

## Nouveaux usages de l'ADN en matière pénale : promesses, savoirs et pratiques

Catherine Bourgain

---



### Édition électronique

URL : <http://journals.openedition.org/cdst/1081>

DOI : 10.4000/cdst.1081

ISSN : 2431-8663

### Éditeur

Presses universitaires d'Aix-Marseille - PUAM

### Édition imprimée

Date de publication : 25 octobre 2019

Pagination : 83-93

ISBN : 978-2-7314-1144-7

ISSN : 1967-0311

Ce document vous est offert par Centre national de la recherche scientifique (CNRS)



### Référence électronique

Catherine Bourgain, « Nouveaux usages de l'ADN en matière pénale : promesses, savoirs et pratiques », *Cahiers Droit, Sciences & Technologies* [En ligne], 9 | 2019, mis en ligne le 31 octobre 2019, consulté le 17 février 2020. URL : <http://journals.openedition.org/cdst/1081> ; DOI : 10.4000/cdst.1081

---

# Nouveaux usages de l'ADN en matière pénale : promesses, savoirs et pratiques

Catherine BOURGAIN\*

---

**L**A GÉNÉTIQUE est un domaine fortement enclin aux promesses. Son usage dans les contextes policier et judiciaire ne fait pas exception. Il est ou sera ainsi possible, nous dit-on, d'identifier toute personne qui aurait juste effleuré une table ou un corps. Un individu, encore jamais mis en cause ou condamné, peut ou pourra être confondu par l'entremise de l'ADN d'un membre de sa famille éloignée. Les origines et le portrait-robot d'un suspect peuvent ou pourront être dressés à partir d'un simple échantillon biologique.

Pour qualifier cette tendance, caractéristique notamment des biotechnologies, à recourir de façon intense aux anticipations dans le gouvernement de la recherche scientifique et technique, P.B. Joly (2015) a introduit le concept de régime des promesses scientifiques. Dans ce régime, les promesses sont d'autant plus importantes que l'innovation considérée nécessite la mobilisation de ressources importantes pour être financée, pour implémenter les adaptations réglementaires requises, pour concevoir les cadres politiques adéquats, ou encore pour fabriquer de nouveaux usages et de nouveaux utilisateurs. L'insistance sur la gravité des enjeux à l'œuvre, et souvent sur la compétition économique internationale, conduit en outre à imposer une logique d'urgence.

Ce régime se caractérise également par un optimisme technologique (Hjorlieffson *et al.*, 2008) très fort. Les incertitudes sont largement sous-estimées, voire négligées. Au contraire, les potentialités à court et moyen termes des technologies sont surestimées, et leurs valeurs sur le temps long restent le plus souvent incomprises (Brown, 2003). Ces promesses génériques ne sont que rarement soumises à des tests de validité.

La présente contribution vise à discuter de nouveaux usages et pratiques de la génétique en contextes policier et judiciaire, associés aux développements technologiques et aux évolutions réglementaires récents. Il s'agit, après avoir présenté quatre usages de ce type – les recherches dites familiales, la caractérisation d'ADN de contact, la détermination des origines et celle de caractéristiques apparentes à partir de l'ADN –, d'analyser les déplacements de savoirs et de pratiques qui les caractérisent. Nous montrons notamment une montée en puissance de l'incertitude au cœur du travail de production et d'interprétation de l'information moléculaire. La forme d'analyse génétique, massivement pratiquée depuis la mise au point des premières empreintes génétiques au milieu des années 1980, vise à produire et à comparer des données génétiques

---

\* Université de Paris Descartes, Cermes3, Inserm, Cnrs, EHESS.

considérées comme simples, solides et univoques, dès lors que les précautions suffisantes ont été prises pour éviter les contaminations ou les erreurs de manipulation. Si la réalité est souvent plus complexe, notamment lorsque les ADN à analyser sont dégradés ou mélangés, la fiction selon laquelle la fonction du laboratoire de biologie est de révéler une information génétique indiscutable est très largement partagée par de nombreux acteurs. Au contraire, avec les nouveaux usages, l'existence d'incertitudes irréductibles ne peut plus être ignorée. Des procédures *ad hoc* doivent être adoptées pour les mesurer, en tenir compte et les communiquer. Le recours à des méthodes probabilistes, adossées à des modélisations statistiques complexes, en est un marqueur, tout comme l'importance accrue du travail requis au laboratoire pour interpréter les données.

