

## Utilisation des profils ADN dans un contexte médico-légal

### Compte rendu de la conférence donnée aux XIX<sup>èmes</sup> journées de biologie clinique tunisienne

V. CASTELLA

Laboratoire de génétique  
forensique de l'Institut  
universitaire de médecine  
légale de Lausanne, Suisse  
(Vincent.castella@chuv.ch)

**Résumé** : les outils de biologie moléculaire sont utilisés dans de nombreux laboratoires afin de fournir des éléments objectifs aux autorités chargées de l'application des textes législatifs ainsi qu'aux personnes souhaitant connaître leur origine. Des fichiers ADN nationaux sont par exemple en fonction depuis 1995 en Angleterre, 1998 en Allemagne et aux Etats-Unis et 2000 en Suisse. Le présent compte rendu décrit la situation rencontrée en Suisse en 2005 et fournit quelques éclaircissements sur la méthodologie associée à ce champ d'activité.

### L'ADN

L'ADN (acide désoxyribonucléique), support de l'hérédité, est présent dans le noyau des cellules. Il est constitué de quatre bases ou nucléotides qui sont l'adénine (A), la thymine (T), la guanine (G) et la cytosine (C). Ces bases se lient les unes aux autres pour former des brins d'ADN. La structure moléculaire de l'ADN peut être comparée à une échelle vrillée dont les montants sont constitués de brins d'ADN complémentaires. Il s'agit de la structure en double hélice découverte par Watson et Crick en 1953. Les «marches» de cette échelle sont constituées par l'appariement des bases A-T ou G-C. Ceci signifie qu'en connaissant la séquence d'un brin d'ADN on peut en déduire la séquence du brin complémentaire. Chez l'être humain, l'ADN nucléaire est composé d'environ 6 milliards de bases réparties entre 23 paires de chromosomes.

### La variabilité génétique

A l'exception des vrais jumeaux, chaque personne possède un ADN nucléaire qui lui est propre. Cet ADN se trouve dans toutes les cellules nucléées du corps et il est invariable au cours du temps. Il existe pourtant une très grande variabilité génétique entre individus qui est principalement localisée dans les régions non codantes de l'ADN. Ces régions représentent environ 98% de l'ADN et ne contiennent pas d'information directement utile pour la synthèse de protéines. Elles peuvent donc être modifiées sans que des gènes perdent leur fonction. La

variabilité génétique est une conséquence de la reproduction sexuée. La femme et l'homme produisent des gamètes, ovules ou spermatozoïdes, qui ne contiennent qu'une moitié de leur patrimoine génétique, soit un lot de 23 chromosomes. Ainsi, l'embryon issu de la fusion d'un ovule et d'un spermatozoïde possèdera à nouveau 23 paires de chromosomes. La production des gamètes se fait au cours de divisions cellulaires appelées méioses. Au cours de ces dernières, des fragments de chromosomes peuvent être échangés au sein d'une même paire. Ces échanges, ou recombinaisons, constituent une source de variabilité génétique en créant des chromosomes différents de leurs «modèles». Une autre source de variabilité est créée lors de la méiose par le tirage aléatoire d'un seul des deux chromosomes de chacune des 23 paires. Finalement, il existe des modifications spontanées de la séquence d'ADN que l'on appelle mutations. Il s'agit d'événements naturels relativement rares par lesquelles une ou plusieurs bases peuvent être changées, ajoutées ou enlevées.

### Les profils ADN

Les microsatellites ou STRs (short tandem repeats) sont les marqueurs les plus utilisés en génétique forensique humaine. Ce sont des fragments d'ADN nucléaire situés à des endroits précis dans le génome (locus). Les STRs sont constitués de motifs de 2 à 8 bases répétés à la suite, par exemple : CATGCATGCATGCATG. Leur polymorphisme est généré par des mutations qui créent des

variations du nombre de répétitions de ces motifs. Le nom de chaque variant ou allèle est déterminé par le nombre de répétitions du motif qu'il contient. L'allèle donné en exemple ci-dessus serait l'allèle 4 puisqu'il contient quatre répétitions du motif «CATG». Un individu homozygote pour un locus possède le même allèle sur le chromosome d'origine paternelle et sur celui d'origine maternelle (exemple de notation : 4 ou 4-4). A l'inverse, un individu hétérozygote a hérité d'allèles différents de ses parents (ex : 4-7). Un profil ADN ou profil génétique recense les allèles d'un individu à un ou plusieurs locus.

