

Rapports et communiqués

Rapport adopté le 20 novembre 2007

Identification des personnes par des analyses biométriques et génétiques.

Bull. Acad. Natle Méd., 2007, 191, no 8, p 1779-1782

RAPPORT

au nom d'un groupe de travail issu de la Commission I (Biologie)

Identification des personnes par des analyses biométriques et génétiques

Mots-clés (Index medicus) : Biométrie. Dépistage génétique.

Emmanuel-Alain CABANIS**, Jean-Yves LE GALL*** et Raymond ARDAILLOU***

RÉSUMÉ

L'identification des personnes entre de plus en plus dans les nécessités de la vie moderne. Identifier, c'est relier des caractères personnels à ceux préalablement rassemblés dans une banque de données. Authentifier, c'est décider si un individu est bien celui qu'il prétend être. Ces deux objectifs sont assurés en analysant les données biométriques et les empreintes génétiques. Toutes les techniques biométriques se déroulent en plusieurs étapes : saisir l'image ou les paramètres physiques, les coder selon un modèle mathématique, comparer ce modèle à ceux contenus dans la base et calculer le risque d'erreur. Ces techniques doivent s'appliquer de façon universelle et recueillir des données spécifiques et permanentes d'un individu. Les plus utilisées sont la reconnaissance faciale, les empreintes digitales (plis de flexion et dermatoglyphes qui ont l'avantage de laisser des traces), la surface et la texture de l'iris. D'autres techniques biométriques analysent des comportements : voix, signature, frappe sur un clavier, démarche. Enfin, on peut être identifié par radio-émetteur implantable. Tous ces systèmes sont caractérisés par les mêmes paramètres : taux de faux rejets, taux de fausse acceptation, taux d'impossibilité d'enregistrement. Les utilisations de la biométrie sont multiples et en constante extension : systèmes d'identification nationaux et internationaux, contrôle d'accès dans des lieux protégés, identification des délinquants et de leurs victimes, sécurité des transactions. La génétique identifie sans risque d'erreur les individus par leurs empreintes génétiques. Il s'agit de courtes séquences de 2 à 5 nucléotides répétées un nombre variable de fois et appelées microsatellites. Les troupes les plus courantes permettent d'analyser 11 à 16 marqueurs autosomiques indépendants. Plus rarement, on analyse l'ADN mitochondrial et celui du chromosome Y. Ces tests génétiques sont utilisés pour identifier des suspects ou des victimes à partir des tissus biologiques retrouvés ou des cadavres et pour établir des liens de filiation. L'identification des personnes pose de nombreux problèmes: dans quels cas constituer et comment exploiter une banque de données selon les instructions de la Commission Nationale de l'Informatique et des Libertés (CNIL) ? Comment réglementer l'accès aux données génétiques ? Le polymorphisme génétique permet-il d'identifier des races parmi les humains ? Jusqu'où doit aller l'interopérabilité avec le foisonnement des fichiers ? Quelles limites apporter aux fiches individuelles consultables sur Internet et à l'identification par radiofréquences à partir de puces implantées ? L'équilibre doit être troublé entre la nécessité de garantir la sécurité des individus et des transactions et, d'autre part, l'autre nécessité de protéger les libertés individuelles et le secret de la vie privée.

INTRODUCTION

Identifier une personne, c'est comparer ses caractères individuels uniques à ceux préalablement rassemblés dans une base de données. Cette définition s'est appliquée dès le début de l'histoire humaine, la banque de données étant alors la mémoire. C'est ainsi que la nourrice Eurycle reconnaît Ulysse, dans le mendiant débarqué à Ithaque parce qu'il porte à son genou la cicatrice d'une ancienne blessure. Depuis le 19^{ème} siècle, les français sont identifiés par leur état civil, d'abord inscrit sur les registres paroissiaux, puis sur ceux des mairies [1]. L'Institut National de la Statistique et des Études Économiques (INSEE) rassemble aujourd'hui ces données dans le Répertoire National d'Identification des Personnes Physiques (RNIPP). Ce répertoire comporte nom, prénoms, date et lieu de naissance, numéro d'acte de naissance, date et lieu du décès éventuel. Le code INSEE à 13 chiffres (1934) [1] indique le sexe, l'année de naissance, le mois de naissance, le département et la commune de naissance et le numéro d'inscription de la naissance au registre d'état civil. Ce code est utilisé par la plupart des services administratifs dont nous dépendons. Le fichier INSEE est régulièrement mis à jour et son accès est protégé. Rattacher une personne à son identité administrative impose d'analyser ses caractères anatomiques et/ou comportementaux et/ou génétiques, avec le minimum d'erreurs. Ces données sont comparées à d'autres, préalablement recueillies, et classées dans une base de données nominatives (analyse « 1 contre n »). La concordance entre la personne contrôlée et celle dont l'identité est stockée est alors vérifiée.