Ces liens intimes entre savoirs et technologies de la génétique et incertitude, loin de constituer une exception, rapprochent ces usages de la génétique en contextes policier et judiciaire de nombreux autres contextes, en premier lieu les mondes biomédicaux. Il ne manque toutefois pas d'interroger les promesses, les simplifications qui les sous-tendent et la pertinence des investissements et des évolutions réglementaires qui sont décidées en leur nom.

84

## I. LES NOUVEAUX USAGES

Les grands programmes d'investissement public dans la génétique, au premier rang desquels le Projet Génome Humain, ont été des accélérateurs d'innovation en matière de description moléculaire de l'ADN (Hilgartner, 2017). De nouvelles machines capables de décrire l'enchaînement précis des nucléotides qui constituent une séquence d'ADN ont été mises au point, en lien avec une automatisation des étapes de préparation des échantillons biologiques. Ces évolutions sont également marquées par un changement de stratégie dans la démarche de production des données. L'approche historique visait à obtenir des données précises par une amélioration graduelle des technologies d'analyse biologique. Elle a été peu à peu remplacée par une approche où la précision résulte de la production massive de données incomplètes, corrigées par des algorithmes statistiques adossés à une puissance de traitement informatique nouvelle. Il est ainsi devenu envisageable de caractériser exhaustivement l'ADN d'un grand nombre d'individus dans un temps court, comme de le faire à partir de quantités infimes de matériel biologique, en partie dégradé, dès lors que les capacités technologiques et l'expertise en matière d'analyse statistique et bioinformatique sont disponibles. Au-delà de ces prouesses au-delà de ces prouesses technologiques en matière d'exhaustivité et de traitement de signaux très faibles, ces innovations ont contribué à une augmentation notable de la productivité de toutes les formes d'analyses génétiques, y compris les plus simples, qui sont aujourd'hui réalisées de façon massive, rapide et à moindre coût, avec comme corollaire la constitution de grandes banques de données génétiques à usages et visées multiples.

Pour présenter, dans les paragraphes qui suivent, les nouveaux usages de l'analyse génétique concomitants de ces évolutions technologiques, deux registres sont

successivement considérés. Dans le premier, c'est la capacité de l'analyse génétique à identifier les individus en les distinguant les uns des autres qui est centrale. L'ADN est mobilisé comme un traceur. Dans le second, l'analyse vise à dire et à décrire l'individu physique absent à partir de sa trace génétique. L'ADN est utilisé comme un prédicteur.

## A. LES EXTENSIONS D'USAGES DE L'ADN TRACEUR

L'ADN traceur est la forme historique de l'analyse génétique en contextes policier et judiciaire. La mise au point des premières techniques d'empreintes génétiques (Jeffreys, 1985) constitue en effet la démonstration qu'il est possible d'identifier des signatures génétiques uniques à chaque individu, en utilisant des procédures relativement simples et sans avoir à analyser l'intégralité de l'ADN. Quelques segments génétiques bien choisis peuvent suffire. Cette économie de moyens a été déterminante dans la forte montée en puissance de ces usages dans les années qui ont suivi, et notamment pour la mise en place de larges banques et fichiers d'empreintes génétiques. Deux évolutions récentes de ces usages d'ADN traceur sont successivement considérées ici. La première, qualifiée de recherche familiale, consiste, en l'absence du profil génétique d'un individu recherché, à utiliser ceux de personnes qui lui sont biologiquement apparentées. Cette fonction d'intermédiaire, jouée par la famille, capitalise sur le fait que l'ADN est transmis au cours de la reproduction sexuée, et qu'en conséquence les signatures génétiques sont en partie partagées entre apparentés. La seconde évolution consiste à reconstruire des profils génétiques à partir d'ADN dits de contact, présents en quantité extrêmement minime.

85

*Les recherches familiales* reposent sur un principe simple : l'ADN nucléaire d'un embryon est constitué pour moitié de l'ADN du spermatozoïde et pour moitié de celui de l'ovule qui ont fusionné. Un enfant partage donc toujours précisément 50 % de son ADN avec son père biologique et 50 % avec sa mère biologique. En chaque point de l'ADN, il possède une copie qu'il a reçue de son père et une copie qu'il a reçue de sa mère. Pour conclure à un lien filial entre deux individus à partir de l'analyse de leurs empreintes génétiques il faut, pour chacun des segments qui les composent, que les deux profils génétiques présentent au moins un allèle identique. Le fait de partager les mêmes allèles peut résulter du hasard. Lorsqu'un allèle est fréquent, des individus peuvent le partager sans être apparentés. Le nombre et le choix des segments actuellement utilisés pour réaliser les empreintes permettent toutefois, lorsqu'ils sont tous informatifs, de contrôler ces situations. Les risques de conclure faussement à un lien de filiation directe et, encore plus, de l'exclure à tort, sont rendus négligeables.

