Chroniques génomiques

Nous avons les moyens de faire parler votre ADN...

Bertrand Jordan

CoReBio PACA,
case 901, Parc scientifique
de Luminy, 13288 Marseille
Cedex 9, France.
brjordan@club-internet.fr

Des empreintes génétiques... au phénotypage policier

Depuis quelque temps, les analyses d'ADN à visée criminologique prennent un nouveau tournant. L'utilisation des empreintes ADN, fondées sur l'analyse d'une série de régions hypervariables (*short tandem repeats* ou STR¹), est largement entrée dans les mœurs, même si son extension à des suspects ou même à de simples témoins reste, à juste titre, très discutée. Cette relative acceptation doit beaucoup à deux caractéristiques de ces analyses. Tout d'abord, leur fiabilité : si la technique est correctement pratiquée, en examinant un nombre suffisant de STR (le standard actuel est de treize), elle est fiable à 100 % en exclusion (cet ADN ne provient pas de cette personne) et à presque 100 % en inclusion. Ce sont là des résultats sans ambiguïté, ne laissant quasiment pas de place à l'interprétation et directement utilisables dans le cadre juridique. Mais l'acceptabilité des empreintes génétiques doit aussi beaucoup au fait qu'elles portent uniquement sur des régions non codantes de l'ADN et sont censées ne donner aucune information sur la personne dont provient le prélèvement – c'est d'ailleurs inscrit dans la loi, en France comme dans la plupart des nations. En ce sens, le parallèle avec les empreintes digitales est pertinent, ce qui a largement contribué à un accord général sur l'inclusion de ces analyses dans la panoplie policière et judiciaire.

Mais une évolution apparaît depuis plusieurs années, avec déjà quelques cas largement médiatisés : la possibilité de « faire parler » un ADN inconnu, d'en tirer des indications sur le phénotype et en particulier l'apparence de son possesseur, et d'orienter ainsi la recherche de suspects. Dans la littérature anglo-saxonne, on parle de *forensic DNA phenotyping* (FDP), phénotypage criminologique à partir de l'ADN, et de détermination des *externally visible characters* (ECV), caractères visibles externes. L'intérêt de ces techniques pour les enquêteurs est évident, tout comme l'est le fait qu'elles posent de sérieux problèmes éthiques. Cette chronique vise, à partir d'un récent article de *Nature Reviews Genetics* [1], à faire le point sur les possibilités actuelles de ces approches, leur développement probable et les questions sociétales qu'elles soulèvent.

De l'ascendance au phénotype : une fausse piste ?

On sait depuis au moins cinq ou six ans que l'analyse fine d'un ADN humain à l'aide d'un jeu de snip² permet de rattacher son possesseur à un (ou éventuellement plusieurs) groupes d'ascendance, en somme d'indiquer si ses ancêtres sont majoritairement Africains, Asiatiques, Européens ou Amérindiens [2]. Une analyse plus poussée permet même des différenciations à l'intérieur de ces grands groupes géographiques, entre Italiens et

¹ Séquences d'ADN contenant un nombre très variable de courtes séquences répétées en tandem.

² *Single nucleotide polymorphisms*, SNP ou snip, polymorphismes portant sur un seul nucléotide, très fréquents dans notre génome et très utilisés dans les études génétiques actuelles.

Figure 1. Prédiction de la couleur des yeux selon le système IrisPlex. Cette figure montre des yeux humains rangés selon la prédiction de couleur fondée sur l'ADN et établie avec le système IrisPlex. Elle débute avec la plus forte probabilité pour le bleu à gauche et se termine avec la plus forte probabilité pour le brun à droite. Noter que l'on n'a pas utilisé pour ce classement la couleur perçue, uniquement la probabilité établie par IrisPlex. Ce n'est bien sûr qu'un exemple de la manière dont le système IrisPlex peut être utilisé en analyse criminologique (© Forensic Molecular Biology, Erasmus MC-University Medical Center Rotterdam).