L'analyse métrique des caractères anatomiques définit la biométrie [2]. En plus d'identifier, la biométrie sert aussi à authentifier, c'est-à-dire à prouver que le porteur d'un document est bien celui qu'il affirme être (analyse « 1 contre 1 »). Les techniques et leur utilisation se développent sans cesse avec les progrès de l'informatique (algorithmes et programmes, électronique, réseaux et mémoires) et la diffusion d'Internet dans la population. S'y ajoutent le besoin de sécuriser les transactions, répondre à l'expansion du « e-commerce » immatériel, assumer la demande identitaire en filiation des personnes et, depuis le 11 septembre 2001, la hantise sécuritaire (lutte contre le terrorisme international et la criminalité financière) [1]. Le prix du progrès se paie par le risque d'appropriation de l'état civil d'autrui (fraude à l'identité donnant accès à des fichiers informatiques cryptés) qui augmente dans des proportions préoccupantes [3]. En santé publique, il est essentiel de connaître les caractères d'une population afin de mieux définir une politique de prévention et de soins. Pour la santé du patient, la carte « Vitale » utilisée en France va s'enrichir de nouvelles données (photographie faciale) et, peut être, du dossier médical numérique du patient. Cette explosion dans l'utilisation des techniques biométriques et des fichiers exige de tenter de la réglementer. Sinon, elle portera atteinte aux libertés individuelles, en ne préservant plus le secret de l'individu et de sa vie privée. D'où le contrôle exercé, en France, par la Commission Nationale de l'Informatique et des Libertés (CNIL) et le souci du Législateur de réfléchir à la question [4]. Ainsi, l'Office Parlementaire d'Évaluation des Choix Scientifiques et Technologiques (OPECST) a rédigé un rapport édité par C. Cabal [5, 6].

L'Académie Nationale de Médecine a souhaité compléter ce rapport, limité aux techniques biométriques, en l'élargissant aux

tests génétiques, tout en détaillant les données proprement médicales de l'identification des personnes. Comme le Comité Consultatif National d'Éthique (CCNE) le souligne dans son avis d'avril 2007 [7], l'identité de la personne ne se résume pas à des caractères anatomiques. Mais capacités cognitives et conscience que l'on a de soi sont difficilement mesurables, échappent à notre étude et sont indépendants des paramètres biométriques et génétiques.

Techniques biométriques d'identification et D'authentification

La biométrie regroupe l'ensemble des techniques d'acquisition et d'exploitation statistique des résultats des mesures faites chez les êtres vivants. La biométrie humaine est née au 19^{ème} siècle avec Paul Broca, fondateur de l'École d'anthropologie physique ou anthropométrie [8]. Elle quantifie les dimensions d'éléments du squelette, dont le crâne, à partir des points dits « anthropologiques » et mesure distances et indices, angles et plans d'orientation. Ensuite viennent les caractères anatomiques externes : taille en position debout (stature), formes et dimensions de la tête, de la face, du nez, des oreilles et des lèvres, longueurs et indices des membres. Déçu par les approximations du « signalement » facial né au 17^{ème} siècle, Alphonse Bertillon a découvert la supériorité de la biométrie et créé l'anthropométrie d'identification judiciaire des délinquants récidivants et des victimes (1875) [9]. Ajoutées au « Bertillonage », la photographie systématisée de la tête et les empreintes digitales furent à la base du premier service d' « Identité Judiciaire » du monde (Préfecture de Police de Paris, 1887) [10]. Dès son invention (1895), la radiologie élargit le champ de l'investigation anatomique in vivo avec de multiples valeurs et indices corporels. La biométrie devient « anatomie quantitative » avec André Delmas en 1969, puis numérique avec l'informatique des années 70-80, en aval des détecteurs qui remplacent le film photographique (scan RX puis IRM). La numérisation photographique des années 90-00 transforme en « gabarits » spécifiques l'anatomie de surface du visage, de l'iris et de la rétine, de la main, des empreintes digitales, ainsi que nombre de caractères comportementaux (voix, frappe d'un clavier, démarche ...) avec, en conséquence, une croissance exponentielle des bases de données.

Le processus d'« enrôlement » conduisant à l'identification ou l'authentification se déroule en trois étapes informatiques: 1- la saisie des formes (acquisition ou capture) par photographie numérique (caméra CCD, «*Coupled Charged Device* ») ou par capteur de paramètres physiques variables, tels que radiofréquences, ultrasons, températures, résistances, qui nécessite un appareillage de haute qualité permettant ainsi, une saisie correcte, même chez un individu en mouvement ; 2- le codage en modèle mathématique («*template* », gabarit ou signature), l'original, lui, n'étant pas enregistré ; 3- la comparaison de ce modèle à ceux présents dans la base et le calcul du risque d'erreur («*matching* »). Ces trois étapes bénéficient des progrès constants de l'électronique et de l'informatique (sophistication des capteurs, vitesse du calcul, extension de la mémoire, algorithmes mathématiques...). Trois critères principaux commandent le choix de l'organe et des caractères identifiants : 1- l'universalité (présence chez tout individu) ; 2- l'unicité (spécificité maximum d'un individu à l'autre) ; 3- la permanence (avec le temps et l'avancée en âge). Quatre critères secondaires sont l'acceptabilité par le sujet, plus facile à obtenir pour les méthodes non intrusives, la vitesse d'acquisition (donc le débit possible des individus soumis à vérification, par exemple dans les aéroports), l'interopérabilité d'un pays à l'autre et le coût.

