの情報管理機能を使うことになる．

図 9 は一例として医療機関で扱う医療用情報を例に，DNA 個人情報を採取してからこれを医療行為に

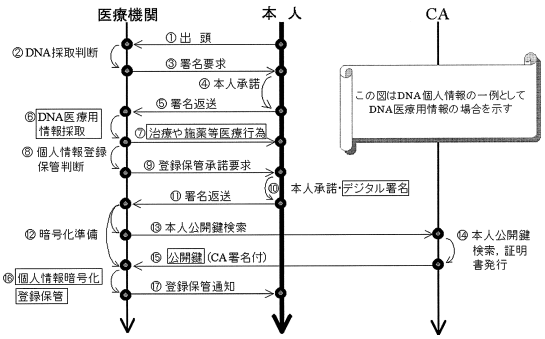


図 9 DNA 医療用情報の採取と登録
Fig. 9 Acquisition and registration of DNA medical information.

使うまでの本人および CA とのプロトコルを示している．

まず本人が医療機関に出頭する．医師は問診の結果本人の DNA 個人情報の採取が必要と判断した場合は，本人の承諾を得るために同意書に署名を要求する．本人は自分の秘密鍵（実際には専用の IC カード）を提示して医師のデスクに設置してある端末装置に差し込み，画面上の同意書に署名入力を行う．

医師は本人の承諾が得られたことをシステム上で確認し，DNA 個人情報を本人から採取・分析する．その情報に基づき必要な治療や施薬等の医療行為を行う．

医師は診療後継続的な治療のために，採取した本人の DNA 個人情報の保管が必要と判断する場合は，再び本人に対し登録・保管の承諾を要求する．

本人は承諾の是非を考え，是認の場合は前と同様本人の秘密鍵入りの IC カードを提示する．ここで暗号技術的には本人の DNA 個人情報に対してデジタル署名を行う．

実際は秘密鍵の入った IC カードを医師のデスクにある端末装置に差し込み，画面上のボタンをクリックすることによってデジタル署名が行われる．

この署名の処理には，DNA 個人情報がデータとして含まれているので，以降他者がなんらかの改ざんを行うことはできない．DNA 倫理法ではデジタル署名のない DNA 個人情報は無意味なもので，かつ署名なしの情報を取り扱うことを違法と規定する．

ここでシステムは CA から本人の公開鍵を取り寄せる．この公開鍵でデジタル署名を含む本人の DNA 個人情報を暗号化し DB に登録・保管する．

この情報は，本人の秘密鍵がないと復号化できないので，以降は DNA 個人情報は写しを取って使われることを未然に防ぐことができる．

現状では医療機関や研究機関におけるシステムの実体が十分把握できていないので，より実用的な DNA 個人情報管理システムの具体的設計は，今後の課題とする．

4.4 DNA 個人情報以外の DNA 情報の管理

本章で管理の対象とした DNA 個人情報とは，本人の身体的特徴や病因に関係する遺伝子のように，プライバシ保護を厳密に行う必要のある情報である．

一方，遺伝子のほとんどの部分は人類共通の部分であり，現に世界中の多くのヒトゲノム DB がインターネットで結ばれている．このような個人のプライバシにかかわらない DNA 情報はむしろ人類の共有財産として一定のルールの下に公開し，医学や薬学の研究に利用すべきものである．

また DNA-ID 自身も法医学分野で社会的活用の動きが始まっている．

4.5 一般的な個人情報管理への適用

前節までは DNA-ID を用いてもっぱら DNA の世界に閉じた情報管理方式論を展開したが，医療カルテや病歴記録のような DNA 以外の一般的な個人情報に関するセキュアな管理方式に本提案の考え方を適用することは可能である．ただし提案するシステムは相当高度で，厳密なプライバシ保護を必要とする情報管理方式であるから，コストと運用性についての検討と改善が必要と思われる．

5. 考　　察

（1） 法医学分野における研究との連携の必要性

法医学関連分野では，犯罪捜査や親子鑑定の立場から DNA 情報をヒューマン ID として適用する研究開発が従来から行われており，実用が進んでいる．特に米国の FBI は，CODIS（Combined Database Index System）という DNA 個人 ID 検索用の本格的な分散データベースネットワークシステムを整備する計画を示している．

