

La génétique, un outil essentiel pour la police scientifique et les recherches de parenté

Daniel LOCKER

Résumé

Une révolution dans la police scientifique intervient en 1985 grâce à un chercheur anglais sir Alec Jeffreys qui développe la technique des empreintes génétiques. Ces empreintes sont utilisées pour la première fois en 1986.

L'objectif de cette communication est de présenter le développement de ce test au cours des 20 dernières années et de montrer que dans le cadre de la police scientifique le test ADN n'est sans doute pas une preuve absolue mais un élément utilisable au même titre que d'autres analyses. Nous reviendrons également sur le développement de la généalogie génétique et sur son utilisation pour la recherche des origines et la résolution d'anciennes affaires criminelles.

Abstract

A revolution in forensic science took place in 1985 thanks to sir Alec Jeffreys who developed the genetic fingerprints technique. These fingerprints were used for the first time in 1986. The objective of this communication is to present the development of this test over the last 20 years and to show that the DNA test is undoubtedly not an absolute proof within the framework of forensic science but a suitable element in the same way than other proofs. We will also present the development of genetic genealogy and its use to trace the origins of different populations and to solve old criminal cases (cold cases).



Les séquençages de l'ADN des génomes humains ont permis de mettre en évidence deux types de marqueurs qui différencient les individus, les SNP¹ (polymorphisme d'un nucléotide) et les VNTR² (nombre variable de répétitions en tandem).

Les SNP

Entre deux personnes prises au hasard, il y a environ 3 millions de différences ponctuelles dans l'ADN.

Ces différences ou SNP correspondent au changement d'une base en un point précis de l'ADN (Figure 1). Cette position précise dans l'ADN peut exister sous deux formes appelées allèles : dans notre exemple la position de A et celle de G sont dites allèles. Les SNP sont retrouvés en moyenne une fois tous les 1000 bases.

A1 ...GTTACCTGGCAT**G**GCACATTGCTTTAA... dans l'ADN de Pierr
A2 ...GTTACCTGGCAT**A**GCACATTGCTTTAA... dans l'ADN de Paul

Figure 1. Exemple de SNP

¹ *Single nucleotide polymorphism*

² *Variable number tandem repeat*

Les VNTR

Les VNTR correspondent à des répétitions de bases en tandem présentes en un endroit précis d'un chromosome et dont le nombre varie d'un individu à un autre (Figure 2).

A1 ...**GTT**ACCTGGTGGTGGTGGATTGCTT**TAA**... ADN de Pierre TGGx5
A2 ...**GTT**ACCTGGTGGTGGTGGATTGCTT**TAA**... ADN de Paul TGGx4

Figure 2. Exemple de VNTR avec des répétitions TGG

Marqueurs VNTR et police scientifique

En 1985 A. Jeffreys, intéressé par l'évolution des gènes humains entreprend le séquençage de gènes humains notamment celui de la myoglobine³. Il met en évidence dans un intron⁴ de ce gène une séquence de 33 paires de bases répétée 4 fois. Il émet l'hypothèse que cette répétition n'est pas soumise à une pression de sélection et peut varier en nombre d'un individu à l'autre. Cette hypothèse est rapidement validée grâce aux méthodes, à l'époque classiques, d'hybridation de l'ADN avec des sondes radioactives. Il montre que cette séquence ou des séquences apparentées se retrouvent à différents endroits du génome humain. Elles ségrégent de génération en génération comme des allèles mendéliens⁵. Cette observation donne le signal de départ d'une utilisation courante des tests ADN par la police scientifique. On a pu rapidement s'affranchir de l'utilisation de la radioactivité grâce à la technique PCR (polymérase chain reaction) et normaliser l'étude des empreintes génétiques. Le principe en est simple : on détermine la distance entre deux points fixes de l'ADN encadrant le VNTR. Plus le nombre de répétitions est élevé plus la distance entre les points fixes est grande.

La validation de cette technique par la police a été réalisée dans le cas des crimes de Narborough en Angleterre. Depuis, on l'utilise régulièrement : la démarche est décrite dans la figure 3. L'empreinte génétique est obtenue entre 1 et 2

jours. Le coût en est maintenant d'une centaine d'euros.