Plusieurs étapes sont nécessaires à l'établissement d'un profil ADN. Dans un premier temps, l'ADN doit être extrait de l'échantillon à analyser. Au cours du processus d'extraction, les parois des cellules sont lysées puis l'ADN est purifié des protéines et autres débris cellulaires. Dans un deuxième temps, l'ADN extrait doit être amplifié afin de devenir détectable. Cette amplification est réalisée au moyen d'une PCR. Il s'agit d'une réaction de polymérisation en chaîne au cours de laquelle une enzyme, la Taq polymérase, copie un grand nombre de fois un fragment cible d'ADN. La spécificité de cette amplification est déterminée par le choix d'une paire d'amorces (courtes séquences d'ADN) qui sont complémentaires aux séquences se trouvant de part et d'autre du fragment cible. La PCR peut permettre de synthétiser des millions de copies à partir d'un seul fragment d'ADN initial. Finalement, la séparation et la détection de l'ADN amplifié se fait au moyen d'appareils d'électrophorèse capillaire ou séquenceurs automatiques. Ces appareils contiennent des micro-tubes (les capillaires) remplis d'un liquide visqueux dans lesquels migrent les fragments d'ADN soumis à un champ électrique. Cette migration permet d'identifier les différents allèles selon leur vitesse de migration. La détection des fragments d'ADN est assurée par un laser et une caméra spéciale.

Des kits d'analyse commerciaux ont été développés afin de pouvoir analyser plusieurs STRs en une seule réaction PCR. Les kits SGM Plus (Applied Biosystems) et PowerPlex 16 (Promega) contiennent respectivement 10

et 15 STRs ainsi que l'amélogénine qui est un marqueur utilisé pour déterminer le sexe de la personne source de l'échantillon analysé. Les STRs qui font partie de ces kits sont situés dans des régions non codantes de l'ADN. Ils ne permettent donc pas d'obtenir d'informations sur l'état de santé des individus (ex : prédisposition à des maladies) ou sur leurs caractéristiques physiques (ex : couleur des yeux, taille) hormis le sexe et certaines aberrations chromosomiques.

### Les expertises en filiation

Des tests ADN peuvent être utilisés pour confirmer ou réfuter l'existence d'un lien de parenté spécifié. Les tests de paternité sont basés sur le fait que chaque enfant reçoit une moitié du matériel génétique de sa mère et l'autre moitié de son père. En comparant les profils ADN des personnes concernées, il est donc possible de déterminer si un individu peut être ou non le parent biologique d'un enfant. Par exemple, si l'enfant est 4-5 et sa mère 5-7 au même locus, on peut remarquer qu'ils possèdent un allèle 5 en commun. L'enfant a donc hérité cet allèle de sa mère. L'allèle 4 de l'enfant doit donc forcément provenir du père biologique. Ce dernier pourrait par exemple être 4-4 ou 4-6. Si une personne et un enfant présentent des profils ADN compatibles, une probabilité de paternité est calculée. Les probabilités de paternité obtenues en routine avec 15 STRs sont généralement supérieures à 99.999%. Selon la jurisprudence suisse, une valeur supérieure à 99.8% permet de conclure que le lien de parenté testé est pratiquement prouvé. Si une personne présente un profil génétique qui n'est pas compatible avec celui de l'enfant, et que ces incompatibilités ne sont pas dues à des mutations, cette personne ne peut pas être le parent biologique de l'enfant. Dans un tel cas, il est conclu que le parent présumé est exclu comme étant le parent biologique. Cette conclusion est une certitude et aucun calcul statistique n'est nécessaire.

Ce principe peut être étendu à d'autres liens de parenté. Par exemple des tests de fraternité pour lesquels on peut vérifier que de vrais frères, sœurs ou frère et sœur ont en moyenne 25% de leur ADN en commun. Pour des tests de gemellité, il s'agit de contrôler que les jumeaux uni-

vitellins ont des profils ADN 100% identiques. Le principe de ces tests ADN est également utilisé pour identifier des cadavres inconnus. Pour se faire, le profil ADN établi à partir d'un défunt peut être comparé avec les profils ADN de parents présumés afin de vérifier son identité. Sous certaines conditions, il est également possible de comparer le profil ADN du défunt avec le profil ADN établi à partir d'effets personnels, comme une brosse à dents ayant appartenu à une personne disparue. Il n'est pas possible d'identifier un corps par l'ADN sans avoir une identité présumée ou un fichier contenant les profils ADN des personnes disparues.