Deux groupes de techniques se distinguent selon qu'elles relèvent de 1- l'anatomie statique, ne laissant pas de traces et captable à distance (analyse géométrique du visage, de l'iris, de la rétine si la coopération du sujet est obtenue, du réseau veineux superficiel du dos de la main) ou laissant des traces (dermatoglyphes de la main, empreintes des pavillons de l'oreille sur la literie) ou de 2- l'anatomie dynamique ou comportementale, nécessitant (frappe sur clavier) ou non (voix, démarche) la coopération du sujet

Analyse des caractères anatomiques, statiques et dynamiques

Reconnaissance faciale et photographie numérique du visage (Fig. 1).

Première reconnaissance de « l'autre » dans la mémoire, celle du visage et de sa photographie, complète l'anthropométrie de Bertillon dès 1840 et s'ajoute au « signalement ». Elle est la plus présente sur les multiples supports d'identification. La capture numérique est simple, possible à distance et peu coûteuse. En capture, les nouveaux dispositifs utilisent souvent le système Wi-Fi («*Wireless Fidelity* ») pour la transmission des images. La performance des caméras de vidéo-surveillance (capture de l'image à la volée à l'insu de l'intéressé) peut laisser à désirer car la qualité de l'image est sensible au mouvement, à l'éclairage et à l'orientation. En revanche, la qualité est bonne si on obtient la collaboration du sujet qui doit rester immobile, être bien éclairé et positionné de façon convenable. Les logiciels d'analyse morphologique, de type local (« déformable » examinant les yeux, le nez, le front, puis une suite de points caractéristiques) ou global (« vectoriel ») construisent le gabarit en 2D. Des systèmes 3D apparaissent et se substituent aux techniques connues : balayage par faisceau laser des reliefs du masque facial ou, dernier en date, succession rapide (quelques ms) de plusieurs flashes provoquant les ombres portées du nez, des pommettes et des lèvres... [12, 13].

Fig. 1 – Visage et reconnaissance de la face.

a b e
c d

a, b) Photographie ana-logique (b) comparée à l'image de capture numérique du visage (a) (doc. Expo. Musée de La Villette); c) Analyse vectorielle de l'image numérique ; d) Imagerie faciale ombrée, donc 3D, travail en cours [12 et 13] ; e) thermographie faciale.

La surface de l'iris et sa texture, la rétine et les autres techniques de biométrie céphalique(Fig 2).

La biométrie de la texture irienne en surface [14, 15] exploite sa photographie numérique. La technique est fiable car elle prend en compte 250 paramètres et points, environ. Les caractères distingués sont la forme de la collerette, le nombre et la localisation des cryptes de Fuchs et des tâches pigmentées, la forme des plis de contraction circulaire et radiaire. La caméra d'acquisition, monochrome, photographie les parties latérales de l'iris. Elles sont délimitées en rectangles, numérisées en

gabarits individuels et comparées à la base de données en utilisant, notamment, « Iriscan », un logiciel conçu dès 1987. Cette analyse biométrique a de nombreux avantages : 1- la texture irienne est permanente et spécifique allant jusqu'à différencier deux jumeaux monozygotes ; 2- la technique est fiable car analysant plus de points que celle des empreintes digitales ; 3- elle est rapide (10 s environ) et facile, si le sujet examiné accepte de coopérer ; 4- elle est peu intrusive, la photographie pouvant être prise « à la volée » à l'insu de l'individu ; 5- elle ne se prête pas à la fraude puisque prothèses oculaires et lentilles colorées sont facilement détectées. Deux inconvénients limitent la méthode : 1- une éventuelle pathologie oculaire congénitale ou acquise (hétérochromie de Fuchs, mélanome, délabrement irien post-traumatique, taie ou dystrophie cornéenne, désorganisation du segment antérieur après chirurgie de cataracte...), et 2- son coût, plus élevé que celui de la prise des empreintes digitales. La coloration de l'iris n'est pas un critère biométrique car elle évolue avec l'âge contrairement à la texture. Elle dépend du nombre de mélanosomes contenus dans les couches cellulaires. Établie vers l'âge de 6 ans, la couleur devient maximale à 15 ans, puis diminue ensuite progressivement. De plus, elle est susceptible d'être modifiée par des thérapeutiques, comme celle du glaucome par les prostaglandines. L'anatomie vasculaire de la rétine, individuelle et variant peu dans le temps, relève de l'ophtalmoscopie directe du fond d'œil, maintenant numérisée. Mais cet examen rétinien, intrusif, exige la coopération du sujet, dont la tête doit rester immobile et le regard fixe.