このシステムは捜査権を持った警察当局の専用 DB であるから，暗号化等の情報セキュリティに対する対処は現状では行われていないようだが，個人識別のアルゴリズムや DB の検索方法については，相当の研究実用化の蓄積があるので，この分野との相補的な連携が必要と考えられる．

（2） 臨床医学分野との連携の必要性

本論文で取り上げた遺伝子情報のセキュリティについては，医療行為を行う臨床医学分野に深くかかわる問題である．21 世紀において遺伝子解明が進むとともに，個人情報としての遺伝子情報の適切な管理はますます重要な課題となると考えられる．すなわち，管

理が甘すぎると個人情報が本人の意志に反して悪用される危険性があり，堅すぎると究極の人体情報である遺伝子情報の活用を阻害する恐れがある．

今後はこの分野との実務面，研究面，情報技術面，法制面での情報交換と連携が必要である．

（3） DNA 個人情報管理システムの実用性

本論文では DNA 個人情報の管理を目的としたシステムのモデルを提言したが，現実にはすでに多くのゲノムデータベースが構築され，運用されているので，その中でこのシステムをどう位置づけるかを検討する必要がある．

また，医療用情報という面からは，本システムは医師のカルテ情報を取り扱うシステムの延長線上にあるとも考えられるので，双方の有機的な関係についてさらに整理する必要がある．

6. ま と め

本論文は情報セキュリティの視点で DNA 情報をとらえ，現代暗号技術を用いてセキュアな管理を必要とする DNA 個人情報の管理システムについて提案するものである．

まず現代暗号技術の視点で DNA 情報を眺め，2 つの属性を明らかにした．

その 1 つは，DNA 情報は DNA-ID としてきわめて識別率の高い個人識別情報をデジタル情報として持つことである．

これについては DNA 情報から個人の ID 情報を生成する方法を提案し，その統計的検証を行った．またこの ID を秘密鍵に組み込んだ構造を有する公開鍵暗号方式を提案し，あわせてバイオメトリックスによる本人認証方式を提案した．

その 2 つは DNA 情報の中の一部がいわゆる遺伝子情報といわれる領域であり，その中に個人の身体的特徴や病因にかかわるプライバシ度の高い情報が含まれていることである．

これについては最初に提案した公開鍵暗号方式をそのまま適用し，自分の鍵で本人の署名を行い，自分の DNA 個人情報を守るという社会システムを提案した．もちろんその実現には DNA 倫理法のような法的規定と道徳の確立が主役であり，システムはそのサポート手段として “従” の役割を果たすことが前提である．

21 世紀はポストゲノム時代の本格的幕開けを迎えるが，情報技術を駆使した遺伝子の配列，構造および機能解析が進むことと呼応して，ますますコンピュータおよび情報セキュリティ技術との連携が重要な研究課題となると思われる．本論文はその第 1 歩を示すものである．