Allemands, par exemple [3, 4], à condition que l'on se limite aux véritables autochtones³. On peut donc imaginer que l'analyse d'un nombre suffisant de snip (non codants dans leur très grande majorité) permette de rattacher l'ADN d'une personne à un groupe géographique et ce faisant d'indiquer, dans les grandes lignes, quelle peut être son apparence. On obtiendrait ainsi des indications phénotypiques à partir d'analyses portant sur des régions non codantes, en respectant la lettre sinon l'esprit de la loi. De telles études ont effectivement été réalisées, notamment en 1999, par l'un des auteurs de l'article cité [1]. Il s'agissait dans ce cas d'un viol perpétré au voisinage d'un foyer d'immigrés, aux Pays-Bas, et l'analyse d'ADN a indiqué, contrairement à la rumeur populaire, que le coupable était d'ascendance Nord-Européenne. La police, aux États-Unis, a eu assez largement recours à ces analyses, et une entreprise appelée *DNAPrint* – aujourd'hui disparue – en a fait un de ses produits commerciaux.

Cette approche, qui peut paraître séduisante, présente en fait de sérieuses faiblesses. L'hétérogénéité interne des groupes d'ascendance – rappelons que 80 % de la variabilité génétique humaine se retrouve à l'intérieur de chaque groupe – fait que l'appartenance à l'un d'eux n'est que faiblement prédictive pour l'aspect de la personne. Et la prévision est encore plus problématique lorsque les ascendances sont mixtes, ce qui est de plus en plus fréquent ! Les auteurs [1] plaident donc la prudence et recommandent de n'y recourir que lorsque l'analyse d'ADN indique une origine géographique unique (au niveau des grands groupes continentaux), et seulement pour avoir une idée générale de l'apparence de la personne en cause.

Les « gènes de l'apparence »

L'idéal bien sûr serait d'examiner directement, dans l'ADN du prélèvement, les gènes qui déterminent les caractères visibles externes. Le problème, c'est que les caractères en cause sont fortement multigéniques et que, comme nous l'avons souvent souligné [5], la valeur prédictive des gènes identifiés par les « balayages du génome » (GWAS, *genome-wide association studies*) est généralement assez faible. Néanmoins, la détermination à partir de l'ADN de certains éléments atteint déjà une assez bonne fiabilité. En premier lieu, le sexe, s'il n'est pas déjà connu, peut naturellement être détecté grâce à des marqueurs du chromosome

Y. La couleur des yeux peut être prédite de manière relativement satisfaisante grâce à la détermination d'une quinzaine de snip [6]. La qualité de la détermination est mesurée par un paramètre appelé AUC – égal à 1 en cas de pleine exactitude et à 0,5 pour un résultat au hasard⁴ – et les valeurs obtenues sont supérieures à 0,90, indiquant donc une bonne fiabilité. Un test commercial appelé *IrisPlex* utilise les six snip les plus prédictifs, exploités grâce à un modèle statistique et une base de données associée, ce qui permet une prévision d'assez bonne qualité (AUC proche de 0,9) et insensible à l'origine géographique (Figure 1). Notons au passage que, comme pour d'autres caractères externes, la définition de la couleur des yeux est fortement subjective et que les déclarations de témoins sur ce point sont très peu fiables. La couleur des cheveux peut également être évaluée par un test analogue portant sur onze gènes, avec des valeurs d'AUC qui vont de 0,8 à 0,9, sauf pour les cheveux roux (associés à une peau très claire) qui dépendent essentiellement d'un seul gène, appelé *MC1R*, et sont donc facilement repérables. Pour la couleur de la peau, les choses sont nettement plus compliquées. La recherche des gènes qui l'influencent ne peut se faire qu'à l'intérieur de chaque groupe d'ascendance, pour éviter les innombrables faux positifs qui apparaîtraient sinon, en raison des différences de fréquence d'allèles entre groupes. Chaque étude ne couvre donc qu'une partie de la gamme observée au sein de l'ensemble de la population et, à l'heure actuelle, les snip connus ne rendent compte que d'un peu plus de la moitié de la variabilité existante. La taille, caractère fortement héréditaire, est sous l'influence d'un nombre très élevé de gènes – on en est à près de deux cents – et sa prédiction à partir de l'ADN a une AUC de 0,65, guère plus que la valeur de 0,5 qui correspond au hasard. La forme

³ Dans l'étude citée, les personnes dont les parents et grands-parents habitaient le même pays.

