DNAバイオメトリックス本人認証方式の提案

板倉征男[†,††] 長嶋登志夫[†] 辻井重男[††]

個人識別 ID のために用いる DNA 情報としては，全塩基配列のなかで STR（Short Tandem Repeat）と呼ばれる数塩基の繰返し回数の個人差を用いることが考えられる．筆者らは STR 座位（ローカスという）を複数箇所指定しそこで得られる繰返し回数情報を一定の順序で並べて個人識別子（以下 DNA 個人 ID と呼ぶ）を生成することを提案し，実用化のための数々の基本的考察を行った．本論文ではこの DNA 個人 ID の原理を用いたバイオメトリックス本人認証およびバイオメトリックス署名について実用的システムを提案する．また，提案の方式を検証するために実証実験を行った．すなわち，500 人以上の提供者の協力を得て実際の人体の DNA を採取し，本方式によりバイオメトリックス本人認証が可能であることを検証した．実用化のために，リアルタイムによる DNA 分析装置の開発が条件となるが，本装置の実現までの間は 2 枚の IC カードを用いて認証を行う方式を考案した．

Biometric Personal Authentication Using DNA Data

Yukio Itakura,[†,††] Toshio Nagashima[†] and Shigeo Tsujii[††]

Biometric verification/identification, which seeks to identify individuals accurately using biological information obtained from them, has attracted increasing attention. Research on applications of this method has progressed from a variety of angles. This paper focuses on DNA data that can produce unique digital information for the purpose of personal identification. It discusses how to collect that information and describes the procedures for processing it to generate identification (ID) data. Based on statistical theory, this paper demonstrates that such information can be applied adequately to personal identification. In addition, the paper proposes a biometric personal verification/identification and digital signature system, and describes its implimentation overall features. In this paper we build a public key encryption method that incorporates DNA data into a secret key and authenticates individuals according to the public key encryption scheme.

1. はじめに

バイオメトリックス認証技術は，近年急速に研究実用化が進んでいる．指紋および虹彩によるものはすでに商品化され，オンサイトにおける正確な本人認証を必要とするシステムに適用されつつある．

このほか網膜，顔貌，声紋，筆跡などによるものがあるが，いずれもアナログ量のパターンマッチングや特徴点比較が基本原理であるため，万国共通の普遍的で絶対的な識別機能を持つシステムの実現には至らず，ローカルなレベル，たとえば端末における相対的な本人認証に主として使われている[1)]．

一方生体情報の中で DNA 情報はその採取・分析が難しく，プライバシの問題もあるので，これまでバイオメトリックス認証の要素としては鑑定などのごく限られた用途に限られていた[2)]．

しかし DNA 情報は，原理的にデジタル情報であり個人差の著しい部分，たとえば STR（Short Tandem Repeat）と呼ばれる数個の塩基配列の繰返し回数の採取箇所を多重化すれば容易に識別精度を上げることができるので，これを個人識別子（DNA 個人 ID と呼ぶ）として応用すれば，これまでの指紋や虹彩などによるものでは得られない高度な機能を有する新たなバイオメトリックス認証方式が考えられる．

本論文では筆者らが提案する DNA 個人 ID を用いた DNA バイオメトリックス本人認証および署名方式について提案し，実証実験を行った結果を述べる．

なお本認証方式の原理を用いて個人のプライバシ情報を管理する社会システムの例を，先に文献 3) で提案しているが，本論文では机上検討段階にあった認証方式に関して実証実験を行い，DNA バイオメトリックス認証方式そのものの実用性の確認を行ったことを

† NTT データテクノロジ株式会社
NTT DATA Technology Corporation
†† 中央大学研究開発機構
Research and Development Initiative, Chuo University

属性＼生体情報	DNA 情報	従来方式
識別子の情報元（属性，データ長）	DNA の STR 繰返し数の個人差（デジタル情報，20 バイト）	指紋，虹彩，網膜，音声等のパターンの個人差（アナログ情報，250〜1500 バイト）
識別アルゴリズム	数値（DNA 個人 ID）比較	パターンマッチング
識別精度	$\sim 10^{-10}$（STR30 多重のときの実用的同値識別率）	$\sim 10^{-6}$（指紋の場合の実用的他人受容率）
識別時間	3時間〜（現状ではリアルタイム処理が不可）	〜数秒（オンサイトによるリアルタイム処理可）

図 1 DNA 情報を用いたバイオメトリックスの特徴
Fig. 1 Characteristic of DNA biometric information.

報告し，その延長として新たに DNA バイオメトリックス署名方式を提案するものである．

2. DNA 情報の採取と個人識別子の生成

2.1 DNA 情報の特徴と応用の動機

図 1 はバイオメトリックス認証から見た DNA 情報の特徴を従来の指紋などと比較したものである．

バイオメトリックス本人認証に DNA 情報を応用する動機は，それが次の 3 つの属性を有することにある．

(1) 情報元の精度と絶対性
DNA 情報から提案する方法で個人識別子（DNA 個人 ID）を生成すると ID が同値となる確率は，STR の多重度を上げれば指数的に減少する．$n = 30$ とすると，あらゆる組合せに対する実用的同値識別率は 10^{-10} 程度となり高精度の識別が可能である[4]．
しかも後述のような工夫をすれば DNA 個人 ID は本人と 1 : 1 の確定数値となるので，様々なバイオメトリックス応用システムを考えることができる．

(2) 生体情報特徴量としてのコンパクト性
後述する DNA 個人 ID は従来のバイオメトリックスシステムにおける生体識別用特徴量情報に相当する．従来 250 バイト（指紋の例）〜1,500 バイト（声紋の例）を要した特徴量情報に対して DNA 個人 ID は 20 バイトで十分である．これは DNA 個人 ID の元がデジタル情報であることの大きな利点である．

