

## «L'empreinte génétique renseigne sur l'appartenance ethnique»

Par **Michel de Pracontal** et **Louise Fessard**

Article publié le mardi 13 septembre 2011

Statisticienne en génétique à l'Inserm, Catherine Bourgain a témoigné lors du procès de plusieurs militants ayant refusé un prélèvement ADN. Au fil de ses interventions, cette chercheuse s'est rendue compte que les 15 à 18 marqueurs génétiques utilisés par le Fnaeg pouvaient livrer des informations sur les maladies, les prédispositions pathologiques ou l'origine géographique des individus fichés.

En contradiction avec l'idée reçue à la création du Fnaeg en 1998, selon laquelle les marqueurs utilisés étaient «neutres» et ne pouvaient fournir aucune autre information que celle permettant l'identification.

Ces avancées pourraient remettre en cause les fondements du fichier français, le code de procédure pénal **disposant** que les empreintes «*ne peuvent être réalisées qu'à partir de segments d'acide désoxyribonucléique non codants*». Entretien.

### Qu'est-ce qu'une empreinte génétique ?

**Catherine Bourgain** : C'est une combinaison de séquences d'ADN qui permet d'identifier une personne. Tous les êtres humains possèdent plus de 99% d'ADN strictement identique. Ce qui distingue deux individus est concentré sur le moins de 1% restant. Plus précisément, ces différences sont localisées dans des régions particulières, qui varient d'un sujet à l'autre, et que l'on appelle «marqueurs génétiques».

Au début des années 1980, des techniques ont été mises au point pour sonder spécifiquement ces régions variables de l'ADN. A cette époque, la communauté scientifique s'est rendu compte que certains marqueurs génétiques pouvaient exister sous plus de 30 formes différentes selon les personnes.

De ces observations est née l'idée qu'on pourrait identifier une personne à partir de son ADN, en examinant un petit nombre (moins d'une dizaine à

l'origine) de marqueurs génétiques choisis parmi les plus polymorphes (c'est-à-dire ceux qui varient le plus d'un individu à l'autre).

En 1984, le Britannique Alec Jeffreys a mis au point une technique pour réaliser une telle identification, en se basant sur certains marqueurs très variables appelés minisatellites. Le procédé fait appel à des enzymes qui découpent l'ADN en certains sites, en fonction des variations individuelles.

On obtient ainsi un profil spécifique de l'individu analysé, qu'on a appelé empreinte génétique.

### Par analogie avec les empreintes digitales ?

En effet. Le profil obtenu par la technique de Jeffreys se présente comme une série de bandes noires qui apparaissent sur un film transparent, et qui font penser à un code barre. Ce motif est particulier à un individu donné, comme l'est une empreinte digitale, d'où l'analogie.

Mais en réalité, ce n'est pas vraiment l'équivalent d'une empreinte digitale. D'abord, elle n'est pas directement accessible.

Pour l'obtenir, il faut recueillir un échantillon de salive, de muqueuse ou encore un cheveu, le traiter pour en extraire l'ADN, puis analyser cet ADN avec des méthodes biologiques, statistiques et informatiques.

Contrairement à l'empreinte digitale, il n'existe pas une empreinte ADN unique pour chaque personne. Le profil obtenu dépend du nombre de marqueurs choisis. Plus on prend de marqueurs, et plus l'identification est fiable.

Ainsi, en 1994, lorsqu'on a commencé à constituer le fichier anglais, on se servait de 6 marqueurs. On en utilise aujourd'hui 10, ce qui donne une grande fiabilité : la probabilité que deux personnes différentes présentent le même profil est seulement de 1 sur 3.000 milliards !

En France, le nombre de marqueurs est passé de 7, lors de la mise en place du fichier Fnaeg en 1998, à 18 (liste fixée par un **arrêté du 23 octobre 2006**)...

### Qu'est-ce qui a permis de recourir à plus de marqueurs ?



Cette augmentation résulte d'abord des progrès techniques qui ont permis d'améliorer les procédures d'analyse, aujourd'hui automatisées, beaucoup plus rapides et moins chères. En recherche, nos études incluent, aujourd'hui, de façon courante, 1 million de marqueurs. Pour les usages de médecine légale et de police, il n'est pas nécessaire d'en avoir autant.