### L'analyse des traces biologiques

En Suisse, les laboratoires qui analysent des traces ADN doivent être reconnus au niveau national. En particulier ces laboratoires doivent être conformes aux normes ISO 17025 et participer régulièrement à des contrôles de compétences nationaux et internationaux. Il est également exigé que toutes les analyses soient effectuées au moins en double. Actuellement les six laboratoires des instituts de médecine légale de Bâle, Berne, Genève, Lausanne, Saint-Gall et Zürich sont actifs. A moyen terme, il est possible que des laboratoires privés s'implantent également.

Le principe de l'analyse des traces peut être résumé comme suit : si le profil ADN réalisé à partir d'un objet trouvé sur les lieux d'un délit correspond au profil ADN d'une personne, il y a des raisons de penser que cette personne était sur les lieux du délit. Ce type d'information est considéré comme un indice et non comme une preuve. Un calcul statistique permet d'évaluer la valeur probante d'une telle correspondance. En Suisse, les profils ADN utilisés dans ce contexte comptent 10 locus et la valeur probante se calcule selon la fréquence des allèles présents à chaque locus. Comme ordre de grandeur, il peut y avoir environ une chance sur 10 que deux personnes non apparentées aient le même profil ADN à un locus. A partir de 5 locus, il peut y avoir moins d'une chance sur 1 million d'observer une telle correspondance de façon fortuite.

Les traces peuvent être regroupées en fonction de leur

nature. Au cours du premier semestre 2004, 901 traces ADN ont été analysées au laboratoire de génétique forensique de Lausanne (LGF). La répartition suivante a été observée :

- 47% de traces de contact (prélèvement sur une poignée de porte, sur le manche d'un outil, sur le volant d'une voiture, sur des vêtements, etc.)
- 30% de traces de salive (mégots, goulots de bouteille, restes alimentaires, etc.)
- 12% de traces de sang sur des supports divers
- 9% de traces de sperme (frottis gynécologiques, sur des vêtements, etc.)
- 1% de cheveux ou poils
- 1% d'excréments

Les prélèvements reçus au laboratoire peuvent être des écouvillons (cotons tiges) que la police a utilisé pour récolter la trace, des fragments de trace (ex : un morceau de vêtement), ou des objets entiers. Dans ce dernier cas, le laboratoire est responsable d'effectuer les prélèvements nécessaires. Avant de procéder à l'analyse génétique proprement dite, le laboratoire peut être appelé à confirmer de façon formelle la nature supposée de la trace. Pour se faire il possède une batterie de tests appelés tests indicatifs ou préliminaires. Il peut ainsi déterminer si la trace contient du sang au moyen d'un test immunochromatographique sensible à la présence d'hémoglobine humaine, si elle contient des spermatozoïdes par microscopie, si elle contient un éjaculat par un test immunochromatographique sensible à des antigènes prostatiques ou si elle contient de la salive par un test enzymatique basé sur l'activité de l' $\alpha$ -amylase.

Toujours selon les statistiques internes du LGF, un profil ADN exploitable a pu être obtenu pour environ 60% des traces analysées en 2004. Le taux de succès varie selon la nature de la trace. En général, les résultats négatifs sont obtenus pour les traces qui contiennent peu ou pas d'ADN. Ceci est plus fréquent pour les traces de contact.

### Le fichier suisse de profils ADN

Le rôle d'un fichier ADN est de stocker des profils ADN et de les comparer entre eux afin de mettre en évidence d'éventuelles correspondances que l'on appelle des hits.

Les critères d'introduction et de retrait de profils ADN d'un fichier varient d'un pays à un autre. Par exemple, des pays n'y introduisent que des personnes reconnues coupables de certaines catégories de crimes. D'autres pays peuvent également y introduire des suspects, parfois selon un catalogue de délits. En Suisse, le système est devenu relativement permissif puisque, depuis le 1er janvier 2005, les profils ADN de toutes personnes soupçonnées d'avoir commis un crime ou un délit ou d'y avoir participé ainsi que les profils ADN des personnes condamnées peuvent être introduits dans le fichier. Ce fichier peut également être utilisé en dehors d'une procédure pénale afin d'identifier des personnes ou des cadavres inconnus.