(3) 情報元の経年不変性と安定性
DNA 情報は人間の細胞すべてが同じ塩基配列，つまり同じ情報であり，一生不変とされている．また塩基そのものは A（アニン），G（グアニン），C（シトシン），T（チミン）の 4 種類の無機質の化合物であるが，これらは焼かれた骨から DNA が採取されるようにきわめて安定な物質である．これらは他のバイオメトリックスの弱点をカバーする優れた属性を有している．

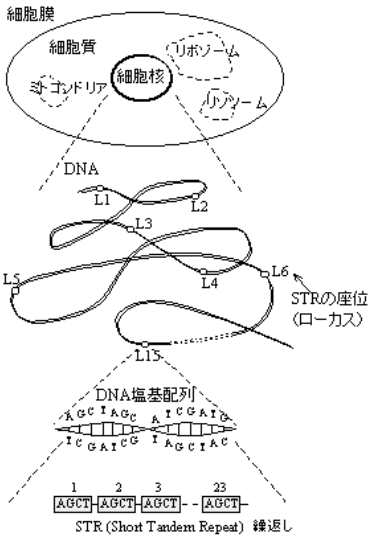


図 2 細胞の構造と DNA
Fig. 2 Cell structure and DNA.

また DNA バイオメトリックス本人認証に関する課題としては，次の 2 点が考えられる．

(1) 即時性および経済性
DNA 情報の分析に現状では 3 時間〜1 日の時間と高額な費用を要する．実用的なバイオメトリックスシステムの開発には約 2 ケタのスピードアップとコストダウンが必要である．

(2) プライバシの保護と倫理的配慮
本方式では DNA 情報のうち病因や人体構造に関係する部分，すなわち遺伝子領域に属する情報には無関係な STR の繰返し回数の個人差情報を扱うものである．しかし指紋と同様に個人の生体情報であるから，本人認証システムへの応用に際して十分なプライバシ保護と倫理的配慮が必要である．

2.2 細胞のしくみと DNA 個人 ID の生成方法

図 2 に細胞のしくみと DNA の概念を示す．人体は約 50 兆の細胞で作られており，その 1 つ 1 つの中にある細胞核に人体の設計図である DNA が畳み込まれている．

DNA は伸ばせば約 1.5 m になる 2 重のらせん構造をした帯に，4 種類の塩基が並んでおり，総計約 30 億個の塩基が示す情報を持っている[5]．

DNA のらせん構造の塩基配列の中で，STR と呼ば

れる箇所（ローカスと呼ぶ）は短い塩基配列の繰返しの対となっており，その回数 (j,k) は個人で著しい差違がある．

なお j および k の一方は父方から，他方は母方から受け継いだ情報である．

A 氏の DNA 個人 ID を α_A とすると，α_A は複数箇所定めたローカスから採取された STR 繰返し回数を示す 1 対の数値を次のように配列することにより生成する．

Step1： 各ローカスにおける STR 繰返し回数を計測する．

Step2： 各ローカスで得られた 2 つの値（STR 繰返し回数値）を小さい順に並べる．

$$L : j||k, \quad j \leq k$$

これを順序性操作ということにする．

Step3： DNA 個人 ID α_A は各ローカスの $L_i(j,k)$ を順に並べた配列で次のように生成する．

$$\alpha_A = L_1 \parallel L_2 \parallel L_3 \parallel \cdots \parallel L_n \qquad (1)$$

ただし，L_i はローカス i 番目の STR 繰返し回数の $(j,k)_i$ を示す．

生成された α_A は，後述するように一定の確率で一意性のある個人識別情報となる[4]．

2.3 DNA 個人 ID の統計理論的特性

2.3.1 DNA 個人 ID の同値確率[6),7)]

ローカス L における STR 繰返し回数 (j,k) がこの組合せで出現する確率を q_{jk} とする．また，j，k の個々の出現確率を q_j，q_k とすると各々は次のように表せる[4]．

$$\begin{aligned} q_{jk} &= q_j \cdot q_k + q_k \cdot q_j = 2q_j q_k \cdots j \neq k (j < k) \\ &= q_j \cdot q_j \qquad \cdots\cdots\cdots\cdots j = k \end{aligned}$$

任意の 2 人が 1 つのローカス i において STR 繰返し回数が同値となる確率 P_i は次のようになる．

$$P_i = \sum_{j=1}^{m} {q_j}^4 + \sum_{1 \leq j < k \leq m}^{m} 4(q_j \cdot q_k)^2 \qquad (2)$$

ここで m は j および k の上限値で，これまで報告された情報では $m = 60$ である．

ローカスを n 段重ねたとき，任意の 2 人が同値となる確率 P_n（いわゆる DNA 個人 ID の同値確率）は

$$P_n = \prod_{i=1}^{n} P_i \qquad (3)$$

となる．

STR ローカス間の相関がある場合相関係数を $\rho(\rho \ll 1)$ とすると，p_n は，次のように表せる．

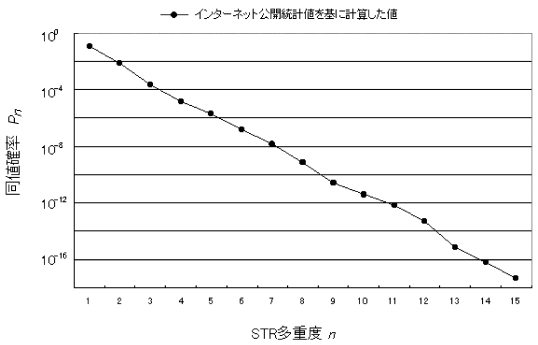


図 3 STR 多重度 n と DNA 個人 ID の同値確率 P_n の関係
Fig. 3 Relationship between STR multiplicity n and DNA personal ID matching probability P_n.