Réalisation d'un profil génétique

Dès le début, on utilise dans tous les pays 15 VNTR répartis sur l'ensemble des chromosomes pour réaliser l'empreinte génétique. La tendance actuelle est d'augmenter le nombre de ces marqueurs jusqu'à 17. Ces marqueurs sont « anonymes » : cela signifie qu'ils ne sont pas dans des régions codantes de l'ADN à l'exception de celui qui sert à déterminer le sexe. Le problème de l'anonymat des marqueurs est soulevé car actuellement, avec les progrès de la génétique des populations, ces marqueurs peuvent indiquer une origine ethnique de l'ADN examiné.

La démarche est toujours la même : on récupère de l'ADN sur les lieux d'un crime à partir de traces de sang, de sperme, de mégot de cigarette, de cheveux ... Le profil obtenu est ensuite comparé à ceux présents dans une banque de données ou correspondant à de possibles coupables (Figure 3)⁶.

Le fichier national automatisé des empreintes génétiques (FNAEG)

Le fichier national des empreintes génétiques a été créé en 1998 (loi Guigou, affaire Guy Georges). Il contient à l'époque uniquement les profils ADN des auteurs de crimes sexuels. Depuis 2001 (loi Lebranchu) on procède à l'extension du fichier. A ce jour, il contient plus de 3 millions d'empreintes génétiques.

L'analyse des résultats doit différencier :

1) Une confirmation d'identité

Des preuves autres que l'ADN conduisent à suspecter des individus. Le test ADN confirme l'identité.

³ Protéine des muscles impliquée dans le transport et le stockage de l'oxygène.

⁴Un intron est un segment d'ADN copié dans l'ARN, puis éliminé, sans fonction apparente dans de multiples situations.

⁵ Jeffreys, A., Wilson, V. & Thein, S. Individual-specific 'fingerprints' of human DNA. *Nature* **316**, 76-79 (1985)

⁶ Locker D., Mémoires de l'Académie d'Orléans Les tests ADN, ou la science mène l'enquête, 47-62 (2010)



Figure 3. Protocole de récupération et d'utilisation des empreintes ADN

2) L'exclusion de suspects

La comparaison de l'empreinte ADN avec des banques de données conduira d'une part à exclure des suspects, d'autre part à chercher d'autres preuves pour prouver la culpabilité d'un suspect.

Dans le cas de correspondance d'un échantillon de la banque avec un prélèvement ADN sur les lieux du crime, il s'agit seulement d'une coïncidence statistique ! Attention au « sophisme du procureur » : c'est-à-dire la confusion entre « la probabilité qu'un événement survienne » et « la probabilité d'innocence de l'accusé ».

Exemple du cas Sally Clark⁷

On peut donner comme exemple le cas Sally Clark en Angleterre. Elle est avocate ; en 1996 son premier enfant est victime de mort subite du nourrisson puis en 1998 son deuxième enfant l'est lui aussi. Sally Clark est alors inculpée de double infanticide. En 1999, au cours de son

procès ; le Pr. Meadow, pédiatre, expert pour l'accusation, estime que la probabilité que deux enfants meurent de mort subite dans la même famille est égale à $1/64\,000\,000$ ($1/8000 \times 1/8000$). Cette probabilité est comprise comme une probabilité de son innocence et est tellement faible que Sally Clark est reconnue coupable. Elle est condamnée à la prison à perpétuité. En 2000, elle fait appel de cette condamnation ; l'appel est rejeté. En 2003, second appel, cette fois Sally Clark est enfin libérée.

Analyse d'un autre cas : l'affaire Diana Sylvester

L'affaire date de 1972 : une infirmière est violée et tuée par un homme inconnu. Des échantillons de sperme de l'assassin sont récupérés, l'empreinte génétique est comparée en 2003 à une liste de 380 000 empreintes génétiques de délinquants sexuels californiens. Elle correspond à celui d'un homme de 65 ans : John Puckett. Mais l'échantillon d'ADN prélevé en 1972 étant dégradé, seuls 5 marqueurs sur 15 sont

⁷ <https://lejournel.cnrs.fr/articles/de-erreur-de-calcul-a-erreur-judiciaire> De l'erreur de calcul à l'erreur judiciaire 2017

étudiés. La probabilité qu'une personne prise au hasard dans la population ait un tel profil est de 1 sur 1,1 million. Encore une fois, cette probabilité n'est pas une probabilité de culpabilité, au contraire : puisqu'il y a 300 millions d'habitants aux États-Unis, 300 environ ont un profil ADN compatible avec celui du coupable et la probabilité que l'un d'eux pris au hasard soit l'assassin est donc d'environ seulement 1 sur 300.