⁴ Le paramètre AUC (*area under receiver characteristic operating curve*) est tiré d'une représentation graphique de la qualité de la prédiction.

Figure 2. Estimation de l'âge en fonction de l'analyse ADN. Corrélation entre l'abondance relative des cercles d'excision (en abscisse, échelle logarithmique par rapport au gène de l'albumine) et l'âge de la personne dont provient l'ADN (en ordonnée) (adapté de [6]).

du visage, enfin, est sous contrôle génétique, comme en témoignent l'observation courante mais aussi l'extrême ressemblance entre vrais jumeaux ; plusieurs équipes tentent actuellement d'identifier des gènes impliqués dans sa détermination, mais la question est à l'évidence d'une grande complexité et les perspectives d'utilisation en criminologie encore très lointaines. L'idée d'un portrait robot déduit de la seule analyse de l'ADN reste donc du domaine de la science-fiction.

Une surprise...

De manière assez étonnante, il semble qu'il soit possible d'évaluer l'âge d'une personne à partir de son ADN ! Et la technique la plus performante ne repose pas, comme on pourrait l'attendre, sur l'analyse d'altérations de l'ADN mitochondrial ou de la taille des télomères, mais sur la mesure du nombre de produits de réarrangement (cercles d'excision) des gènes du récepteur T⁵, qui diminue avec l'âge et l'involution du thymus, organe où ont lieu ces événements. Selon les dernières études [7], on peut ainsi prédire la catégorie d'âge (par groupe de 20 ans, correspondant approximativement à une génération) avec des valeurs d'AUC supérieures à 0,9, donc de manière assez fiable (Figure 2).

Des possibilités séduisantes, des dérapages très probables

Il s'avère donc possible, à partir de l'analyse d'une trace d'ADN, d'obtenir une idée de l'apparence de la personne concernée, et d'orienter ainsi la recherche d'un suspect – ou d'ailleurs l'identification de restes humains suite à une catastrophe. Ces déductions s'appuient dès à présent sur l'analyse de gènes directement impliqués dans des caractères visibles, et non seulement sur les indications assez vagues données par l'appartenance à un groupe d'ascendance. On peut prévoir, sans risque d'être démenti, que ces prévisions vont rapidement s'affiner, gagner en crédibilité et s'étendre, à terme, à des caractéristiques aujourd'hui inaccessibles. Il reste certes de sérieux problèmes techniques, notamment la fiabilité de ces méthodes en présence d'échantillons dégradés ou infinitésimaux, ou encore la qualité des déductions lorsqu'elles sont appliquées à un ADN provenant d'un groupe d'ascendance différent de celui dans lequel la corrélation a été établie, mais tout ceci va certainement s'améliorer au fil des mois et des années.

S'améliorer – ou empirer ? Le plus n'est pas forcément le mieux en ce domaine, et l'on peut facilement envisager des améliorations fort problématiques. Lorsque l'on cherche à déduire un phénotype à partir de l'ADN, rien n'oblige à se limiter à l'apparence : on peut tout aussi bien s'intéresser à des informations d'ordre médical ou même à des prédispositions d'ordre comportemental. Le fait que, dans un procès récent en Italie [8], une réduction de peine ait été demandée et obtenue sur la base d'une (prétendue) prédisposition génétique à l'agressivité⁶ montre qu'il ne s'agit pas là d'un pur fantasme. La description du suspect à partir de son ADN pourrait ainsi inclure, par exemple, le fait qu'il soit probablement diabétique, ou qu'il présente un comportement très impulsif.