Fig. 2 – Reconnaissance de l'iris

a
b
c d

a) Variabilité individuelle des textures et couleurs iriennes, d'un sujet à l'autre ; b) anatomie de surface de l'iris (doc. M. Rigal) ; c) passage aéroport : numérisation par caméra ; d) les trois étapes du codage de l'image numérique irienne (de haut en bas, filtrage des reflets, délimitation du contour irien puis encodage du gabarit (doc. Sagem).

L'anatomie de surface des sillons du cortex cérébral, telle qu'explorée en imagerie par résonance magnétique (IRM) en 3D, présente une haute variabilité entre hémisphères droit et gauche chez un même sujet et entre les personnes. Cette identification, précise, appartient encore à la recherche. Il en va de même pour la thermographie faciale.

Empreintes des crêtes papillaires (dermatoglyphes), digitales et palmaires, et géométrie de la main (Fig.3 et 4).

L'examen des empreintes laissées par les dermatoglyphes de la main (lophoscopie), souvent limité aux seules empreintes digitales, est, avec le portrait, la plus ancienne des techniques morphologiques (Égypte, Chine). Les empreintes sont celles du réseau de crêtes et de sillons de l'épiderme des doigts et des mains, formé dès la 8^{me} semaine de la vie intra-utérine [16, 17]. Après capture de l'image, le logiciel en distingue les caractères ou « minuties » (fins de lignes ou bifurcations formant des arches, boucles et tourbillons). Selon le niveau de discrimination recherché, un nombre croissant de minuties, de 15 jusqu'à 85, est à la base du gabarit biométrique (logiciel « Minutiae » par exemple). Un taux de ressemblance entre 0 et 100% est déterminé, le contrôleur choisissant le seuil d'acceptation. Les gabarits d'empreintes digitales sont inter-opérables, entre équipements et pays. L'examen des plis de flexion complète l'analyse des dermatoglyphes digitaux. Ces structures sont stables et ne sont pas modifiées par l'âge. Les dermatoglyphes font aussi distinguer entre jumeaux mono- ou dizygotes en fonction du nombre de différences constatées. Les jumeaux monozygotes présentent des empreintes légèrement dissemblables parce que la formation de ces empreintes s'effectue au cours de la vie intra-utérine et dépend en partie de la position du fœtus et des contacts inter-tissulaires, différents pour les deux jumeaux. Les dermatoglyphes aident aussi au diagnostic de maladies génétiques telles que trisomies, syndrome de Turner, syndrome de Klinefelter... Une agénésie digitale préserve des dermatoglyphes embryonnaires, à la différence d'une amputation (fig. 4).

Fig. 3 – Em-preinte digitale

a
b c

a) Minuties et caractères de classification des empreintes ; b) capture numérique de l'empreinte d'un l'index gauche ; c) affichage numérique comparé de deux empreintes, de l'index, droit et gauche

(doc. Thalès).

La supériorité des empreintes digitales est leur trace, laissée par les doigts sur les objets manipulés (enquêtes criminelles). Au-delà d'une utilisation policière ou médicale, la demande de systèmes d'identification par empreintes digitales (*Finger Prints Identification Systems*) augmente dans les domaines de l'accès sécurisé et de la domotique. Ces systèmes remplacent, de plus en plus, clés ou cartes d'accès à un lieu et codes d'accès aux ordinateurs [18].

Autre anatomie de surface, la géométrie de la main suppose la bonne volonté de l'intéressé pour obtenir une photographie en conditions optimales. Environ 90 dimensions et caractères morphologiques (longueur et largeur des doigts, forme des articulations...) sont retenus. Simple et inoffensive, la technique se développe avec succès en milieu scolaire, la présentation de la main (2s) remplaçant le ticket de cantine quotidien. La vascularisation des doigts et le réseau veineux dorsal de la main faciles à analyser, mais d'une moindre fiabilité, sont également exploités.

Anatomie dynamique et comportements

Le rythme de frappe sur clavier remplace de plus en plus l'étude de l'écriture. La frappe sur clavier peut augmenter le degré de sûreté d'un mot de passe en étudiant la façon dont ce mot de passe est saisi.

La trajectoire et la dynamique de la signature peuvent être également analysées [19]. Le grand avantage de ce système réside dans la valeur juridique de la méthode.

Les mouvements des lèvres et la reconnaissance vocale ont été aussi retenus comme tests d'identification. La voix est une information utilisable au téléphone, mais elle dépend de plusieurs paramètres, d'où sa moindre sûreté.

D'autres techniques, en voie d'émergence, ne sont pas encore validées, telle l'étude de l'odeur corporelle, l'analyse automatique de la démarche et de la trace du pavillon de l'oreille sur la literie. Cette dernière participe, avec les empreintes digitales, à l'identification criminelle.