Mais de fait, les améliorations techniques rendent les analyses de plus en plus accessibles. Ce qui n'est pas sans poser problème, car à la différence des empreintes digitales qui sont un motif sans signification biologique, les empreintes génétiques livrent de nombreuses informations sur l'individu analysé.

Le profil génétique en dit beaucoup plus qu'une empreinte digitale. Il permet de savoir si le sujet peut être touché par une maladie génétique, s'il a telle ou telle prédisposition, et même quelle est son origine ethnique...

### **Cela était-il prévu lorsqu'on a créé les fichiers d'empreintes génétiques ?**

Non, justement. Les marqueurs choisis au niveau international étaient censés ne pas donner d'information biologique sur l'individu. Au départ, les scientifiques britanniques ont opté pour des régions de l'ADN qui offraient une grande variabilité d'un individu à l'autre et se situaient sur différents chromosomes.

Comme les polices des différents pays coopèrent et vont voir ce qui se faisait chez le voisin, ces premières options ont été reprises dans des normes internationales. Selon la conception initiale, ces choix avaient été faits en utilisant des régions de l'ADN dit «non-codant», sans signification biologique. C'est ce que l'on pensait en 1998, lorsqu'on a créé le Fnaeg. Les parlementaires pensaient que les marqueurs choisis ne donnaient guère de renseignement sur la personne analysée.

### **Et la vision a changé ?**

Oui, parce qu'on a réalisé que les marqueurs en disent beaucoup plus que ce que l'on croyait. Pendant longtemps, on a eu une vision du génome très

compartimentée : telle région de l'ADN sert à quelque chose, c'est un gène qui code pour une protéine, tel segment est de l'ADN poubelle, dont on ignore à quoi il sert et dont la variation n'a pas, a priori, de conséquences sur l'individu.

Comme les marqueurs choisis pour l'expertise judiciaire étaient pris dans l'ADN non codant, on pensait qu'ils ne pouvaient servir qu'à l'identification de personnes ou de traces, sans fournir d'information autre sur les traits génétiques de la personne elle-même.

Or, la communauté scientifique s'est rendu compte que la dichotomie codant/non codant n'était pas aussi nette qu'on le croyait. Il y a de fortes interférences entre ADN non-codant et ADN codant. Parfois, l'action d'un gène est modulée par une séquence «non codante» : autrement dit, un même gène peut avoir des effets différents en fonction d'une séquence variable située assez loin de ce gène et qui est censée ne pas avoir de rôle biologique.

Bref, les résultats qui s'accumulent depuis une bonne dizaine d'années tendent à remettre en question l'idée qu'il existerait un ADN «totalement neutre».

### **Donc, il n'y a pas non plus de marqueurs neutres ?**

Prenons un exemple : l'un des marqueurs utilisés par le Fnaeg, appelé D2S1338, a été étudié par l'équipe du Professeur Gasparini de l'institut TIGEM de Naples, l'équivalent italien du Téléthon français. Ces chercheurs ont travaillé sur une famille de Lille dans laquelle se transmet un dysfonctionnement très spécifique des globules rouges qu'on appelle la pseudokaliémie.

Ils ont montré que le marqueur D2S1338 était le point de l'ADN qui permettait le mieux de déterminer qui dans cette famille était atteint de pseudokaliémie, et qui ne l'était pas. Donc, cette séquence d'ADN, réputée neutre et non informative, donne en fait une indication précise.

Dans ce cas, il s'agit d'une maladie rare et il est donc peu vraisemblable que l'information soit largement utilisable. Il n'en reste pas moins qu'en 2004, un marqueur «non codant» est devenu «codant». En



conséquence, il n'est pas exclu que les empreintes génétiques, aujourd'hui muettes, deviennent plus bavardes demain.

### Quel type d'informations pourraient-elles livrer ?

En 2010, une équipe dirigée par la chercheuse portugaise Luisa Pereira a **montré** qu'en se servant uniquement des 17 marqueurs couramment utilisés à travers le monde dans les enquêtes criminelles, on pouvait obtenir une indication assez précise sur l'origine géographique d'un individu.