La situation est plus complexe entre deux enfants d'un même couple biologique. Parce qu'il n'y a pas de transmission directe d'ADN entre eux et que deux séries d'événements aléatoires (les méioses) sont à l'origine respectivement de chacun des enfants, la proportion d'ADN qu'ils partagent n'est pas fixe. La théorie génétique prédit qu'en moyenne cette proportion est de 50 % et que, pour un segment donné, 25 % des frères et sœurs portent les mêmes deux allèles de ce segment (l'un provenant du père et l'autre de la mère), 50 % ne portent qu'un seul même allèle (du père ou de la mère)

et 25 % ne partagent aucun allèle commun. La comparaison d'empreintes requiert donc une double prise en compte du hasard : hasard conduisant deux individus à porter les mêmes allèles en l'absence d'apparentement direct et hasard conduisant deux individus apparentés à partager plus ou moins d'allèles. Avec les mêmes segments, les incertitudes sont donc toujours plus importantes pour un test de fraternité que pour un test de paternité. De façon générale, plus l'apparentement à tester entre deux personnes est éloigné et plus l'incertitude est élevée.

Disposer d'une description plus fine et exhaustive de l'ADN peut contribuer à réduire ces incertitudes. Des données sur le génome entier permettent de proposer des tests d'apparentement solides pour des cousins au deuxième (issus de germains), voire au troisième degré (Erlich, 2018). L'approche requiert toutefois d'utiliser des modèles statistiques pour tenir compte de tous les moments de « tirage au sort » qui jalonnent la généalogie reliant les individus. Ces modèles doivent également intégrer le fait qu'une histoire ancienne commune – appartenir à la même *population* dans le vocabulaire de la génétique – participe de cette ressemblance entre individus.

86

Ces recherches familiales sont ainsi particulièrement sensibles au choix et à la qualité des bases de données de référence utilisées pour modéliser les ressemblances attendues en l'absence d'apparentement. Si la population de référence choisie représente mal la population d'origine des individus, un apparentement peut être jugé plus probable qu'il ne l'est et conduire à des enquêtes autour d'individus non apparentés (Rohlf, 2012). Le choix des modèles statistiques utilisés et des hypothèses qui les fondent a également un impact important sur les conclusions des tests, certaines méthodes présentant un biais qui conduit à conclure trop souvent, à tort, à l'existence d'apparentements (Balding *et al*, 2013). Ces enjeux méthodologiques qui combinent des choix de méthodes et des questions autour des banques de données disponibles font l'objet de débats assez vifs en lien avec les risques de discrimination associée (Garrison *et al*, 2013).

La seconde grande évolution d'usages en matière d'ADN traceur est liée à l'amélioration de la sensibilité de détection de l'ADN dans les échantillons biologiques recueillis. Il est aujourd'hui possible de proposer une analyse génétique avec des quantités infimes d'ADN, correspondant à une dizaine de cellules (moins de 100 pico-grammes), voire même à une unique cellule, pouvant avoir été déposées par un simple contact. Cette montée en puissance des *ADN dits de contact* va de pair avec une augmentation notable des échantillons collectés. Entre 2004 et 2008, la proportion d'échantillons soumis pour analyse génétique en Belgique et contenant moins de 100 pico-grammes serait ainsi passée de 19 à 45 %. Mais seuls 10 % des plus de 9 500 échantillons considérés comme correspondant à des ADN de contact, récoltés entre 2007 et 2011 à New York, ont effectivement conduit à des profils génétiques utilisables (Butler, 2016). En effet, ces faibles quantités posent des problèmes complexes. Les ADN recueillis sont souvent dégradés. Ils correspondent régulièrement à des mélanges d'ADN hétérogènes provenant d'individus différents. Enfin, les instruments et procédures utilisés pour la caractérisation des segments peuvent avoir des comportements inattendus lorsque les

quantités à analyser sont très faibles. Des variations aléatoires peuvent ainsi conduire à ne pas détecter des allèles pourtant bien présents dans l'échantillon et habituellement systématiquement détectés. C'est alors la fiabilité première des données « brutes » qui est mise en cause par l'irruption de cette nouvelle forme d'incertitude logée au cœur de leur production.