$$P_n \approx (1 + n\rho) \prod_{i=1}^{n} P_i \qquad (4)$$

ここで q_j または q_k については各国別の実測データがインターネットにより公開されている[8]．STR 繰返し回数の分布が幅広い数値に分布するほど P_i の値は小さくなる．この値は国により有意差があることが統計的に実証されているが，日本人のように歴史的に多くの血を交えていない国における P_i の値は大きくなる傾向があり，同値確率の観点での条件は悪くなる[9]．したがって日本人の統計値で識別能力を検証しておけば，万国共通に適用する場合に備えてより安全サイドで検証を行ったことになる．

インターネットによって公開されている q_j または q_k の日本人のデータを基に P_i を計算し，その P_i に基づいて DNA 個人 ID の同値確率を求めたものが，**図 3** に示すグラフである．

これによると，n と P_n の関係は統計的に

$$P_n \approx 10^{-n} \qquad (5)$$

で示されることが分かる．

一方ローカス間の相関については，実際の DNA 情報を採取分析した結果，式 (4) の $(1+n\rho)$ の値は STR の多重度 n に対して $n = 15$ のとき 1.78，$n = 30$ のとき 2.56 程度であり，P_n への影響は 1 桁以下であることが分かった[4]．

2.3.2 実用的同値識別率と必要な STR 多重度

前項で論じた DNA 個人 ID の同値確率 P_n は，任意の 2 人の間の ID が同じ値になる場合の確率である．

実際にこれを N 人の集団で ID として使用する場合には，あらゆる組合せにおいて ID が同値となる場合を考慮しなくてはならない．

すなわち STR 多重度を n，生成した DNA 個人 ID を使用する集団の人数を N とすると，N 人のあらゆ

n (STRローカス多重度)	N (適用人口)	P (実用的同値識別率)
15	10^5	$\frac{1}{2}10^{-5}$
21	10^7	$\frac{1}{2}10^{-7}$
24	10^8	$\frac{1}{2}10^{-8}$
30	10^{10}	$\frac{1}{2}10^{-10}$

図 4 STR 多重度 n と実用的同値識別率 P の関係
Fig. 4 Relationship between STR multiplicity n and practical matching value recognition rate P.

る組合せに対する同値確率 P（これを実用的同値識別率と呼ぶこととする）は，

$$P \approx \frac{1}{2}N(N-1)P_n \tag{6}$$

となる．ここで P_n は実証実験などから $P_n \approx 10^{-n}$ が得られている[4]．

N 人の集団で提案する DNA 個人 ID を使って情報セキュリティシステムを設計する場合，実用的同値識別率 P の値を $\approx \frac{1}{N}$ と設定すると，必要な STR 多重度 n は上記の実験式で求められる．(n, N, P) の関係について数例を示すと，**図 4** のようになる．

2.3.3 DNA 個人 ID が同値となる可能性

DNA 個人 ID は，実用的同値識別率以下の確率で発生する同値の DNA 個人 ID の間や一卵性双生児の間で同値となる可能性がある．

このような場合においても識別を可能とするために，何らかの対策が必要である．このため本方式では登録しようとする DNA 個人 ID を登録済の他人の数値と比較し，同値となった場合は 3.1 節の式 (8) 生成の過程で乱数を振り直して別の値になるような処置を行うこととしている．

2.4 DNA 個人 ID のマッピング

2.4.1 ハッシュ関数処理

生成された DNA 個人 ID α_A は，式 (1) のように DNA 情報をそのまま並べたものである．これにハッシュ関数処理を行い，その出力を δ_A とする．

$$\delta_A = h(\alpha_A) \tag{7}$$

ここで $h()$ としては汎用一方向性ハッシュ関数である SHA (Secure Hash Algorithm)-1 を適用する[10]．α_A は n(STR の多重度) $= 15$ の場合 10^{50} 程度，$n = 30$ の場合 10^{100} 程度のビット長の情報となる．SHA-1 の場合入力が 2^{64} ビット長未満のビット列で出力が 160 ビット長のメッセージダイジェストとなる一方向性関数演算を行うので，本方式に採用できるハッシュアルゴリズム関数である．なおこの $h()$ は公開する．公開することにより δ_A の普遍性が保証される．すなわち，同一人物の δ_A はどこで DNA を採取しても同じ値となる．

2.4.2 ハッシュ関数処理のメリット

ハッシュ関数を作用させるメリットとしては，個人のプライバシである α_A（DNA 個人 ID）を秘匿できることがあげられる．なお，DNA は他人の髪を毛根ごと採取するなど，人体の細胞の一部を盗むことにより，容易に他人の α_A を分析し，手に入れることができるので，たとえハッシュ関数処理を行っても秘密情報とすることはできない．したがって上記の $h(\alpha_A)$ の処理はもっぱら倫理目的のために行うものである．

3. DNA バイオメトリックス本人認証方式

3.1 DNA 個人 ID の暗号鍵への組込み

3.1.1 暗号鍵へ組み込む意義

DNA バイオメトリックス本人認証方式および後章のバイオメトリックス署名方式をシステム化するにあたり，基盤となる個人情報である DNA 個人 ID について，次の理由でこれをまず暗号鍵に組み込むことを提案する．

(1) プライバシの保護
DNA 個人 ID は，人間の基本的生体情報であり，プライバシ保護の考慮を第 1 義としてシステム化を論じなければならない．したがって生成した DNA 個人 ID はたとえハッシュ関数を通したものでも識別判定の情報としてそのまま適用することは，倫理上好ましくない．このため，DNA 個人 ID 情報が組み込んであることを証明することはできるが，DNA 個人 ID 情報そのものは秘密とすることができるようなしくみが必要である．本論文では δ_A を秘密鍵に組み込み，そのペアとなる公開鍵を生成して，それを CA（認証局）に登録する方式を提案する．バイオメトリックス本人認証は，DNA 個人 ID を生体情報で判別するのではなく，ここで提案する生体情報組込み型公開鍵が生成できるか否かで判別する方式で行う．