Pourquoi une empreinte ADN ne suffit – elle pas pour boucler une enquête ?

Certaines traces, comme un morceau d'ongle ou un cheveu sans bulbe, ne contiennent pas d'ADN nucléaire. Elles permettent en revanche d'isoler l'ADN mitochondrial. Cet ADN ne permet pas de déterminer clairement qu'une empreinte ADN appartient à un suspect, car le même ADN mitochondrial peut être partagé par plusieurs individus. En revanche, il peut permettre d'écartier la piste d'un suspect dont le génome mitochondrial ne correspond pas aux empreintes recherchées. L'analyse est tellement

puissante que les contaminations sont inévitables. L'ADN est parfois très dégradé et ne permet pas de faire une analyse correcte.

Conclusions sur l'analyse des tests ADN

Ceux-ci doivent être interprétés avec prudence. Il ne faut pas faire dire n'importe quoi à la statistique.

Evolution des analyses des empreintes ADN

Premier exemple: l'affaire Elodie Kulik

Cette infirmière a été violée et assassinée et des traces ADN des coupables ont été trouvées sur les lieux du crime. Sur les 6000 tests ADN réalisés on ne trouve aucune correspondance avec ces traces. De retour vers le FNAEG, les gendarmes ont l'idée de rechercher des profils ADN apparentés montrant par exemple 50% de correspondance avec l'ADN trouvé sur les lieux du crime. Cette recherche conduit en 2013 à



Figure 4. Comparaison des profils ADN de la famille Guerinoni avec le prélèvement fait à l'endroit du crime.

l'identification du père d'un individu ayant laissé une trace d'ADN sur les lieux du crime.

Deuxième exemple : le meurtre de Yara Gambirasio en Italie

L'affaire a conduit à la réalisation de 18 000 tests ADN. Yara Gambirasio disparaît le 26 novembre 2010 à l'âge de 13 ans à Brembate en Italie. Elle est retrouvée morte le 26 février 2011 dans un terrain vague. Les seuls indices trouvés sont : des traces de sang sur les sous-vêtements de la jeune fille (sang de l'inconnu n°1), de fines particules de plâtre et de ciment dans ses poumons ainsi que des fibres de tissu sur son corps.

Dans un premier temps les carabinieri concentrent leurs investigations sur les clients d'une boîte de nuit proche de la scène du crime, ils appréhendent un certain Damiano Guerinoni, dont l'ADN est proche de celui de "l'inconnu n°1". Toute la famille Guerinoni présente plusieurs marqueurs montrant une concordance avec le suspect sans nom (Figure 4). Pour les généticiens, les marqueurs d'origine pourraient venir du père, Giuseppe Guerinoni, chauffeur d'autobus mort en 1999. Par bonheur, la veuve Guerinoni a conservé le permis de conduire de son époux au dos duquel était collé un timbre

fiscal. La salive du timbre fiscal est analysée et on arrive à la conclusion suivante : il y a 99,999987% de probabilités que son fils soit l'« inconnu n°1 ». Mais un problème se pose : les enfants de Giuseppe Guerinoni sur lesquels des tests ADN ont déjà été effectués ne correspondent pas au profil génétique de l'« inconnu n°1 ». Les enquêteurs sont convaincus que l'assassin de Yara est un enfant illégitime de Giuseppe Guerinoni. La chance d'un policier conduit à un certain Giovanni Bossetti qui a vécu dans le même village que Giuseppe Guerinoni.

Le 15 juin 2014, au prétexte d'un test d'alcoolémie, les gendarmes font souffler G. Bossetti. L'examen fait sur sa salive démontre que son ADN correspond à 99,999987% à celui retrouvé sur le corps de Yara. L'« inconnu n°1 » pourrait être G. Bossetti. La marge d'erreur est de 0,000013%. Il faut alors rechercher d'autres éléments de preuves de la culpabilité de Bossetti. Rapidement on montre que les particules de plâtre/ciment trouvées sur Yara proviennent du chantier où il travaille et on apprend que sa camionnette aurait été vue plusieurs fois près du gymnase où s'entraînait Yara (sa fourgonnette a été filmée par des caméras de vidéosurveillance dans le quartier de l'adolescente). Il faisait de fréquents passages à Brembate di Sopra (ville où vivait Yara) pour revenir de son chantier à son domicile alors que cela rallongeait son chemin de 5 km. Dernier indice, on a retrouvé sur les

Haplotypes = un ensemble de marqueurs porté par un chromosome non séparable facilement par recombinaison et transmis en bloc