Le fichier ADN suisse est en fonction depuis l'été 2000. Les profils ADN qui y sont introduits doivent comporter un nombre minimum de locus afin de limiter les risques de hits fortuits. Ces limites sont de 10 locus pour les profils de personnes, de 6 locus pour les profils de traces comportant de l'ADN d'une seule personne et de 8 locus pour des profils de traces comportant de l'ADN de deux personnes. Sur demande des autorités compétentes, les profils ADN qui ne satisfont pas à ces critères peuvent être utilisés pour réaliser des comparaisons «manuelles» avec d'autres profils ADN. Au 31 juillet 2005 le fichier ADN suisse comptait 72 762 profils ADN dont 62 636 profils de personnes et 10 126 profils de traces.

Une distinction est faite entre les hits trace-personne qui permettent d'identifier une trace et les hits trace-trace qui permettent de mettre en relation différents délits pour lesquels une même personne est impliquée. Ces derniers sont également utiles puisque la police peut posséder des éléments d'enquête (témoignages, empreintes digitales, etc.) pour une affaire qui peuvent dès lors permettre d'élucider une série d'affaires liées par l'ADN. Au 31 juillet 2005, 8 401 hits trace-personne et 8 653 hits trace-trace ont été constatés depuis la mise en service du fichier ADN suisse. Une trace identifiée est retirée du fichier national afin de ne pas alourdir le système.

### **Marqueurs génétiques complémentaires**

Dans certains cas, l'utilisation de marqueurs génétiques

complémentaires peut aider à interpréter des cas difficiles ou peut augmenter la valeur probante d'un test de paternité ou d'une identification. Les plus utilisés sont les STRs situés sur le chromosome Y (Y-STRs) et les marqueurs d'ADN mitochondrial (ADNmt).

**Les Y-STRs :** parmi les 23 paires de chromosomes présents chez l'être humain, se trouve une paire de chromosomes sexuels, les hétérosomes, qui sont les chromosomes X et Y. Les femmes possèdent les chromosomes XX et les hommes XY. Chez les femmes, des échanges de matériel génétique (recombinaisons), peuvent avoir lieu sur la totalité des deux chromosomes X. Par contre, chez les hommes, seul 5% du chromosome Y peut se recombiner avec un chromosome X. Les Y-STRs utilisés en génétique forensique font partie de régions non recombinantes. Ceci signifie, qu'en l'absence de mutation, le profil Y-STR est transmis inchangé de père en fils. On s'attend donc à ce que tous les hommes d'une lignée paternelle possèdent un même profil Y-STR. Ces marqueurs sont donc utilisés pour tester certains liens de parenté, comme des liens de fraternité ou de grand-père paternel à petit fils.

De par leur spécificité masculine, les Y-STRs peuvent également être utilisés pour des affaires de mœurs. Les échantillons analysés pour ces cas (prélèvements gynécologiques, traces sur des vêtements, etc.) contiennent généralement un mélange de matériel féminin et masculin. Si l'on utilise des marqueurs classiques, le profil masculin risque d'être «masqué» par le profil féminin et ne pourra pas être détecté. Des protocoles d'extraction d'ADN spéciaux incluant une lyse différentielle des cellules épithéliales et des spermatozoïdes peuvent permettre de séparer grossièrement l'ADN d'origine féminine et l'ADN d'origine masculine. Cependant, seule l'utilisation de Y-STR peut permettre de révéler les composantes masculines d'un mélange qui contient beaucoup plus d'ADN féminin que d'ADN masculin.

Du point de vue analytique, l'établissement de profils Y-STRs est semblable à celui des STRs autosomaux décrits plus haut. Il existe des kits commerciaux comme par exemple le kit PowerPlex-Y (Promega) qui permet d'amplifier 12 locus en une seule réaction PCR.