De nouvelles approches, développées pour tenter de produire des résultats tenant compte de ces différents niveaux de complexité, proposent de produire un « génotypage en probabilités » – en anglais *probabilistic genotyping* (Balding *et al.*, 2009). Il s'agit non plus de produire des profils génétiques dans lesquels les allèles des segments sont présents ou absents, mais des profils avec une probabilité pour chaque allèle d'être présent ou absent. Ces méthodes reposent notamment sur des modélisations statistiques assez complexes et le recours à des simulations informatiques importantes. Comme pour les recherches familiales, les résultats sont variables selon les logiciels utilisés. Si certains, développés dans des contextes académiques, sont librement accessibles, d'autres, vendus par des sociétés privées, sont plus opaques. En 2016, la Société internationale de génétique médico-légale (ISFG, International Society for Forensic Genetics) a ainsi publié des recommandations sur les procédures de validation des logiciels proposant des calculs biostatistiques pour les analyses génétiques (Cobe *et al.*, 2016). Cette position de l'ISFG insiste sur deux niveaux de responsabilité : celle des développeurs et fournisseurs de logiciels, mais aussi celle des utilisateurs de ces logiciels invités à procéder à différents contrôles « internes ». John M. Butler, assistant pour les questions médico-légales du directeur du National Institute of Standards and Technology américain, et figure influente du milieu, appelle ainsi les scientifiques qui utilisent ces logiciels « à ne pas les considérer comme des boîtes noires produisant des résultats qu'ils sont incapables de comprendre » (Butler, 2015a).

## B. L'INTRODUCTION DE NOUVEAUX USAGES D'ADN PRÉDICTEUR

La seconde catégorie de nouvelles formes d'usages de l'ADN en contextes policier et judiciaire constitue une évolution notable dans ce champ. L'objectif n'est plus d'identifier les individus en les distinguant les uns des autres par la comparaison d'empreintes génétiques, mais en prédisant certaines de leurs caractéristiques, au premier rang desquelles les origines et certains traits morphologiques apparents.

Les allèles des différents segments et variations de l'ADN ne sont pas présents dans toutes les populations humaines, et leurs fréquences varient d'une population à l'autre. Depuis le début du xx<sup>e</sup> siècle, la génétique des populations s'est constituée en discipline attachée à décrire cette variabilité génétique entre groupes humains (Bourgain et Darlu, 2013 ; Sommer, 2016) pour étudier l'histoire humaine (Cavalli-Sforza *et al.*, 1994). En partant de groupes socialement construits (ethnies, nations, langues, religions...), auxquels ces généticiens ont ajouté leurs propres critères – reflet de leurs représentations et des contraintes pratiques auxquelles ils étaient confrontés (M'Charek, 2005) –, des populations ont été définies et les individus qui les composent ont été caractérisés génétiquement.

Comme nous l'avons mentionné, les évolutions technologiques et réglementaires et les investissements massifs dans le secteur des biotechnologies ont contribué ces dernières années à la mise en place de nombreuses banques de données, publiques ou privées, dans lesquelles les caractéristiques génétiques d'individus rattachés à différents groupes humains sont conservées. Proposer de révéler les origines d'un ADN particulier consiste ainsi à caractériser certains segments bien choisis sur cet ADN et à comparer ensuite cette information avec celle des bases de données. Les origines correspondront au(x) groupe(s) présents dans les bases, dont l'ADN considéré est le plus proche.

88 Plusieurs éléments viennent compliquer la mise en pratique de ces reconstitutions. D'une part, les informations génétiques disponibles dans ces bases sont extrêmement diverses. Si les plus anciennes comportent une petite dizaine de segments par individu, les plus récentes peuvent en contenir plusieurs millions. Selon les segments caractérisés sur un ADN donné, les bases de données utilisables ne sont donc pas les mêmes. D'autre part, les critères utilisés pour définir les mêmes groupes peuvent être variables d'une base de données à l'autre, et l'information disponible sur les différents groupes est souvent hétérogène. Quand certains sont représentés par plusieurs milliers d'individus, d'autres sont en effectif très réduit. Les résultats des analyses sont ainsi susceptibles de varier selon les bases de données utilisées et leur géographie interne. Enfin, ces comparaisons n'ont rien d'automatique. Des hypothèses de simplification, des modèles statistiques et des algorithmes pour les implémenter sont requis pour les mener à bien. Selon le logiciel utilisé, les résultats de comparaison entre un profil génétique et une base donnée sont susceptibles de varier de façon substantielle (Bourgain *et al.*, 2019). Plus les origines des individus sont multiples et peu congruentes avec les catégories retenues dans les bases de données, et moins les résultats produits seront robustes et convaincants. Au total, donc, plutôt que de « révélations », c'est bien de procédures d'inférence complexe dont il est question. Ces procédures dépendent des choix qui ont présidé à la constitution des grandes infrastructures de données, des choix implémentés par les développeurs de logiciels et des choix des utilisateurs finaux.