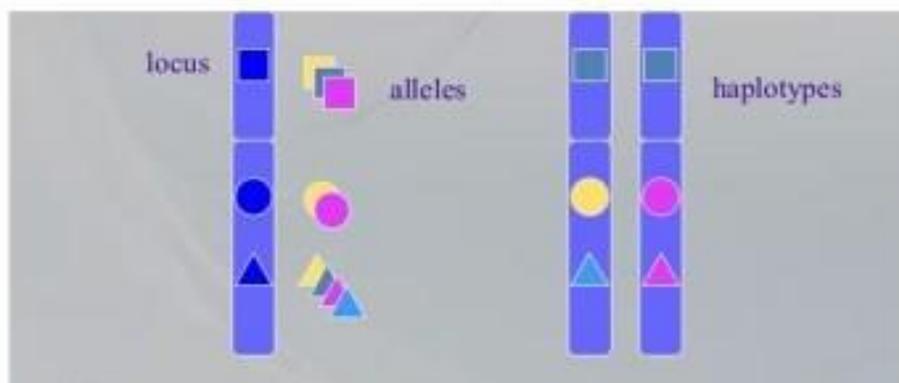


Figure 5. Les haplotypes

vêtements de Yara un fil de tissu identique à ceux des sièges de la camionnette de Massimo Bossetti.

Tests ADN et généalogie génétique⁸

Ces tests permettent de comparer les séquences d'ADN (les SNP) afin d'estimer la probabilité que des individus partagent un ou plusieurs ancêtres communs. Il existe plusieurs types de tests :

1. Les tests sur l'ADN-Y qui testent uniquement la lignée paternelle
2. Les tests sur l'ADN-mitochondrial qui testent la lignée maternelle
3. Les tests sur l'ADN autosomal qui permettent de faire de la généalogie

Ces tests permettent de découvrir à quel haplotype (Figure 5), c'est-à-dire à quel grand groupe génétique issu d'ancêtres communs, vous appartenez tout en vous donnant un pourcentage de vos origines génétiques. Représenté sur une carte du monde, le résultat permet de connaître le parcours migratoire de l'haplotype auquel vous appartenez, et donc celui de vos ancêtres.

Parenthèse sur la qualité des tests des origines pratiqués par les laboratoires privés : un exemple

Des tests ont été réalisés sur des triplées pour connaître leurs origines. Le test a montré que le pourcentage du patrimoine génétique français et allemand des triplées était complètement différent : 11% pour Nicole, 18% pour Jaclyn et 22,3% pour Erica. Comment trois sœurs ayant le même ADN pouvaient soudainement avoir des origines différentes ? On peut clairement douter du sérieux scientifique des analyses.

Utilisation des marqueurs haplotypes du chromosome Y et controverse autour de la tête d'Henri IV⁹



Figure 6. Joseph-Emile Bourdais brocanteur et propriétaire de la tête momifiée qu'il ne cesse d'attribuer à Henri IV.

En 1925, un article de la *Gazette des arts* présente un crâne momifié supposé être la tête du roi Henri IV (Figure 6).

En 2013, une controverse surgit autour de cette tête momifiée. Deux thèses s'affrontent alors : la première prétend qu'il s'agit de la tête du roi, la seconde réfute cette attribution. Dans un premier temps, les arguments pour l'attribution de ce crâne à Henri IV sont basés sur des analyses morphologiques. Les arguments tirés de l'analyse de l'empreinte ADN seront par la suite utilisés pour étayer cette attribution. En 2013, un article scientifique paru dans une revue à comité de lecture soutient que la comparaison des marqueurs ADN du chromosome Y du crâne d'Henri IV avec ceux du sang récolté au pied de l'échafaud de Louis XVI sont en accord avec une parenté entre les deux. Problème : l'ADN du crâne est très dégradé et l'accord n'est démontré que sur 5 marqueurs puis seulement 2 après un nouveau test. En 2014, un autre article scientifique montre que les marqueurs du chromosome Y du sang présumé de Louis XVI sont incompatibles avec les marqueurs retrouvés chez les descendants actuels de Louis XVI. Pour le moment le débat est loin d'être tranché et chacun campe sur ses positions.