(2) 直接組込み可能な DNA 情報の属性
従来のバイオメトリックスの特徴量情報は，ID としては相対的な情報であり，かつデータ長も大きいので鍵に直接組み込むことは困難であっ

た．一方これに相当する DNA 個人 ID は個人識別のための絶対的な情報であり，かつデータ長も 8 バイトに圧縮できる属性を有するため，暗号鍵に直接組み込むことが実現可能となる．

(3) 生体情報 DB の割愛

暗号鍵に組み込むことにより，認証のためのリアルタイムの生体情報 DB を特別に構築する必要はなくなり，既存の PKI のしくみの中でバイオメトリックス本人認証を実現することが可能となる．

3.1.2 秘密鍵への組込みと公開鍵の生成方法[11)]

(1) 秘密鍵への組込み

ここでは公開鍵暗号方式における秘密鍵に DNA 個人 ID を組み込むことを述べる．

秘密鍵を X_A とする．X_A は 160 ビット長程度のビット列を考える．X_A は次の計算により生成する．

秘密鍵： X_A（160 ビット長程度のビット列）

$$X_A = \delta_A + r_A \tag{8}$$

ここで

δ_A**：** α_A（生体情報より生成した DNA 個人 ID）のハッシュ値をとった値．

r_A**：** 個人の秘密乱数．160 ビット長程度のビット列の乱数．登録用専用端末で生成する．

r_A を必要とする理由は，生体情報としての α_A は，他人の DNA を盗むことが可能なので，そのままでは秘密情報とはなりえないこと，したがって個人の秘密情報 r_A を加えて秘密鍵とする必要があるためである．

(2) 公開鍵への組込み

ここでは離散対数問題に基づく公開鍵暗号である ElGamal 暗号をとることとし，(1) で生成した秘密鍵から次のようにして公開鍵 Y_A を生成する．

公開鍵： Y_A

$$Y_A = g^{X_A} (\mathrm{mod\ p}) \tag{9}$$

（p は大きな素数，g は位数 p の原始元であり，システムに共用な値である）

認証局（CA）への登録： Y_A，g^{r_A} および個人情報を登録する．

生成した公開鍵は，他の関係情報とあわせて認証局（CA）に登録する．この際 CA は過去に登録済の CA を調べ，同一の公開鍵がある場合は r_A を生成し直して新たな Y_A を生成し登録し直すことを要求する．これにより 2.3.3 項で課題とした同一の DNA 個人 ID を有する者が現れた場合の課題が解決できる．

なお CA に g^{r_A} を登録する理由は，δ_A とは異なる δ_B なる DNA 個人 ID を持つ者が

$$X_A(=\delta_A + r_A) = \delta_B + r_B \tag{10}$$

となるような r_B を選び生成した公開鍵を Y_A と同一になるように見せかける不正を，本人認証の際にチェックできるようにするためである．

3.2 DNA バイオメトリックス本人認証方式

本人の個人情報と前項で述べた秘密鍵およびそのペアとなる公開鍵などを耐タンパー性のある IC カードに記録し，携行する．これをバイオメトリックス実印カードと名づける．

ここではこの IC カードの持主の正当性を検証する意味での本人認証方式について述べる．

本人認証を行うシステムは，生体情報（DNA）センサ装置を備え，本人の口腔を軽く擦った綿棒を入力情報としてカードの持ち主が本人であることをバイオメトリックス認証技術により確認する．すなわち綿棒に付着した口腔細胞の分析により DNA 個人 ID を生成し，これから本人の公開鍵が生成できるか否かをテストする．

今自分は A であると称する人物 A' から採取した $\alpha_A{}'$ にハッシュをかけた DNA 個人 ID を $\delta_A{}'$ とする．これは生体情報センサ端末からの生体情報入力となる．

次に CA（認証局）より登録済みの Y_A および g^{r_A} を取り寄せ，バイオメトリックス実印カードに記憶されている情報と一致することを照合し確認する．

このようにして CA と照合済の g^{r_A}，およびシステム共通情報である g，p を使い，上記 $\delta_A{}'$ から自分は A であると称する人物の公開鍵 $Y_A{}'$ を生成する．計算方法は以下のとおりである．

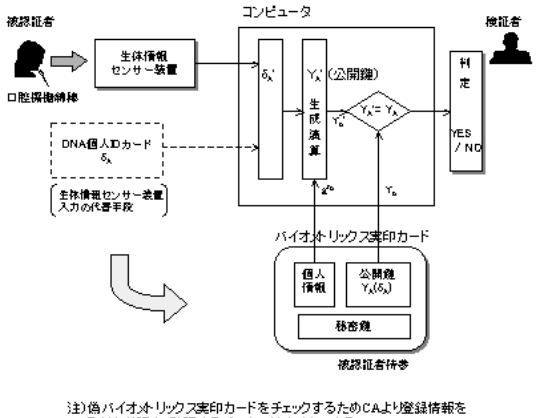
$$Y_A{}' = g^{X_A{}'} = g^{\delta_A{}' + r_A} = g^{\delta_A{}'} \cdot g^{r_A} (\mathrm{mod\ p}) \tag{11}$$

生成した $Y_A{}'$ がバイオメトリックス実印カードに記録されている Y_A と等しいか検証する．

$$Y_A = Y_A{}' \quad ? \tag{12}$$

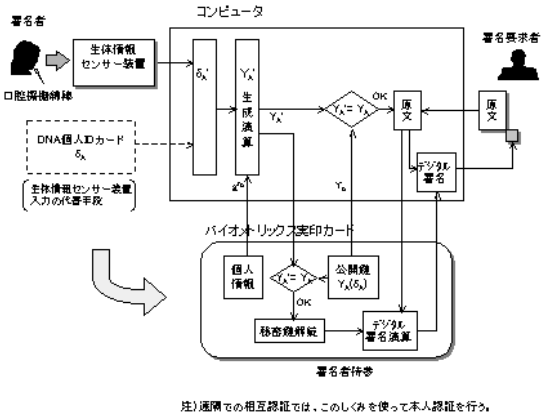
一致すれば A' は A 本人であることが認証される．もし A 以外の人物が偽って A を主張しても，自分の $\delta_A{}'$ で Y_A が生成できなければ本人になりすますことは不可能である．

現状技術では生体情報センサ装置は分析結果を出すのに 3 時間以上かかり，リアルタイムによるオンサイトチェックには使えない．このため自分の α_A から生成した $\delta_A = h(\alpha_A)$ を別の IC カード（DNA 個人 ID カードと名づける）に記録しておき，この 2 枚目のカードで生体情報センサ装置からの生体情報入力を代



注)偽バイオメトリックス実印カードをチェックするためCAより登録情報を取り寄せ照合・確認するプロセスは省略してある.