Mais l'aspect le plus préoccupant est sans nul doute le risque de surinterprétation, risque présent dès que l'on parle d'ADN, qui évoque inmanquablement le fantasme du « destin inscrit dans les gènes », mais encore augmenté dans le cas présent. Les empreintes ADN telles qu'elles ont été mises en œuvre jusqu'ici fournissent des résultats très dichotomiques et ne comportant guère de latitude d'interprétation, et l'on tendra tout naturellement à attribuer la même certitude aux indications d'apparence, alors que la prévision d'un phénotype à partir de l'ADN est et restera à coup sûr de nature probabiliste. Ses indications sont par nature approximatives, avec une marge d'erreur plus ou moins large mais jamais négligeable, or le public, comme les institutions policières et judiciaires, risque fort de leur accorder un grand crédit et, à la limite, d'en faire un instrument de preuve – alors qu'elles devraient être limitées à un rôle en tant qu'outil d'investigation. Une fois de plus, une bonne compréhension des limites de la technique tout comme un encadrement de ses utilisations s'avèrent indispensables pour limiter les dérives. ♦

Your DNA will talk to us!

⁵ Cercles d'excision de la jonction signal ou sjTREC, témoignant de la production récente de nouveaux lymphocytes T par le thymus.

⁶ Présence d'allèles des gènes *MAOA*, *COMT*, *SCL6A4* et *DRD4* dont certaines études suggèrent qu'ils accroissent le risque de telles conduites.

CONFLIT D'INTÉRÊTS

L'auteur déclare n'avoir aucun conflit d'intérêts concernant les données publiées dans cet article.

REMERCIEMENTS

Nous remercions le Pr Manfred Kayser (Department of Forensic Molecular Biology, Erasmus MC - University Medical Center Rotterdam) de nous permettre d'utiliser la Figure 1 pour illustrer cette chronique génomique.

RÉFÉRENCES

1. Kayser M, de Knijff P. Improving human forensics through advances in genetics, genomics and molecular biology. *Nat Rev Genet* 2011 ; 12 : 179-92.
2. Jordan B. *L'humanité au pluriel*. Paris : Seuil, 2008.
3. Lao O, Lu TT, Nothnagel M, et al. Correlation between genetic and geographic structure in Europe. *Curr Biol* 2008 ; 18 : 1241-8.
4. Novembre J, Johnson T, Bryc K, et al. Genes mirror geography within Europe. *Nature* 2008 ; 456 : 98-101.
5. Jordan B. À la recherche de l'héritabilité perdue. *Med Sci (Paris)* 2010 ; 26 : 541-3.
6. Walsh S, Liu F, Ballantyne KN, et al. IrisPlex: a sensitive DNA tool for accurate prediction of blue and brown eye colour in the absence of ancestry information. *Forensic Sci Int Genet* 2011 ; 5 : 170-80.
7. Zubakov D, Liu F, van Zelm MC, et al. Estimating human age from T-cell DNA rearrangements. *Curr Biol* 2010 ; 20 : R970-1.
8. Forzano F, Borry P, Cambon-Thomsen A, et al. Italian appeal court: a genetic predisposition to commit murder? *Eur J Hum Genet* 2010 ; 18 : 519-21.

TIRÉS À PART

B. Jordan

Collection SCIENCE ET BIOMÉDECINE

ISBN : 2-84254-107-3 64 pages

ISBN : 2-84254-108-1 80 pages

ISBN : 2-84254-111-1 86 pages

Bon de commande

À retourner à EDK, 2, rue Troyon - 92316 Sèvres Cedex
Tél. : 01 55 64 13 93 - Fax : 01 55 64 13 94 - E-mail : edk@edk.fr

NOM : Prénom :
Adresse :
Code postal : Ville :
Pays :
Fonction :

Je souhaite recevoir l'ouvrage **Les oligo-éléments** : 10 € + 3 € de port = **13 € TTC**

Je souhaite recevoir l'ouvrage **Acides gras, acides aminés et peptides** : 12 € + 3 € de port = **15 € TTC**

Je souhaite recevoir l'ouvrage **Stress oxydatif et alicaments** : 14 € + 3 € de port = **17 € TTC**

en exemplaire, soit un total de €

☐ Par chèque, à l'ordre de **EDK**

☐ Par carte bancaire : ☐ Visa ☐ Eurocard/Mastercard

Carte n°

Signature :

Date d'expiration :

N° de contrôle au dos de la carte :