Le **principe** de base utilisé par ces chercheurs est le suivant : dans chaque population, certains allèles sont plus fréquents que d'autres (on appelle allèles les différentes formes que peut prendre un segment d'ADN variable). Autrement dit, chaque groupe humain est caractérisée par une répartition statistique des allèles particulière.

Ainsi, les groupes sanguins ne sont pas répartis de la même manière selon les populations. Si je connais uniquement votre groupe sanguin et que vous êtes B, je peux en conclure que vous avez plus de chance d'être Russe que Français, parce que l'allèle B est plus fréquent chez les Russes.

Bien sûr, il ne suffit pas de considérer un seul marqueur ou un seul gène variable pour déterminer l'origine d'un individu. Mais si on en croise plusieurs, on arrive à augmenter la probabilité d'appartenance à un groupe donné.

Pour en revenir au travail de Luisa Pereira et ses collègues, ils ont collecté les empreintes génétiques déjà existantes de plus de 50.000 individus, issus de 40 populations différentes dans le monde.

À partir de cette base de données d'empreintes génétiques dont ils connaissaient l'origine, ils ont mis au point un algorithme, baptisé Population Affiliator. Ce logiciel permet de calculer la probabilité qu'une empreinte corresponde à une personne asiatique, eurasiennne ou sub-saharienne.

Or, il apparaît que dans 86 % des cas, le logiciel donne une réponse exacte. Cette recherche très intéressante a été publiée en 2010 dans la revue *«International journal of legal medicine»*. Elle démontre que l'empreinte génétique peut donner une information sur l'appartenance ethnique d'une personne.

Même si, pour l'instant, c'est une indication assez grossière, il n'est pas exclu qu'en affinant la méthode on puisse, par exemple, distinguer un Basque d'un Nord-Africain. Et on pourrait aller encore plus loin, en se servant de l'information sur l'origine pour accéder à des caractéristiques physiques : les Asiatiques ou les Sub-Sahariens ont rarement les yeux bleus, par exemple.

Bref, l'empreinte génétique est beaucoup plus révélatrice que ce que l'on pensait lorsqu'on a défini le cadre légal du Fnaeg et des fichiers analogues dans les autres pays.

**Lire la réaction de la Cnil sous l'onglet Prolonger.**

**Directeur de la publication** : Edwy Plenel

**Directeur éditorial** : François Bonnet

**Directrice général** : Marie-Hélène Smiéjan

**Le journal MEDIAPART est édité par la Société Editrice de Mediapart (SAS).**

Durée de la société : quatre-vingt-dix-neuf ans à compter du 24 octobre 2007.

Capital social : 1 538 587,60€

Immatriculée sous le numéro 500 631 932 RCS PARIS. Numéro de Commission paritaire des publications et agences de presse : 1214Y90071.

Conseil de direction : François Bonnet, Jean-Louis Bouchard, Michel Broué, Laurent Mauduit, Edwy Plenel (Président), Marie-Hélène Smiéjan, Thierry Wilhelm. Actionnaires directs et indirects : Godefroy Beauvallet, François Bonnet, Gérard Desportes, Laurent Mauduit, Edwy Plenel, Marie-Hélène Smiéjan ; Laurent Chemla, F. Vitrani ; Société Ecofinance, Société Doxa ; Société des Amis de Mediapart.

Rédaction et administration : 8 passage Brulon 75012 Paris

**Courriel** : contact@mediapart.fr

**Téléphone** : + 33 (0) 1 44 68 99 08

**Télécopie** : + 33 (0) 1 44 68 01 80 ou 01 90

**Propriétaire, éditeur et prestataire des services proposés sur ce site web** : la Société Editrice de Mediapart, Société par actions simplifiée au capital de 1 538 587,60 euros, immatriculée sous le numéro 500 631 932 RCS PARIS, dont le siège social est situé au 8 passage Brulon, 75012 Paris.

Abonnement : pour toute information, question ou conseil, le service abonné de Mediapart peut être contacté par courriel à l'adresse : serviceabonnement@mediapart.fr. Vous pouvez également adresser vos courriers à Société Editrice de Mediapart, 8 passage Brulon, 75012 Paris.