図 5 DNA バイオメトリックス本人認証方式の原理
Fig. 5 Principle of DNA biometrics-based personal authentication system.



注)通信での相互認証では、このしくみを使って本人認証を行う。

図 6 DNA バイオメトリックス・デジタル署名方式の原理
Fig. 6 Principle of DNA biometrics-based digital signature system.

行させる方式が考えられる．**図 5** に DNA バイオメトリックス本人認証方式の原理を示す．

以上は IC カードの持主の正当性を確認する方式である．これに対してネットワークを介した相互認証により本人確認を行う場合，すなわち遠隔での本人認証については次節で述べる DNA バイオメトリックス署名方式を使い，検証者からのチャレンジ電文にデジタル署名をつけてリターンする交信プロトコルにより実現する．

3.3 DNA バイオメトリックス署名方式

本方式におけるバイオメトリックス署名は，次のように行う．

すなわち「バイオメトリックス実印カード」の持ち主の正当性を前節と同じ手順で確認し，OK となるとカードに記録された秘密鍵が解錠され，「バイオメトリックス実印カード」にあるデジタル署名プログラムが入力された文書情報に署名をつける．署名が終わると秘密鍵の施錠を行う．IC カードの中にデジタル署名機能を持つので，秘密鍵の情報を外に漏らさずに署名ができる．

DNA バイオメトリックス署名では，自分の DNA 情報がデジタル署名に埋め込まれることになるので，血判を押印したような心理的効果も期待される．**図 6** に DNA バイオメトリックス・デジタル署名方式の原理を示す．

3.4 他と比べた DNA バイオメトリックス方式の特徴

DNA 個人 ID を適用した DNA バイオメトリックス・システムを構築する基本的検討を行ったが，これを指紋のような従来のバイオメトリックス識別子を適用する方法と比較すると，特徴が明白になる．

DNA 個人 ID は，測定分析装置の照合アルゴリズムに依存することのない普遍的なデジタル情報であることから，これを識別子とするバイオメトリックスシステムは従来の方式とは大幅に異なった斬新な機能を有することが期待される．他の方式と DNA 個人 ID を適用したバイオメトリックス・システムの比較を**図 7** に示す．

4. DNA バイオメトリックス・システムの具体的構成[12),13)]

4.1 DNA バイオメトリックス登録システム

本節では 3 章の方式を使って具体的に DNA バイオメトリックス・システムを構築する場合の具体例を提案する．

まず DNA の生体情報を登録するしくみが必要であるが，本方式では生体情報をそのまま扱うのではなく，暗号鍵に組み込む方法をとるため，**図 8** のような公開鍵暗号方式における秘密鍵と公開鍵の生成および CA（認証局）に公開鍵を登録するのと類似なしくみとなる．この装置を発行用専用端末と呼ぶこととする．

CA には公開鍵 Y_A のほかに暗号化された個人秘密乱数 g^{r_A} を合わせて登録する．この値はバイオメトリックス本人認証において，DNA 個人 ID δ_A から公開鍵 Y_A が生成できるか検証演算を行うとき必要なものである．

発行用専用端末で登録用に生成した δ_A，X_A，Y_A，g^{r_A} およびキーボードから入力した氏名，生年月日などの個人情報は，バイオメトリックス実印カード（図

方式 / 項目	従来の方式(指紋の例)	本方式の場合(DNA 個人 ID の例)	本方式の特徴
個人認証の原理	指紋隆起模様のマニューシャ(特徴点)の個人差	DNA STR の繰返し数の個人差	・本質的に個人差のあるデジタル情報をベースとする。
登録処理	特徴点抽出 ↓ 生体情報(テンプレート)DB(登録) これとは独立に秘密鍵生成(個人秘密乱数) ↓ 公開鍵(登録)	STR 分析→多種 STR 配列法による DNA 個人 ID 生成(α_A) ↓ ハッシュ関数処理 $h(\alpha_A) \rightarrow \delta_A$ ↓ 秘密鍵生成 $x_A = \delta_A + r_A$ (DNA 個人 ID＋個人秘密乱数) ↓ 公開鍵生成 $Y_A = g^{x_A} = g^{\delta_A + r_A}$ ● Y_A, g^{r_A}, g, p(大きな素数)を CA に登録	・暗号鍵に生体情報を埋め込む (個人識別子の精度が高く、本質的にデジタル情報であるため秘密鍵/公開鍵に生体情報を組込むことが可能。) ・セキュアな生体情報の管理が可能 (生体情報は公開鍵に埋め込んだ形で登録するが、登録情報から生体情報を計算することは困難。)
照合アルゴリズム	メーカ固有の照合アルゴリズム	共通の照合アルゴリズム($Y_A = Y_A'$)	・パターンマッチングアルゴリズムが不要 ・照合方式は世界標準化容易 (多重ローカス位置の定義を行うだけでよい。)
認証処理	1. X.509 証明書の正当性検証 2. テンプレートの正当性検証 3. テンプレート :バイオメトリックス技術を用いた本人認証	1. 同左 2. Y_A, g^{r_A} の正当性検証 ●CA の署名の正当性確認 3. バイオメトリックス技術を用いた本人認証 ● $Y_A = Y_A'$ の判定	・認証処理のあいまいさが無い (照合アルゴリズムにおいて本人の DNA 個人 ID(α_A')から生成する公開鍵 Y_A'が、IC カードの中に登録された Y_Aに一致するか否かの単純なロジックで Yes/No 判定が行われる。)
デジタル署名	1. バイオメトリックス(指紋)認証による本人確認 2. 秘密鍵解錠 3. デジタル署名処理	1. バイオメトリックス(DNA)認証による本人確認 2. 同左 3. 同左	・血判効果がある (秘密鍵で署名された文書に生体情報が埋め込まれていることになるので、血判で署名するような心理効果がある。)
精度	$< 10^{-6}$ (実用的本人拒否率／他人受入率)	$10^{-15} \sim 10^{-30}$ (多重 STR ローカス数 15〜30 の場合の同値確率)	・ローカス数の増加により容易に精度の向上が可能
認証に要する時間	数秒 (リアルタイム処理実用化済)	3 時間 (オンサイトによるリアルタイム処理が現状では不可)	・リアルタイム認証及びそのコストダウンが本方式の今後の課題