Avec la prédiction de caractéristiques morphologiques apparentes, l'analyse entre dans le champ de la prédiction génétique des caractères, amplement labouré par les généticiens humains depuis les débuts de leur discipline dans le dernier tiers du XIX<sup>e</sup> siècle (Bourgain, 2017). Deux dimensions sont importantes pour les applications discutées ici. La première a trait à la modalité de mesure retenue pour le caractère d'intérêt. Si certains traits se laissent aisément décrire sur une échelle unidimensionnelle, comme c'est le cas de la taille ou du poids, il en va différemment de données plus complexes comme la forme d'un visage, la couleur de certains yeux, cheveux ou peaux. Les outils de mesure objective qui sont utilisés pour définir ces caractéristiques peuvent avoir une certaine efficacité – les catégories fondées sur ces mesures correspondent plutôt bien aux catégories généralement perçues par l'œil observateur –, ils n'en constituent pas moins des simplifications. Il n'existe pas une façon unique d'avoir les yeux verts, et des observateurs différents peuvent associer une même paire d'yeux à des catégories de couleurs différentes. La perception des caractéristiques morphologiques apparentes est

en outre largement socialement construite, et nombre d'entre elles (couleur des cheveux, de la peau, poids, taille...) évoluent au cours de la vie. La seconde dimension à souligner porte sur les liens entre ADN et caractéristiques morphologiques. Prédire à partir de segments génétiques implique que le caractère considéré est largement déterminé par la génétique et que les facteurs d'environnement n'ont qu'une influence marginale et négligeable. Or les influences respectives des facteurs génétiques et environnementaux sont souvent très interdépendantes, comme l'illustre bien l'exemple de la taille. Les dernières études cherchant à préciser les segments génétiques permettant de prédire la taille ont identifié près de 800 segments différents. Mais l'information cumulée de tous ces segments ne permet d'expliquer que 27 % des variations interindividuelles de taille (Marouli *et al.*, 2017). *A contrario*, des évolutions importantes de taille sur des temps courts ont été décrites dans certaines populations (Ezzati *et al.*, 2015) – entre 1896 et 1996, les femmes de Corée du Sud ont ainsi grandi en moyenne de 20,5 cm et les hommes d'Iran de 16,5 cm –, qui ne sauraient être expliquées par des changements génétiques. Elles témoignent au contraire de l'importance du contexte environnemental et culturel – principalement en matière de nutrition et d'hygiène – dans lequel se trouvent les mères pendant leur grossesse et les individus jusqu'à la fin de leur croissance. Loin d'être une exception, cet exemple reflète bien la situation de nombreuses caractéristiques morphologiques apparentes, dont le déterminisme très largement multifactoriel est aujourd'hui encore très mal décrit.

89

Certains traits morphologiques constituent toutefois des exceptions partielles. C'est notamment le cas de la couleur des yeux. Une quinzaine de segments sur l'ADN ont ainsi été associés aux variations de couleur entre le bleu et le marron (Kayser et de Knijff, 2011). Dans un grand échantillon d'individus néerlandais d'ascendance européenne, une équipe de recherche a montré que ces segments permettaient de prédire correctement la couleur pour 93 % des individus dont les yeux sont effectivement bleus et pour 91 % des individus dont les yeux sont effectivement marrons (Liu *et al.*, 2009). Pour les individus dont la couleur est jugée « intermédiaire », les résultats sont moins convaincants. Les prédictions ne sont correctes que dans 73 % des cas. Les individus ne sont pas « égaux » devant la prédiction génétique de la couleur de leurs yeux. Ainsi, non seulement les prédictions génétiques sont plus ou moins précises selon les caractéristiques, mais, pour une caractéristique donnée, la qualité de prédiction est variable d'un individu à l'autre.