Les kits de STRs standards contiennent des marqueurs qui

sont généralement situés sur des chromosomes différents. Leur transmission est donc indépendante et l'information obtenue pour chacun d'eux peut être combinée. A l'inverse, les marqueurs d'un kit STRs-Y ne sont pas indépendants puisque tous sont sur le même chromosome. Ainsi, la valeur probante d'une correspondance entre deux profils n'est pas estimée en prenant en compte la fréquence des allèles présents à chaque locus, mais en prenant en compte la fréquence du profil complet. La valeur probante d'une identification par ce type de marqueur est donc bien moindre que celle qui peut être obtenue avec des marqueurs standards. Il faut également garder à l'esprit que les Y-STRs ne permettent pas d'identifier un individu, mais bien une lignée paternelle. Il n'est pas possible d'utiliser des fiches de profils Y-STRs dans le but d'identifier une trace. Le profil Y-STRs obtenu pour une trace peut par contre être comparé avec les profils d'éventuels suspects. Dans ce contexte, des bases de données sont utilisées pour estimer des valeurs probantes en cas de correspondances.

**L'ADNmt** : les mitochondries sont des organelles responsables de la production d'énergie dans les cellules. Elles contiennent leur propre ADN circulaire d'une longueur de 16'569 bases. Le polymorphisme de l'ADNmt est principalement présent dans sa région non codante, appelée région de contrôle ou D-loop. Les laboratoires de génétique forensique s'intéressent plus particulièrement à deux domaines hypervariables, appelés HVI et HVII, qui font partie de la région de contrôle et qui contiennent la plupart des mutations. L'information collectée pour dresser un profil d'ADNmt est généralement composée de deux séries d'environ 400 bases qui correspondent à la détermination des séquences des domaines HVI et HVII. Plutôt que de retranscrire ces 800 bases, les séquences obtenues sont comparées à une séquence de référence et seules les différences sont relevées. La séquence de référence internationale est celle qu'Anderson et collaborateurs ont publié en 1981 dans le journal Nature (numéro 290, pages 457-465). Elle est appelée séquence d'Anderson ou séquence de Cambridge.

L'intérêt de l'ADNmt est double. Cet ADN ne se recombine pas et se transmet maternellement. Ceci signifie, qu'en l'absence de mutations, toutes les personnes d'une

même lignée maternelle, femme ou homme, partagent un même profil d'ADNmt. A l'instar des Y-STRs, les profils d'ADNmt peuvent donc être utilisés pour tester des liens de parenté impliquant des personnes d'une même lignée maternelle (frère et sœur, grand-mère maternelle et petits enfants, etc.). Un autre avantage de l'ADNmt est son nombre élevé de copies par cellule. Comme vu précédemment, chaque cellule contient deux copies d'ADN nucléaire : une copie d'origine maternelle et l'autre paternelle. Contrairement à cela, chaque cellule contient des centaines de mitochondries qui elles-mêmes contiennent une dizaine de copies d'ADNmt. Au total il peut donc y avoir des centaines, voire des milliers de copies d'ADNmt dans une seule cellule. L'ADNmt peut donc s'avérer très utile pour analyser des échantillons qui contiennent très peu d'ADN nucléaire et/ou qui ont été stockés dans de mauvaises conditions. Typiquement, ces marqueurs sont utilisés pour l'analyse de cheveux ou d'ossements anciens.

Au niveau analytique, l'étude de l'ADNmt est beaucoup plus fastidieuse que celle des STRs. Pour établir les séquences des régions HVI et HVII d'une trace ou d'une personne, il faut réaliser plusieurs amplifications et plusieurs réactions de séquençage. Comme pour les Y-STRs, la valeur probante d'une correspondance entre profils d'ADNmt est beaucoup plus basse que celle obtenue par l'analyse de STRs standards. De plus l'ADNmt permet d'identifier une lignée maternelle et non pas un individu.

L'interprétation des analyses d'ADNmt est compliquée par l'existence d'hétéroplasmie. Ce terme signifie que différentes lignées d'ADNmt peuvent coexister chez un individu ou dans une seule cellule. Il est donc possible que deux séquences divergeant par une ou deux mutations puissent provenir du même individu. Au-delà de deux mutations, il est beaucoup plus vraisemblable que les profils d'ADNmt proviennent de personnes différentes.

### Références

- Butler JM (2001) Forensic DNA typing. Academic press, London, UK
- Coquoz R (2003) Preuve par l'ADN. Presses polytechniques et universitaires romandes, Lausanne, CH
- Aitken CGG and Taroni F (2004) Statistics and the evaluation of evidence for forensic scientists. Wiley, Chichester, UK
- Jobling MA, Hurles ME and Tyler-Smith C (2004) Human evolutionary genetics. Garland Science, New York