図 7 他方式と比較した DNA バイオメトリックス・システムの特徴
Fig. 7 Characteristic of DNA personal ID for biometrics-based authentication.

中のカード 2）に書き出し記録する．この IC カードは耐タンパー性のある構造であることが必要である．

バイオメトリックス実印カードは，これらの情報のほかに Y_A の生成・比較演算およびデジタル署名演算機能を有するプログラムを内蔵する．前者はデジタル署名時に秘密鍵を解錠するために，また後者はデジタル署名時に秘密鍵を外に漏らさないために，IC カード内に内蔵するプログラムである．

生体情報センサ装置の代替または補完機能として，DNA 個人 ID カード（図中のカード 1）を併用する場合は，これに δ_A および個人情報を書き出し記憶する．

4.2 DNA バイオメトリックス認証システム

DNA 情報を用いたバイオメトリックス本人認証およびバイオメトリックス署名を行うシステムの構成例を**図 9** に示す．

ここではクライアント PC にこれらの 2 つの処理機能を持たせ，アプリケーションサーバにあるアプリケーションプログラムからこれらの機能をクライアント PC に要求し，処理を行わせる設計としている．

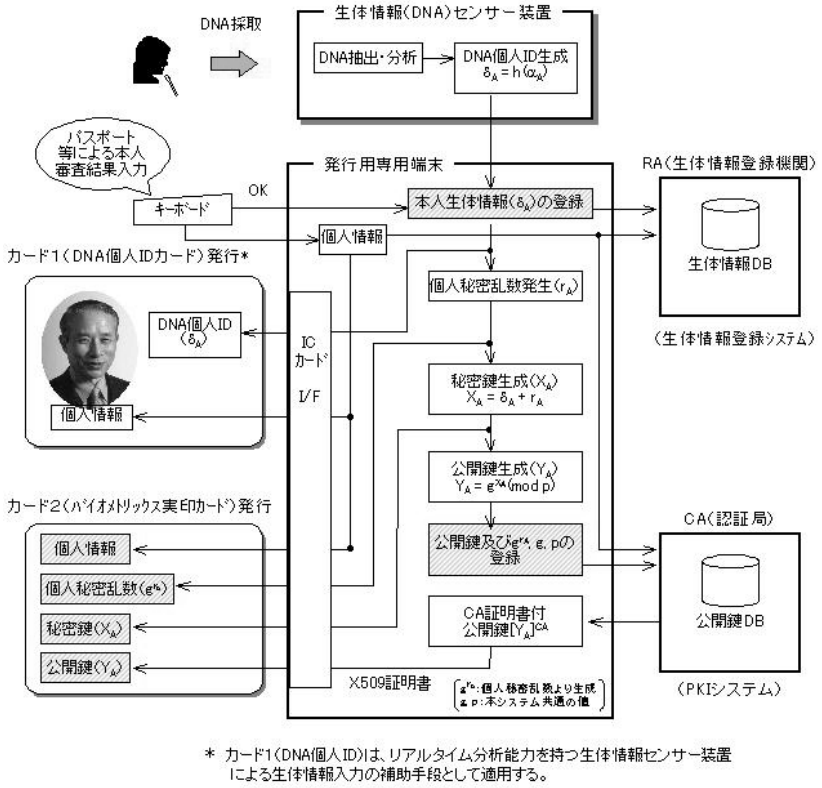


図 8 DNA 情報を用いたバイオメトリックス登録システム
Fig. 8 Biometrics-based registration system using DNA information.

5. DNA バイオメトリックス・システムの実証実験

5.1 実験システムの構築

本方式を検証するために，実験システムを構築した．ただし生体情報センサ装置は，現在の技術ではリアルタイム処理が不可能であるために仮想のものとし，実際の生体情報（DNA）の分析は既存の法医学鑑定に用いる分析機械を用いてオフラインで行った[14]．

また実験システムではバイオメトリックス実印カード（カード 2）の機能のうち公開鍵生成処理とデジタル署名処理のプログラムを除く機能を 1 枚の IC カードにインプリメントして実証実験を行った．

実験の第 1 ステップとして，本人認証機能のテストを行うこととし，1,000 人分の PKI（公開鍵基盤）機能を有する CA を実現した．

実験システムの概要は次のとおりである．

- 認証システムクライアント PC
  機種：Endeavor MT-4000（エプソン製）
  OS：Windows 2000，NT-WS4.0
  Java 開発環境・実行環境：JDK1.2.2_007
  実験用プログラム開発規模：8.4 Kstep
- CA（認証局）サーバ PC
  機種：EdicubeTC730（エプソン製）
  OS：FreeBSD 4.2
  Java 開発環境・実行環境：JDK1.2.2_006
  データベース：Postgre SQL
  実験用プログラム開発規模：4.9 Kstep
- IC カード
  BULL 製　Java Card
  大日本印刷製　Standard-9