Proposer de révéler les caractéristiques morphologiques associées à un ADN particulier requiert en conséquence de choisir les segments informatifs pour chaque caractéristique considérée, de choisir les données de référence et les méthodes statistiques qui permettront de proposer une prédiction sur la base du résultat de l'analyse des segments. Toutes ces étapes sont autant de moments qui peuvent présenter des incertitudes quant à la pertinence de la prédiction strictement génétique, quant à la qualité et la réplicabilité des résultats de recherche mobilisés pour fonder les prédictions, quant à la robustesse des méthodes utilisées et l'interprétation des résultats obtenus. Comme pour les origines géographiques, il s'agit de procédures d'inférence complexe, largement tributaires de la qualité des études



scientifiques publiées et des choix implémentés par les développeurs des logiciels de prédiction, dont beaucoup sont vendus par des firmes privées.

## II. DÉPLACEMENTS D'USAGES ET DE SAVOIRS : VERS UNE MONTÉE DE L'INCERTITUDE ET DE SA GESTION

90

Dans sa forme historique, le recours à l'ADN traceur repose sur un postulat fort. Chaque individu possède un ADN qui lui est propre et dont la spécificité peut être révélée par la caractérisation d'un profil génétique. Dès lors, la fonction du laboratoire de biologie est claire. Il doit prendre toutes les précautions pour que cette révélation des profils génétiques soit exacte : s'assurer que la collecte des échantillons biologiques est réalisée sans contamination pour que le profil produit corresponde à la trace biologique relevée sur la scène de crime ; mettre en place des procédures de contrôle qualité en interne pour éviter toutes les erreurs de procédure (échanges de tubes, erreurs de manipulation, mauvais usage de la procédure d'analyse, erreurs dans le rendu des résultats...). La réalité est souvent plus complexe. Lorsque les ADN sont dégradés, il peut être compliqué d'obtenir un profil complet et d'assurer en conséquence l'identification de l'individu. Des échantillons biologiques peuvent combiner des ADN issus de plusieurs personnes. Le mélange qui en résulte peut rendre l'analyse des profils complexe et incertaine. Mais, pour autant, la fiction selon laquelle la mission du laboratoire est de révéler des données tangibles, indiscutables, et de produire ainsi une preuve scientifique objective, reste largement partagée entre acteurs des analyses et avec les cercles extérieurs – public, monde judiciaire, politiques... Le terme de fiction (Beaudevin *et al.*, 2019) permet d'insister sur le fait que beaucoup d'acteurs, et notamment ceux qui sont en charge de ces analyses, ne sont pas dupes du caractère partiellement construit de ce récit. Mais il permet également de souligner son importance dans la façon dont ils font collectivement sens de leurs pratiques de travail et de leur légitimité.

Avec la montée en puissance des nouveaux usages de l'ADN, les réalités quotidiennes de travail sont amenées à évoluer. Les quatre exemples présentés – recherches familiales, ADN de contact, prédiction des origines et des caractéristiques morphologiques apparentes – mettent tous à mal cette fiction. À l'instar d'autres contextes, notamment médicaux, ces innovations n'évacuent pas l'incertitude, mais au contraire en engendrent de nouvelles (Fox, 1999). Ces dernières peuvent se loger au cœur même du travail de production des données dites « brutes », comme l'illustrent les problèmes d'erreurs aléatoires associées aux quantités infimes des ADN de contact. Elles se logent également aux différents étages de nouvelles étapes d'interprétation devenues incontournables et qui nécessitent la mobilisation de nombreux éléments extérieurs au laboratoire. Une caractéristique commune aux quatre nouveaux usages présentés réside en effet dans leur dépendance accrue à des données et procédures d'analyse produites et évaluées en dehors du laboratoire : banque de données de référence, publications scientifiques pour choisir les segments à caractériser, logiciels pour conduire les analyses...

Ces formes d'incertitudes systémiques requièrent un déplacement de la figure et des compétences de l'expert en analyses génétiques. Les approches probabilistes deviennent ainsi incontournables. Comme le décrit l'historien Harry Marks dans son histoire de l'introduction des statistiques et des essais randomisés en santé, lorsque le hasard est considéré comme inévitable, la posture du statisticien est de le mesurer plutôt que de chercher à le réduire (Marks, 1997). Cette place nouvelle prise par les analyses et les logiciels statistiques dans tous les nouveaux usages présentés peut ainsi être interprétée comme autant de marqueurs de ces incertitudes. L'expertise génétique n'a plus comme objectif de révéler une information largement considérée comme indiscutable. Elle doit désormais produire un résultat présentant une incertitude irréductible et proposer une mesure considérée comme précise et robuste de cette incertitude.