⁸<https://journals.openedition.org/civilisations/3747>
2014

⁹https://www.lemonde.fr/sciences/article/2014/01/20/polemiques-sur-le-crane-suppose-d-henri-iv_4351224_1650684.html 2014

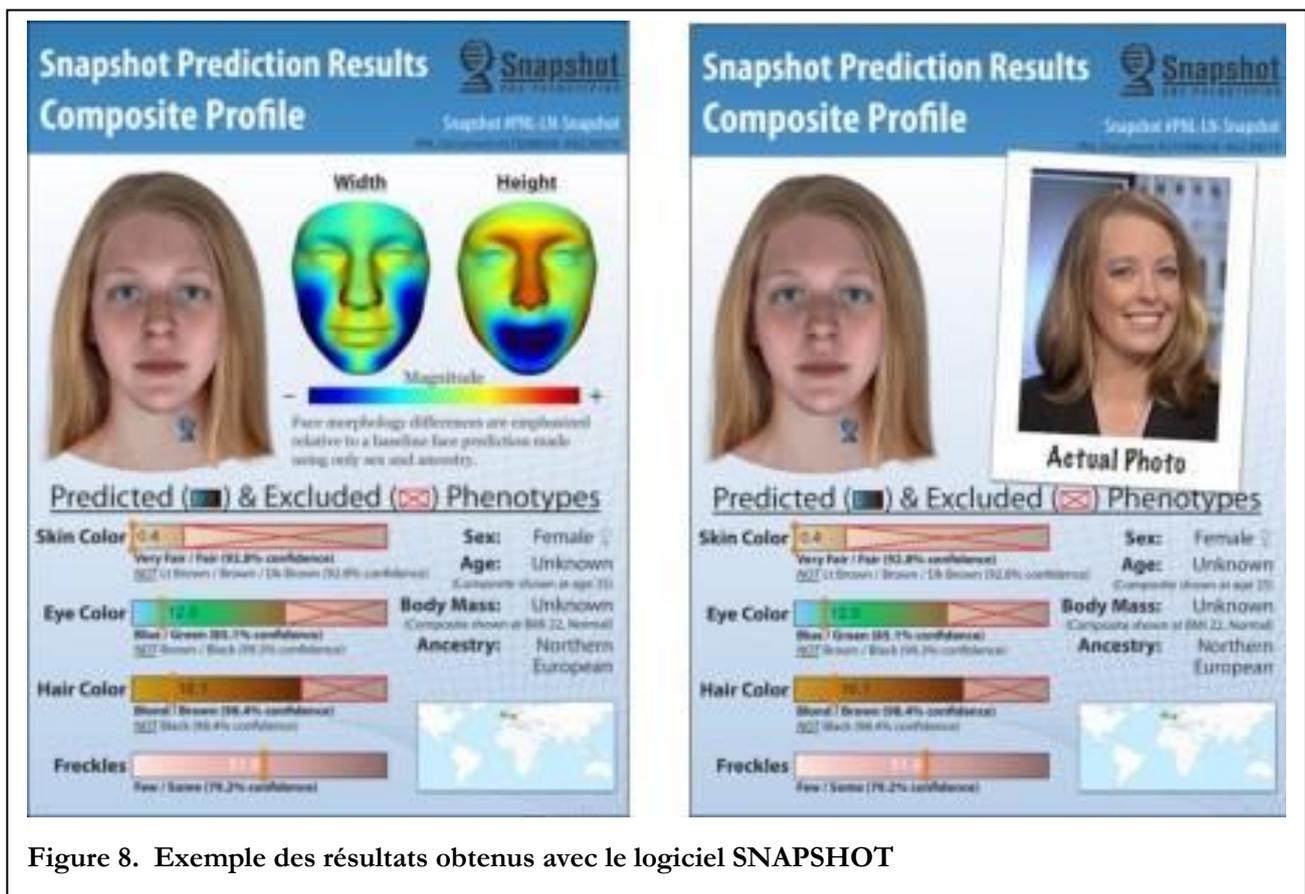


Figure 8. Exemple des résultats obtenus avec le logiciel SNAPSHOT

La Généalogie génétique vers la solution des « cold cases »

Nous prendrons comme premier exemple celui de l'affaire William Talbott. En 1987, deux canadiennes sont retrouvées assassinées avec pour seul indice, une trace de sperme sur les sous-vêtements d'une des jeunes filles. L'enquête a piétiné pendant plus de trente ans. Mais il a suffi de rentrer les données de l'ADN du sperme sur GEDmatch (banque ADN) pour trouver une correspondance avec deux cousins éloignés: l'arbre généalogique de la famille a mené l'équipe d'enquête directement à un chauffeur routier : William Talbott. Le 24 juillet 2019, Talbott a été condamné à une peine d'emprisonnement à perpétuité. Cette affaire est la première dans laquelle CeCe Moore a fourni le nom d'un suspect sur la base de recherches généalogiques approfondies (Figure 7).