5.2 実証実験の方法

本研究では，中央大学研究開発機構と東北大学大学院医学系研究科との共同研究協定を結ぶことにより，515 人の DNA 提供者の協力を得て DNA 個人 ID の統計的検証とバイオメトリックス本人認証システムの実用化の可能性について実証実験を行った．

ただし前述のとおり，個人のプライバシ情報を扱うことになるので，倫理的手続きについては十分注意した．すなわち本研究内容については 2000 年度に出された厚生省の「遺伝子解析研究に付随する倫理的問題に対応するための指針」[15] に基づき，東北大学に設置

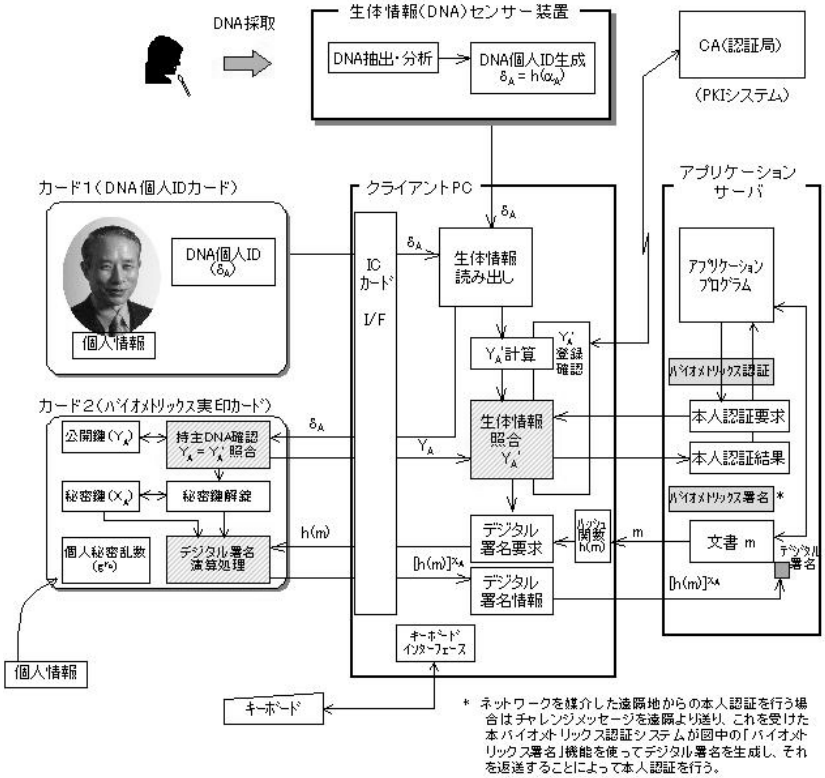


図 9 DNA バイオメトリックス認証システム
Fig. 9 Biometrics-based authentication system using DNA information.

された倫理委員会に提議し，その審査を受けた．規定により，DNA 提供者に十分なインフォームド・コンセントと同意書の署名を得て採取および分析を行った．

STR 解析で得られたデータから本提案による生成方法で DNA 個人 ID を生成し，検証を行った．

実証実験によって得られたデータは次のような手順で解析を行い，検証を行った．

採取し生成した 515 人の DNA 個人 ID について同値確率を調べた．ローカス 15 多重をもとに生成した DNA 個人 ID の同値のものの存在は，すべての組合せにおいて 0 となった．

次にローカスの多重度を 15 から順次減らしていき，それに応じた DNA 個人 ID の同値率を調べていくと，ローカスの多重度が 5 のときはじめて同値のものが表れ，以小の多重度のものについては**図 10** のような同値の DNA 個人 ID が存在することが実証された．

実証実験に基づく同値確率の STR 多重度とインターネット公開データに基づくそれとはよく近似しており，式 (5) および図 3 の正当性が検証できた．これにより DNA バイオメトリックス本人認証は，図 3 のような精度で判別可能なことが検証できた．

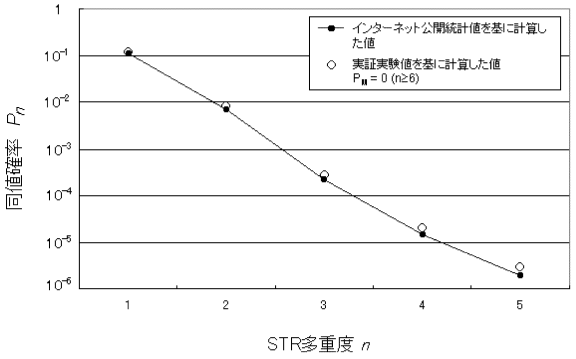


図 10 採取した DNA による実証実験結果
(STR ローカス多重度と DNA 個人 ID の同値確率の関係)
Fig. 10 Result of a validation experiment using actual DNA collected.
(Relations between STR Locus multiplicity n: DNA personal ID matching probability P_n).