Fig. 4 – Empreintes digitales et malformations des doigts et de la main (doc. M.O. Réthoré [17]).

a b c
d e f

a) Agénésie de la main et empreinte distale ; b) Syndactylie ; c) Agénésie des doigts et du pouce ; d) Plis palmaires dans cas de situs inversus ; e,f) Amputation congénitale des doigts, photo (e) et empreinte palmaires (f).

Identification et localisation par GPS, avec micro-émetteur RF implanté (RFID) ou non, contenant le dossier médical.

Au Royaume-Uni, le port obligatoire, depuis 2003, d'un bracelet électronique mobile permet la surveillance électronique par GPS (*Global Positioning System*), de 300 pédophiles ou délinquants sexuels, déjà condamnés. La France semble en voie d'adopter cette mesure.

Les « *Radio-Frequency Identification Devices* (RFID) » sont des microémetteurs implantés sous l'hypoderme, initialement destinés à la surveillance d'animaux domestiques ou sauvages. Le contrôle des déplacements de détenus libérés et de malades a été l'étape suivante. Autorisée aux USA par la FDA (2004), la « Verichip » est utilisée prioritairement dans le domaine médical (hôpitaux du New-Jersey depuis 2006) pour l'identification et la surveillance des patients atteints de la maladie d'Alzheimer ou d'autres troubles psychiatriques. Elle peut contenir des éléments du dossier médical utilisables en cas d'urgence [20, 21]. Élément de 2x6 mm (mémoire, émetteur et antenne) implanté en sous-cutané (main ou avant-bras), la puce électronique émet en continu, sur une radiofréquence proche de 130 kHz, le code d'identification à 16 chiffres de son porteur, permettant ainsi sa géo-localisation par l'intermédiaire d'un satellite géo-stationnaire. Les dérives possibles du système, évidentes, suscitent aux États-Unis un débat qui s'amplifie.

Performances des systèmes biométriques

On doit considérer deux types de qualités pour un système, celles mesurant la probabilité de ne pas se tromper et celles caractérisant sa facilité d'utilisation et son acceptabilité.

Mesure des taux d'erreur

On distingue le taux de faux rejets (FRR pour « *False Rejection Rate* ») et le taux de fausse acceptation (FAR pour « *False Acceptance Rate* »). Les deux varient de 0 à l'infini. Plus la marge d'erreur autorisée augmente, plus les fausses acceptations augmentent et plus les faux rejets diminuent : le système est moins sûr, mais plus fonctionnel. Au contraire, lorsque la marge d'erreur autorisée se réduit, on accepte de moins en moins d'individus essayant de frauder, mais on augmente le taux de rejet des personnes autorisées, ce qui amoindrit la tolérance du système. On appelle taux d'égale erreur le point pour lequel FRR et FAR sont identiques. Selon les conditions et le but suivi, un équilibre est recherché entre FRR et FAR. En plus de ces deux paramètres, intervient le taux d'impossibilité d'enregistrement des données (FTE pour « *Failure To Enroll* »). Dans les techniques les plus utilisées, on constate que le FRR est faible pour les empreintes digitales et l'étude de l'iris (<1%) alors qu'il est plus élevé pour l'étude du visage (1-5%). Le FAR est de 1/10 pour les deux premiers et varie de 1 à 5% pour le troisième. Le FTE est quasi nul pour le visage, très faible pour les empreintes digitales et faible pour l'iris. La performance augmente avec l'utilisation de deux techniques ou plus (« multi-biométrie ») [22]. Mais l'examen est plus long, sa tolérance diminue et son coût augmente. *International Biometric Group* (IBG), société de conseil très présente sur Internet, établit régulièrement le classement multifactoriel (variable au rythme des progrès) des 8 principales techniques biométriques.

Acceptabilité et autres caractéristiques

Fig. 5 – Tableau résumé comparant, selon quatre critères, à importance cotée de 1 à 5, sept des techniques de la biométrie, (doc. EAC, YD [2], d'après C. Cabal [5], Government Accountability Office (GAO) et International Biometric Group (IBG) [19]).

D'autres caractéristiques d'évaluation d'une technique sont aussi à considérer : 1- l'effort requis pour subir le test (acceptabilité) ; 2- son caractère « intrusif » éventuel (perception par l'individu d'une atteinte à sa liberté et/ou à son intégrité physique) ; 3- son universalité ; 4- l'interopérabilité des méthodes entre pays et fabricants ; 5- son coût. Un tableau résume ces données pour sept techniques (Fig. 5). Enfin, en cas de défaillance de la technique, les résultats doivent être contrôlables par un opérateur extérieur (reconnaissance du visage par exemple). Au total, et comme la CNIL y insiste, le choix de la technique dépend, essentiellement, du but que l'on poursuit [4].

Utilisation de la biométrie

Les utilisations de la biométrie sont multiples. Les principales sont l'identification des personnes pour permettre l'accès à un appareil, un lieu (de sa maison à un site atomique), ou un pays, l'identification des délinquants et de leurs victimes, la sécurité des transactions et l'étude scientifique des caractéristiques d'une population.