Le second exemple est celui du Golden State Killer, ou du Tueur du Golden State (Etat de

Californie)¹⁰. C'est le cas d'un tueur en série américain qui a sévi dans les années 1970 et 1980. C'est seulement en avril 2018, une quarantaine d'années plus tard qu'un suspect a été appréhendé ; il s'agit de Joseph James De Angelo qui a fait des études de police scientifique et de droit criminel en Californie. Il s'est marié en 1973 avec Sharon Huddle, une avocate avec laquelle il a eu trois filles. La même année, il est entré dans la police d'Exeter, une petite ville à quelque 300 kilomètres au sud-est de San Francisco. Au cours des trois années qui ont suivi, plus d'une centaine de cambriolages ont été recensés. Les vols se sont arrêtés au moment où Joseph James De Angelo a déménagé à Auburn, une petite commune près de Sacramento où il s'est fait recruter de nouveau par la police locale. C'est là, à partir de juin 1976, qu'une série de viols particulièrement sordides a eu lieu. C'est l'ADN qui permet, en 2001, de relier le violeur et le cambrioleur et de comprendre qu'ils n'étaient qu'une seule et même personne. Les analyses de l'ADN du criminel avaient permis

¹⁰ <https://www.rfgenealogie.com/s-informer/infos/nouveautes/un-nouveau-cold-case->

[resolu-par-la-genealogie-genetique-aux-etats-unis-2018](https://www.rfgenealogie.com/s-informer/infos/nouveautes/un-nouveau-cold-case-)

dès 1986 de découvrir un détail important : il renferme un marqueur rare signifiant que l'individu porteur de ce marqueur est « non-sécréteur » (les antigènes correspondant à son groupe sanguin ne se retrouvent pas dans les fluides corporels). Par ailleurs, le tueur du Golden State a une forme du gène PGM (phosphoglucomutase) plutôt inhabituelle. Or on sait que seulement 1 à 2% de la population détiendrait la combinaison non-sécréteur avec ce type de PGM.

À partir de 2016, le FBI promet 50 000 \$ à qui aurait des informations menant à l'arrestation du tueur. Le « plus froid des cold cases » est rouvert en 2016 par la procureure du comté de Sacramento après 40 ans de recherches infructueuses.

Les enquêteurs utilisent des échantillons d'ADN recueillis sur les lieux des crimes pour déterminer le profil génétique du tueur, puis entrent ce profil dans une base de données généalogiques en ligne, GEDmatch. Ils examinent ensuite avec des généalogistes les familles afin de trouver des correspondances avec ce profil. Si le tueur n'a pas transmis son ADN sur Internet, au moins un membre de sa famille l'a fait.

C'est ce qui a permis de montrer la correspondance entre l'ADN trouvé sur les lieux des crimes et celui de Joseph DeAngelo

En mai 2018, Parabon NanoLabs a nommé la généalogiste CeCe Moore à la tête de son unité de généalogie génétique. En novembre 2018, Parabon déclarait travailler sur 200 cas; 55% d'entre eux avaient produit des pistes et, en mai 2019, ils résolvaient des « cold case » à raison d'un par semaine.

En mai 2019, GEDmatch a demandé aux personnes qui avaient téléchargé leur ADN sur son site de s'inscrire spécifiquement pour autoriser les services répressifs à accéder à leurs informations.

Les nouveaux outils de la police scientifique pour confondre les coupables

Snapshot DNA Phenotyping Service¹¹ (Figure 8) est le nom d'un outil de phénotypage d'ADN développé par Parabon NanoLabs qui permet de créer des schémas composites d'imagerie du visage à partir d'échantillons d'ADN. Outil très contesté par les scientifiques. Il a été rapporté que les prédictions de couleur de peau et des yeux étaient assez fiables mais pas les prédictions de la forme du visage.

Conclusion

Actuellement tous ces tests ne sont faits que sur une partie de l'ADN. Que nous réserve l'avenir ? Un séquençage individuel de la totalité de notre ADN ?

Remerciements

Michel Monsigny a lu et corrigé ce texte, je lui en suis très reconnaissant

¹¹ <https://snapshot.parabon-nanolabs.com/> 2020