6. 考　　察

6.1 分析時間の課題

DNA 個人 ID を適用してバイオメトリックス認証を行う場合，生体情報採取後の実用的な分析時間は数

秒のオーダによる処理が要求される．現状の法医学分野で使用される分析機材では，最新のものでも 3 時間以上を要しているので，情報セキュリティシステムで実用化するには，革新的な技術によるブレークスルーが必要である．

上記のようなオンサイト・リアルタイム分析が不可能な現状では，要求仕様を満たす端末が開発されるまでの代替手段として，δ_A（DNA 個人 ID α_A にハッシュをかけた情報）を専用の IC カードに記憶し，これを使って本人の生体情報を入力する方法を 3.2 節で提案した．

法廷などでバイオメトリックス（DNA）による厳密な本人認証を行う場合は，時間をかけて生体情報を実際に分析するという担保を有することがポイントである．

6.2 なりすまし登録防止の課題

本提案では DNA バイオメトリックス登録システムに登録する際，パスポートや免許証で十分本人確認を行い，不正はない，すなわちなりすまし登録はできないという前提でシステムを考えた．しかし現実には写真の貼替えなどによって他人になりすます不正登録の可能性は存在する．これに関してはバイオメトリックス各方式に共通するもので，本提案の DNA バイオメトリックス認証方式でも解決されていない課題である．

一方本論文では述べていないが，DNA 親子鑑定のアルゴリズムを用いて絶対的な本人確認を行う方式が考えられる．この方式は DNA 情報を用いた独自のなりすまし防止確認機能として新たな有用性が期待されるので，今後引き続き追究することとする．

6.3 倫理的課題

本提案による DNA 個人 ID は，DNA 情報のうち人体の構造や病因に関与しない，いわゆる遺伝子領域以外の部分でマイクロサテライトといわれる STR 情報を用いるので，個人の秘密情報には関与するものではない．しかしながら指紋のように個人識別が可能な本人固有な情報を取り扱うので，プライバシ保護について十分考慮する必要がある．本提案ではプライバシ対策として次の 2 つを考えた．

その 1 つは DNA 個人 ID にハッシュ関数処理を行い，一方向性のマッピングを行った情報を識別子（δ_A）として使用すること，その 2 つは DNA 個人 ID を秘密鍵およびそのペアとなる公開鍵に組み込んでしまい，以降は公開鍵暗号方式の機能を使って署名や認証を行う．

暗号鍵に組み込まれた生体情報が自分のものか否かは，その情報を直接相手に示さなくても 3.2 節に述べた方法で証明することができる．

また総合的見地からみると，指紋と同様生体情報を個人番号と同等に扱い，そのような個人情報で人間を管理することの是非について討議する必要がある．

テロリズムが人間の安全な生活を脅かしている昨今において，適切な倫理法の下で本方式が究極の個人認証システムとして正しい運用が行われれば，21 世紀の新しい情報セキュリティシステムとして威力を発揮することが期待できる．

7. ま と め

本論文は DNA 個人 ID を応用して情報セキュリティシステムにおけるバイオメトリックス認証およびバイオメトリックス署名を実現するための基本的課題について検討した．

特にプライバシ保護の観点からハッシュ関数処理を行うことや，暗号鍵に組み込むことを考察し，課題解決のための方法を提案した．また提案方式を検証するため 500 人以上の DNA 提供者の協力を得て実証実験を行った．

その結果，STR のローカスの多重度 n を増やせば DNA 個人 ID の同値確率を指数的に下げることができ，かつローカス間の情報の相関性も無視できることが確認できた．これにより提案する DNA バイオメトリックス認証システムは世界的規模の人口に対して，高精度で普遍的な識別能力を持つ可能性が検証できた．

最後に倫理的課題と社会システムとして討議すべき問題提起を行った．

参 考 文 献

1) 菅知之ほか：特集 ここまできたバイオメトリクスによる本人認証システム，情報処理，Vol.40, No.11, pp.1072–1103 (1999).
2) Holt, C., et al.: Practical applications of genotypic surveys for forensic STR testing, *Forensic Science International*, Vol.112, pp.91–109 (2000).
3) 板倉征男，辻井重男：DNA-ID を用いた DNA 個人情報管理システムの提案，情報処理学会論文誌「21 世紀のコンピュータセキュリティ技術」，Vol.42, No.8, pp.2134–2143 (2001).
4) 板倉征男，橋谷田真樹，長嶋登志夫，辻井重男：DNA-ID の統計的検証，信学技報，Vol.101, No.214, pp.1–7, ISEC2001-19 (2001).
5) Brown, T.A.：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) Guo, S. and Thomson, E.: Performing the exact test of Hardy-Weinberg Proportion for multiple alleles, *Biometrics*, pp.361–372 (1992).

7) Weir, B.: Independence of VNTR alleles defined by fixed bins, *Genetics*, pp.873–887 (1992).
8) Huckenbeck, W., Kuntze, K. and Scheil, H.: *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. http://www.uni-duesseldorf.de/WWW/MedFak/Serology/database.html
9) 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，情報処理学会CSS-2000シンポジウム論文集，pp.121–126 (2000).
10) 岡本龍明，山本博資：現代暗号，pp.189–195, 産業図書 (1997).
11) 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，電子情報通信学会シンポジウム予稿集，SCIS2000 (2000).
12) 日本自動認識システム協会（編）：これでわかったバイオメトリックス，pp.94–102, 119–126, オーム社 (2001).
13) Isobe, Y., Seto, Y. and Kataoka, M.: Development of Personal Authentication System Using Fignerprint with Digital Signature Technologies, *IEEE Proc. 34th Hawaii International Conference on System Science* (2001).
14) Walsh, P., Metzger, D. and Higuchi, R.: Chelex100 as medium for simple extraction of DNA for PCR based typing from forensic material, *Biotechniques*, pp.506–513 (1991).
15) 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).

(平成13年11月22日受付)
(平成14年 6 月 4 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コース卒業．昭和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，データー通信システムの開発および運用に従事．中央大学研究開発機構客員研究員．工学博士（平成14年）．電子通信学会会員．

長嶋登志夫（正会員）

昭和52年東京理科大学理工学部物理学科卒業．昭和54年同大学院修士課程修了．同年中央大学物理学教室技術員．平成10年NTTデータテクノロジ（株）入社，現在に至る．電子政府システムの開発および情報セキュリティの研究に従事．

辻井 重男（正会員）

昭和33年東京工業大学工学部電気工学コース卒業．中央大学教授，東京工業大学名誉教授．工学博士．電子情報通信学会会長等歴任．総務省電波管理審議会会長．著書『暗号—ポストモダンの情報セキュリティ』（講談社メチエ選書），『暗号と情報社会』（文藝春秋社）等．