Dans le domaine civil, les systèmes d'identification nationaux (cartes d'identité, de sécurité sociale, de santé Vitale, les permis de conduire, de naviguer, de chasser...) répondent progressivement aux mêmes normes – biométriques – que celles exigées pour le contrôle aux frontières et les voyages internationaux (passeports, visas, cartes vertes). Depuis 2005, le passeport biométrique français comporte une photographie faciale, normalisée, en attendant, prochainement, la ou les empreintes digitales. Une puce, comportant données de l'état civil et photographie, ainsi qu'une antenne, sont insérées dans la couverture du passeport. Seuls les lecteurs autorisés et possédant la clé décodent le fichier et le comparent à une base de données [23]. A ce propos, notons que l'obtention de ce passeport en Préfecture exige la présentation de la copie intégrale de l'acte de naissance, i.e. la copie complète du registre d'état civil comportant la filiation. Le rapport d'activité 2004-2005 du Conseil National pour l'Accès aux Origines Personnelles (CNAOP) [24], l'avis n° 90 du CCNE, « Accès aux origines, anonymat et secret de la filiation » du 24 novembre 2005 [25] attirent l'attention sur les problèmes que cela entraîne pour les personnes nées sous X, constat sujet de préoccupation de l'Académie Nationale de Médecine [26].

Dans les aéroports, le FAR du système d'identification choisi doit être le plus petit possible afin de diminuer le nombre des contrôles visuels. Le temps passé à la vérification doit être court pour éviter « l'engorgement ». Il est par exemple de 1,5 s. pour la vérification biométrique et de moins de 2 s. pour vérifier le droit d'accès à la zone protégée de l'aéroport d'Orly. Le FRR est faible, l'autorisation intervenant souvent au 2^e essai. Dans ces systèmes existe un contrôle en temps réel de la validité et des droits d'accès du document présenté. Malgré ces progrès et l'opposition européenne, il semble que les USA reviennent à l'exigence supplémentaire d'un visa électronique [27].

En matière pénale, des bases de données regroupent les éléments d'identification des délinquants connus. Aux États-Unis, 60 millions d'empreintes digitales sont enregistrées dans l'IAFIS (« *Integrated Automated Fingerprint Identification System* »), utilisé pour plus de 50.000 recherches par jour. En France, le FAED (Fichier Automatisé des Empreintes Digitales), créé en 1987, est moins riche que celui des États-Unis, mais la durée de conservation des empreintes y est limitée à 25 ans.

En médecine légale, la biométrie et la génétique, groupant de multiples techniques (radiologie et examen clinique, autopsie, étude de l'ADN) servent à identifier les corps et leurs fragments. A défaut d'identification, ces techniques déterminent le sexe, l'âge, la stature et l'origine géographique du défunt. La détermination de l'âge chronologique clinique et radiologique, telle que requise par les magistrats, a pour but de vérifier la réalité de l'âge affirmée par le sujet et dont dépend l'application de la procédure pénale différente selon les quatre âges de référence, 13, 15, 16 et 18 ans, âge de la majorité légale.

Au quotidien, l'identification des empreintes digitales, rapide et simple, étend ses applications. A la caisse du magasin, poser son index sur une touche pendant moins de 1 s. est plus simple qu'ajuster ses lunettes, fouiller son sac, en extraire sa carte bancaire, s'en remémorer le code et le composer sur un clavier. En outre, l'authentification du sujet élimine le risque de perte ou de vol. Des chaînes de supermarchés, soulageant ainsi les personnes fragiles, âgées et/ou handicapées, unanimement satisfaites, installent ces détecteurs aux caisses, avec débit bancaire automatisé (après enregistrement de leur état civil et autorisation bancaire) [28].

Le polymorphisme génétique (génomique et protéique)

Les empreintes génétiques

La technique consiste à prélever un échantillon (salive, cheveu, sang...) et à en sécuriser le transport au laboratoire afin de pouvoir y extraire l'ADN des cellules qu'il contient [29]. Les marqueurs génomiques, aujourd'hui couramment étudiés, sont des *loci* de courte séquence (2 à 5 nucléotides), répétés un nombre variable de fois, et appelés STR (« Short Tandem Repeat ») ou microsatellites. Il s'agit donc d'un polymorphisme de taille qui présente surtout l'intérêt d'être poly-allélique, donc très informatif, et d'être réparti de façon à peu près uniforme sur l'ensemble du génome dans les espaces inter-géniques, dans les introns et les régions flanquantes des gènes. L'amplification par PCR (*Polymerase Chain Reaction*) à partir d'amorces flanquant le STR conduit à une synthèse exponentielle de fragments d'ADN, dont les tailles sont différentes en fonction du nombre de répétition qu'ils contiennent. Ces fragments sont ensuite séparés d'après leurs tailles par une technique d'électrophorèse capillaire. L'utilisation de la PCR confère à cette technique une extrême sensibilité qui permet la