Les sociologues qui se sont intéressés à des pratiques génétiques présentant un lien similaire à l'incertitude en contexte médical ont montré combien ce travail reposait à la fois sur de nombreux standards (recommandations de bonnes pratiques, bases de données de référence...) et sur la capacité des experts à prendre un recul critique sur ces standards, pour les mobiliser à bon escient, de façon adaptée à chaque situation, toujours particulière (Timmermans, 2015).

91

Les propos de John Butler (2015a), extraits d'une réflexion sur le futur des analyses génétiques en contextes policier et judiciaire, résonnent fortement avec le cas biomédical. Il souligne ainsi que « produire des conclusions plus solides en utilisant des approches probabilistes sur les données va nécessiter de pouvoir valider les logiciels utilisés et de bien comprendre leurs limites. Les scientifiques qui utilisent ces logiciels ne doivent pas les traiter comme des boîtes noires produisant des résultats qu'ils sont incapables de comprendre. » Mais, au-delà de cet appel à de nouvelles pratiques de réflexivité sur les standards mobilisés par les experts, Butler insiste sur une dimension complémentaire : « La nécessité de bien communiquer sur le sens des résultats produits à ceux qui les reçoivent et les utilisent pour prendre des décisions. Ces résultats doivent faire apparaître les limites des méthodes, les modèles, les hypothèses et les étapes d'interprétation [...] La communauté doit être attentive et fournir les efforts requis pour proposer des interprétations sans exagération ». L'idée qu'il existe des incertitudes irréductibles doit sortir des cercles experts du laboratoire pour être communiquée et prise en compte par tous les acteurs des différentes étapes d'utilisation des résultats.

Ce besoin nouveau de mettre en place des formes de travail et de communication pour intégrer et gérer ces incertitudes est peu compatible avec la fiction du laboratoire de génétique révélant la preuve scientifique indiscutable. Cette dernière, trop éloignée des enjeux quotidiens, permet mal de faire sens des pratiques. La question se pose de savoir si des fictions alternatives sont en développement ou pourraient l'être. Autrement dit, cette nouvelle place de l'incertitude est-elle susceptible de remettre en question l'importance que les analyses génétiques ont prise ces dernières années dans les pratiques policières et judiciaires ? Dans le monde biomédical, cette montée de l'incertitude au cœur des nouvelles technologies de génomique n'a pas empêché leur diffusion. Des formes de

régulation professionnelle ont été mises en place, avec notamment des pratiques visant à renforcer la confiance dans les analyses par l'articulation des données génétiques avec des données cliniques, biologiques ou d'imagerie. Mais la situation est différente et plus complexe en contexte pénal. D'une part, l'expert généticien revendique de connaître le moins d'éléments possibles sur l'affaire pour conduire son travail de révélation en toute objectivité. L'articulation de la génétique avec d'autres formes de données imposerait en conséquence une évolution notable de cette posture de l'expert qui pourrait le fragiliser. D'autre part, les épreuves de vérité auxquelles les preuves sont soumises en contexte pénal sont beaucoup moins concentrées dans les mains des seuls experts qu'elles ne le sont en contexte médical. Elles font au contraire une place beaucoup plus grande à la contradiction avec d'autres acteurs, notamment lors du procès. Les régulations professionnelles ne peuvent donc pas être envisagées de façon trop parallèle dans ces deux champs, et les comparaisons sont délicates.

Ces éléments ne manquent toutefois pas d'interroger sur les promesses de la génétique en contextes policier et judiciaire, les simplifications qui les sous-tendent et la pertinence des investissements et des évolutions réglementaires qui sont décidées en leur nom.