参 考 文 献

1) 金久 實：ゲノム情報への招待，pp.13–15, 共立出版 (1996).
2) 小長谷明彦：遺伝子とコンピュータ，pp.29–34, 67–68, 共立出版 (2000).
3) Fisher, C.: Mitochondrial DNA, Today and Tomorrow, *11th International Symposium of Human Identification*, Speaker Abstracts, p.3 (2000).
4) Sprecher, C., Krenke, B., Amiott, B., Rabbach, D. and Grooms, K.: The PowerPlex 16 System, *Profiles in DNA*, Vol.4, No.1, pp.3–6 (2000).
5) Brown, T.A.（著），村松正實（訳）：ゲノム，p.154, メディカル・サイエンス・インターナショナル (2000).
6) 厚生科学審議会先端医療技術評価部会：遺伝子解析研究に付随する倫理的問題等に対応するための指針，厚生省 (2000).
7) 3 省関係者会議：ヒトゲノム・遺伝子解析研究に関する倫理指針，文部科学省，厚生労働省，経済産業省 (2001).
8) 辻井重男：東工大 黒澤馨教授への私信 (1999).
9) 辻井重男：第 3 回 FAIT 打合せ資料 (1999).
10) 辻井重男，板倉征男，山口 浩，北沢 敦，齋藤真也，笠原正雄：生体情報が秘密鍵に埋め込まれた構造を有する公開鍵暗号方式，信学技報，SCIS2000, D07 (2000).
11) 板倉征男，岩田 哲，尾形わかは，黒澤 馨，辻井重男：DNA 情報を組込んだ公開鍵暗号方式，信学技報，ISEC2000-42, pp.137–144 (2000).
12) Huckenbeck, W., Kuntze, K. and Scheil, H.-G.: *The Distribution of the Human DNA-PCR Polymorphisms*, A Cooperation Project of Institute of Forensic Medicine Institute of Human Genetics and Anthropology Heinrich-Heine-University Dusseldorf, Germany (2000).
13) 板倉征男，長嶋登志夫，辻井重男：個人識別用 DNA 情報の統計的検証，コンピュータセキュリティシンポジウム 2000，pp.121–126 (2000).
14) 岡本龍明，山本博資：現代暗号，p.118, 産業図書 (1997).

付録 DNA-ID の同値確率 P_M の推定

個人識別用 ID（α_A）の識別確率が，STR 座位（これをローカスという）の組合せ数 n とどのような関係にあるかを推定する．

ここではローカス間で分布に個人に依存する相関がないと仮定する．i 番目のローカスにおいて，繰返し回数 j が出現する確率を P_{ij} とすると，A，B の 2 人が

このローカスで同値となって識別できない確率〔以下同値確率（Matching Probability）と呼ぶ〕は，${P_{ij}}^2$ である．

A はあらゆる j の値をとる可能性があるから，A，B 2 人が同値となって識別できない確率，すなわち i 番目のローカスにおける同値確率 P_{M_i} はすべての可能な j について和をとればよい．

$$P_{M_i} = \sum_{j=1}^{m} {P_{ij}}^2$$

ここで m は各ローカスにおける繰返し回数の上限値で通常 60 程度である．

ローカス i を複数個組み合わせる場合，全体の同値確率 P_M は，ローカス間の個人に依存する相関がないことを仮定すれば，個々のローカスにおける同値確率の積となるので

$$P_M = \prod_{i=1}^{n} P_{M_i}$$

となる．

したがって，ローカスを n 個組み合わせて生成される個人識別用 ID の識別確率 P_r は

$$P_r = 1 - P_M = 1 - \prod_{i=1}^{n} P_{M_i}$$

となる．

本文図 4 にローカスの組合せ数 n と上式で計算した同値確率 P_M の関係を示す．なお，実際の計算に使用した P_{ij} は，ヨーロッパ系およびアジア系の人種の実データ，すなわち各人種ごとの各ローカスにおける繰返し回数の出現頻度の統計データを使った．

$n = 15$ のとき，$P_M \fallingdotseq 10^{-18}$〜$10^{-21}$ のオーダの同値確率が得られる．この値は現状において実用化された鑑識技術の最先端のレベルのものである．情報セキュリティ分野においては，さらに高度な識別確率を必要とする場合があるが，その識別確率を実現するのに必要なローカスの組合せ数を決めることができる．

(平成 12 年 11 月 29 日受付)
(平成 13 年 6 月 19 日採録)

板倉 征男（正会員）

昭和 39 年東京工業大学工学部電子工学科卒業，41 年同大学院修士課程修了．同年日本電信電話公社入社，（株）NTT データを経て現在 NTT データテクノロジ（株）勤務，データ通信システムの開発およびサービス事業に従事．中央大学研究開発機構客員研究員．電子情報通信学会会員．

辻井 重男（正会員）

昭和 33 年東京工業大学工学部電子工学科卒業．中央大学教授，研究開発機構長．東京工業大学名誉教授．電子情報通信学会会長等歴任．郵政省電波管理審議会委員．著書「暗号—ポストモダンの情報セキュリティ」（講談社選書メチエ），「暗号と情報社会」（文春新書）等．