détermination des empreintes génétiques sur des traces infimes de matériel biologique. Actuellement des trousse commerciales permettent de tester simultanément entre 11 et 16 marqueurs génétiques autosomiques indépendants. Le choix de ces marqueurs a fait l'objet d'un accord international, facilitant ainsi la lecture et les échanges entre laboratoires et services nationaux. L'existence de quelques cas particuliers (mosaïcismes, chimérismes) doit être prise en compte pour l'interprétation des résultats. En outre les empreintes génétiques standard sont rigoureusement identiques chez des jumeaux monozygotes. De nombreux pays sont dotés d'une banque de données stockant les empreintes génétiques de délinquants ou de suspects. C'est le cas de la France avec le Fichier National Automatisé d'Empreintes Génétiques (FNAEG). Créé en 2003, il rassemble actuellement environ 450.000 profils génétiques. Il permet de croiser les informations recueillies avec les identités recensées. Au Royaume Uni, environ 3,5 millions de profils sont enregistrés et 8 à 10 mille rapprochements par an sont effectués. Dans certaines circonstances l'étude porte sur l'ADN mitochondrial et, chez l'homme, sur le chromosome Y. L'ADN mitochondrial, à la différence de l'ADN nucléaire d'origine biparentale, provient uniquement de la mère. Les mitochondries sont nombreuses dans la plupart des tissus. L'étude porte sur la région promotrice de l'ADN mitochondrial appelée *D-Loop* ». La technique utilisée est un séquençage complet de cette région et est donc plus coûteuse. C'est la raison pour laquelle elle est habituellement réservée à des analyses de tissus dégradés (criminalistique). Les variations entre individus sont des différences de quelques nucléotides dans cette séquence; elles définissent des « mitotypes » qui sont comparés à une séquence de référence. L'étude de l'ADN mitochondrial est, bien entendu, beaucoup moins discriminante que l'étude de l'ADN nucléaire. L'analyse de microsatellites du chromosome Y a sa place en cas de viols multiples ainsi que dans le cas d'une recherche de paternité, en l'absence du père biologique supposé, par l'examen des apparentés de sexe masculin. Les techniques d'études des empreintes génétiques évoluent très rapidement : automatisation et miniaturisation des méthodes, utilisation de puces à ADN décelant des variations portant sur un seul nucléotide (SNP ou « *Single Nucleotide Polymorphism* »), détermination du phénotype (portrait robot) à partir de la connaissance des gènes codant pour les paramètres biométriques.

Polymorphisme des protéines

Bien avant l'utilisation des empreintes génétiques (1985), le polymorphisme génétique était appréhendé en étudiant par des méthodes électrophorétiques ou immunologiques les variations structurales, génétiquement déterminées, de protéines plus ou moins complexes : antigènes des groupes sanguins érythrocytaires, protéines du sérum, enzymes érythrocytaires, groupes tissulaires (système HLA). La plupart de ces polymorphismes sont bi ou pauci-alléliques donc peu informatifs. Ce n'est pas le cas du système HLA dont les gènes sont extrêmement poly-alléliques [30], ce qui permet de confirmer ou de rejeter une paternité avec une quasi-certitude ; par contre, la nécessité d'une prise de sang a limité son utilisation pour les usages habituels des techniques d'identification.

Le développement des techniques de spectrométrie de masse permet, théoriquement, l'analyse simultanée de la totalité des protéines (protéome) d'un type cellulaire ou d'un tissu, ainsi que de leurs variations structurales. Cette technologie très lourde est débutante. Elle est également difficile, essentiellement parce que le nombre de protéines (entre 100.000 et 150.000) est beaucoup plus élevé que celui des gènes (23.000). L'analyse protéomique globale n'est donc pas utilisable aujourd'hui pour l'identification des individus.

Applications de la génétique à l'identification des personnes.

Examinées, dès 1989, par l'avis n° 17 du Conseil Consultatif National d'Éthique pour les Sciences de la vie et de la santé (CCNE) [31], organisées par le Décret n° 2000-570 du 23 juin 2000 [32], objet de la Loi n° 2004-800 du 6 août 2004 [33] et de la saisine d'urgence du CCNE le mercredi 3 octobre 2007 avec remise de son avis n° 100 le jeudi 4 octobre 2007 [34], les applications de la génétique à l'identification des personnes soulèvent des questions majeures. On doit rappeler, en outre, que la filiation en droit français ne repose pas sur la génétique, mais sur des normes juridiques et, en particulier, la « possession d'état ». Les applications de la génétique concernent les enquêtes judiciaires en droit pénal, les tests de paternité en droit pénal et civil, les études des populations (recherche médicale et santé publique) et, enfin, les demandes privées de convenance personnelle. Dans les enquêtes judiciaires criminelles, notamment celles qui concernent les actes de terrorisme, les homicides, les agressions sexuelles et les vols à main armée, tout objet supportant un tissu biologique peut être analysé. Le principe de l'analyse consiste à comparer une empreinte génétique relevée sur le lieu d'un crime avec celle d'un individu (suspect), de la victime, ou bien celles contenues dans une banque de données. La conclusion est soit le rejet d'identité qui est absolu, soit la haute probabilité d'identité qui dépend du nombre de marqueurs examinés et de leur répartition dans la population. L'examen de l'ADN d'un cadavre, lorsqu'il est possible, peut l'identifier, par comparaison avec l'ADN de membres de la famille ou de celui contenu dans une base de données. Les tests de paternité sont basés sur la comparaison des STR de l'allèle paternel avec ceux observés chez l'enfant. Là aussi, le rejet est absolu et la probabilité de paternité très grande lorsqu'on a étudié les 16 marqueurs disponibles dans les trousse commerciales.