92

## BIBLIOGRAPHIE

- AULCHENKO Y. S. *et al* (2009), "Predicting human height by Victorian and genomic methods", *EJHG*, 17: 1070-1075.
- BALDING D.J. & BUCKLETON J. (2009), "Interpreting low template DNA profiles", *Forensic Sci. Int. Genet.* 4(1):1-10.
- BALDING D. J., KRAWCZAK M., BUCKLETON J. S., CURRAN J. M. (2013), "Decision-making in familial database searching : KI alone or not alone ?", *Forens. Sci. Int. Genet.* 7, 52–54.
- BEAUDEVIN C., PEERBAYE A., BOURGAIN C. (2019), "It has to become true genetics' Tumour genetics and the division of diagnostic labour in the clinic", *Sociology of Health and Illness*, 41: 643-657.
- BOURGAIN, C., DARLU, P. (2013), *ADN superstar ou superflic ? Les citoyens face à une molécule envahissante*, Paris, Le Seuil.
- BOURGAIN C. (2017), « Agir sur les gènes, est-ce suffisant ? », *Archives de philosophie du droit*, 59 :39-52.
- BOURGAIN C., SABBAGH A., TURRINI M. (2019), « Maladies, paternité, origine ethnique. Faut-il se fier aux tests génétiques ? », *Pour la Science*, 103 :26 :39.
- BROWN, N. (2003), "Hope against Hype. Accountability in Biopasts, Presents and Futures", *Science Studies*, 16:3–21.
- BUTLER J.M. (2015a), "The future of forensic DNA analysis", *Phil. Trans. R. Soc. B* 370 : 20140252.
- BUTLER J.M. (2015), *Advanced Topics in Forensic DNA Typing : Interpretation*, San Diego, Ed. Elsevier Academic Press.

- BUTLER J.M. (2016), Challenges in Forensic Genetics. EuroForGen. International Dissemination Conference. [https://strbase.nist.gov/pub\\_pres/Butler-EuroForGen-June2016.pdf](https://strbase.nist.gov/pub_pres/Butler-EuroForGen-June2016.pdf)
- CAVALLI-SFORZA L., MENOZZI P., PIAZZA A. (1994), *The History and Geography of Human Genes*, Princeton University Press.
- ERLICH Y., SHOR T., PE'ER I., CARMİ S. (2018), "Identity inference of genomic data using long-range familial searches", *Science*, 363:690-694.
- FOX R. C. (1999), "Medical Uncertainty Revisited", in G. ALBRECHT, S. C. SCRIMSHAW & R. FRITZPATRIK (Eds.), *Handbook of Social Studies of Health and Medicine*, Londres, Sage, p. 409-425.
- GARRISON N. A., ROHLFS R. V., FULLERTON S. M. (2013), "Forensic familial searching : scientific and social implications", *Nature Review Genetics*, 14:445.
- HILGARTNER S. (2017), *Reordering Life: Knowledge and Control in the Genomics Revolution*, MIT Press.
- HJORLEIFSSON S., ARNASON V., SCHEIA E. (2008.), "Decoding the Genetics Debate : Hype and Hope in Icelandic News Media in 2000 and 2004", *New Genetics and Society* 27(4):377-94.
- JEFFREYS A., WILSON V., THEIN S. (1985), Hypervariable "minisatellite" region in human DNA, *Nature*, 314: 67.
- JOLY P.B. (2015), « Le régime des promesses technoscientifiques », in M. Audétat (dir.), *Sciences et technologies émergentes : pourquoi tant de promesses ?* Ed. Hermann, Editeurs des Sciences et des Arts.
- KAYSER M., DE KNIJFF P. (2011), "Improving human forensics through advances in genetics, genomics and molecular biology", *Nat Gen Rev.* 12:179 -192.
- LIU F. *et al.* (2009), "Eye color and the prediction of complex phenotypes from genotypes", *Current Biology*, 19: R192-R193.
- M'CHAREK A. (2005), *The Human Genome Diversity Project : an Ethnography of Scientific Practice*, Cambridge University Press.
- MARKS H. (1997), *La médecine des preuves. Histoire et anthropologie des essais cliniques (1900-1990)*, Éditions Les empêcheurs de penser en rond.
- MAROULI E. *et al.* (2017), "Rare and low-frequency coding variants alter human adult height", *Nature*, 542: 186.
- ROHLFS R.V., FULLERTON S.M., WEIR B.S. (2012), "Familial Identification : Population Structure and Relationship Distinguishability", *PLoS Genet* 8(2): e1002469.
- SOMMER M. (2016), *History within: the science, culture, and politics of bones, organisms, and molecules*, The University of Chicago Press.
- TIMMERMANS S. (2015), "Trust in standards: transitioning clinical exome sequencing from bench to bedside", *Social Studies of Science*, 45, 1, 77-99.