Quelques questions soulevées par l'identification des personnes

Personne ne conteste les besoins d'identification, mais personne ne conteste, non plus, que le processus doit être réglementé afin de préserver les libertés individuelles. Plusieurs questions, ici résumées, sont sujettes à débat.

Dans quels cas constituer et comment exploiter une base de données ?

La Loi n° 2004-801 du 6 août 2004, relative à la protection des personnes physiques à l'égard des traitements de données à caractère personnel [35], soumet à l'autorisation de la CNIL la création de fichiers biométriques par des organismes privés. Une saisine préalable de la CNIL s'impose aux projets des pouvoirs publics. La décision prend en compte la finalité de la banque (est-elle acceptable ou non ?) et l'adéquation des moyens au but recherché (ces moyens respectent-ils le principe de proportionnalité ou sont-ils excessifs?). Dans tous les cas, l'équilibre est recherché entre nécessité d'assurer la sécurité des personnes et des biens et obligation de préserver les libertés individuelles. En revanche, la mise en œuvre de dispositifs d'accès, avec enregistrement des données biométriques sur un support individuel (sans création de fichier élargi) fait l'objet d'une simple déclaration.

L'exploitation de la base de données obéit à plusieurs contraintes :

- 1- l'inscription à un fichier doit recueillir le consentement écrit de l'intéressé, sauf exceptions prévues par la loi et l'exactitude des données enregistrées doit être vérifiée avec soin. Le corollaire en est que toute personne doit pouvoir accéder aux données qui le concernent pour en vérifier l'exactitude ;
- 2- les données qui y sont présentes doivent être protégées, ce qui oblige souvent à les stocker sur un support externe ;
- 3- les indications concernant un individu doivent être conservées pour un temps limité et fixé par avance ;
- 4- les données nominatives ne doivent pas être croisées avec d'autres données nominatives contenues dans une autre

banque à finalité différente ;

5- beaucoup de fichiers pouvant contenir des données supplémentaires à celles nécessaires aux objectifs du système, le « gabarit » extrait de la banque doit être conçu de manière à ne pas comporter des renseignements superflus ;

6- enfin, toute dérive ou extension non prévue au moment de la création réclame l'établissement d'un nouveau dossier de demande.

Il est évident que les missions de la CNIL sont considérables et que le nombre de demandes qu'elle a à examiner augmente avec le temps, ce qui ne peut se faire que si les moyens qui lui sont alloués sont en adéquation avec ses responsabilités.

L'Union Européenne et le Conseil de l'Europe portent une vive attention au dossier comme en témoigne la Charte Européenne des Droits Fondamentaux à laquelle fait référence le projet de traité européen simplifié [36 à 38].

Comment doit-on réglementer l'accès aux données génétiques ?

Actuellement, toute analyse génétique à des fins médicales ne peut se faire que dans un laboratoire agréé, sur prescription médicale [35]. Les résultats de l'analyse sont confidentiels ; mais, l'information des membres de la famille par un médecin est prévue lorsque ceux-ci peuvent être concernés par ces résultats et doivent bénéficier d'un conseil génétique. L'Agence de la Biomédecine est chargée de l'application de la réglementation. Les analyses génétiques à des fins judiciaires sont subordonnées à la demande d'un magistrat, dans le cadre de procédures spécifiques, et sont également effectuées dans des laboratoires agréés, sous la responsabilité d'un expert judiciaire inscrit sur les listes de cours d'appel ou agréé par la Cour de Cassation. Le domaine administratif fait appel aux analyses génétiques dans onze des pays de l'Union européenne. L'objectif est de vérifier la filiation à l'occasion d'une demande de regroupement familial lors d'une immigration. Les tests sont volontaires. En France, un projet de loi semblable, après avoir suscité une vive émotion [39], a entraîné la saisine d'urgence du CCNE (mercredi 3 octobre 2007) avec avis rendu le jeudi 4 octobre 2007 [34].

Les analyses génétiques à des fins scientifiques se multiplient, mais à la différence des précédentes, les données peuvent être cryptées, l'étude statistique n'exigeant pas de connaître l'identité des sujets étudiés. Certains imaginent la possibilité de multiplier les génotypages d'une population entière. En France, cela serait possible progressivement en conservant et en analysant l'ADN des nouveau-nés, à partir des échantillons