

## Génétique et recherche des origines

Pierre Darlu, UMR7206, CNRS, MNHN, Université Paris VII

La question de la recherche des origines prend une nouvelle dimension depuis que la génétique s'en mêle. Le temps où la recherche généalogique consistait à fouiller les archives de mairies ou de départements est bien révolu. La génétique semble fournir un outil bien plus performant pour remonter dans le temps, plus précis (car plus biologique ?) et plus loin (jusqu'à Adam et Eve ?). Diverses entreprises commerciales n'hésitent pas à l'utiliser pour leur plus grand bénéfice. Les médias en exploitent les résultats les plus aguichants et les scientifiques tirent profit à la fois des produits commerciaux et de l'effet médiatique, évidemment motivés par leur soif légitime de connaissances. Mais qu'en est-il exactement de l'apport scientifique de la génétique à la recherche des origines et quelles implications peut-il avoir au niveau sociétal.

### I- De la génétique des populations à la génétique personnalisée

Un premier point est le changement de paradigme qui s'opère actuellement dans le champ de la génétique des populations.

I-1) Dans les années 1970, la génétique des populations montrait que la diversité génétique entre populations est extrêmement faible par rapport à la diversité observée à l'intérieur d'une population. Cette constatation, soutenue par les travaux de Lewontin, a largement contribué à discréditer l'idée que la race avait un fondement génétique (voir par exemple une réflexion de Lewontin sur la « race » <http://raceandgenomics.ssrc.org/Lewontin/>) ( voir aussi Edwards, pour une critique de ce calcul de la diversité, *BioEssays*, 2003, 25 :798-801 : « Human genetic diversity : Lewontin's fallacy ») Ces travaux reposent sur deux concepts, celui de population et celui de diversité génétique.

I-1-a) Le concept de population reste difficile à définir ; C'est, au mieux, un ensemble de personnes, non apparentés (mais jusqu'à quel degré ?), ayant les 4 grands-parents appartenant à la même zone géographique (mais quelle étendue ?), et partageant la même langue et/ou la même culture (mais quel niveau d'identité ?). C'est aussi, bien que contradictoire avec l'absence de parenté, un ensemble de personnes qui se marient davantage entre eux qu'avec l'extérieur (mais quel est la taille de ce cercle de mariage définissant « l'extérieur » ?). Donc définir une population soulève beaucoup de questions qui relèvent de la géographie, de la généalogie, de la culture (au sens large). C'est ce concept flou dont les généticiens s'accommodent plutôt facilement (à la différence des ethnologues ou anthropologues, plutôt rétifs) et qui sert de base à toutes les recherches en génétique des populations où l'individu n'est pas l'objet sur lequel se focalisent les recherches, mais l'entité « population » comme ensemble des personnes qui la composent. La plupart des sites commerciaux qui proposent des tests d'origine reste confus sur ces nuances pour ne se placer qu'à un niveau intuitif : « population », « ethnie », « ancestral origin », « ethnic background », « ethnic origins », « ethnic populations » « population groups », « racine », « peuple d'origine », « race »...

I-1-b) Le concept de diversité génétique est habituellement fondé sur des observations multivariées qui consistent à amalgamer les informations données par le plus de marqueurs indépendants possibles (groupes sanguins classiques et groupes immunitaires, marqueurs sur l'ADN, comme les CNV, microsat, SNP...) et mesurer, par divers indices synthétiques, le degré de différence entre les individus au sein d'une population ou entre populations. Concrètement, les résultats sont présentés sous différentes formes : i) des cartes décrivant la diversité génétique comme un continuum (voir les représentations de Boyd , Mourant, Cavalli-Sforza et al.), au sein duquel la localisation de frontières ou de barrières reste difficile ; ii) des d'arbres phylogénétiques ou des plans CPA construisant des clusters de populations.

I-2) Les données, de la population à l'individu. Récemment, cette thématique a pris un évident tournant. Les données génétiques deviennent de plus en plus abondantes, à la fois par le nombre de personnes rentrant dans les bases de données (Family Tree DNA aurait 177 075 Y-DNA and 107 302 mtDNA records dans leur base de données), par le nombre de populations investiguées (DNA Tribes annonce « 1000 « ethnic populations and 39 world regions », et par la multiplicité des marqueurs génétiques dont le polymorphisme est étudié (500000 SNP par personne est maintenant une quantité classique). Bien que les problématiques de la génétique des populations perdurent, avec ses thématiques habituelles, une place particulière est maintenant donnée à l'individu. Plusieurs tendances semblent se dessiner.

I-2-a) La première est la recherche des gènes « typiques » ou marqueur spécifique d'une population (au sens donné plus haut). Ce sont les AIM « Ancestry Informative Markers » ceux qui sont les plus à même de définir tel ou tel population. Certains considèrent qu'un panel de 93 SNP serait suffisant pour identifier les groupes de populations, pointer du doigt les « outliers » et établir des degrés de mélanges (Nassir et al., 2009) [<http://www.biomedcentral.com/content/pdf/1471-2156-10-39.pdf>]. D'autres marqueurs sont considérés comme très populations-spécifiques. A titre d'exemple, citons l'haplogroupe J1 (sur le chromosome Y, Ychr) présent avec une forte fréquence dans l'est anatolien et au Caucase, les haplogroupes R1b1c4 and R1b1c6 chez les Basque; l'haplogroupe K1a1b1a (mtDNA) chez les Ashkenazes (« Study identify genetic signatures for 3.5 million of Ashkenazi jews »). Certaines entreprises commerciales, dans leur démarche DTC (« direct-to-consumer »), font une exploitation systématique de ces AIM ou marqueurs « typiques », sans toujours préciser leur méthodologie, et proposent, par exemple d'établir si vous êtes de la lignée des Levi, des Cohen ou des Pigeons (cf. Charb) (voir plus loin). N'assiste-on pas là à une actualisation savante du marquage ethnique ou racial, à partir d'arguments génétiques ?

I-2-b) Une deuxième tendance consiste à rechercher la façon optimale de regrouper les individus en populations « statistiques », et non plus seulement en « vrai » population (au sens flou donné plus haut). C'est le triomphe du logiciel « Structure » et de ses dérivés dont les diagrammes fourmillent dans les journaux spécialisés. Des hypothèses successives sont formulées sur le nombre possible de groupes, ou clusters, d'individus : 2, 3, 4 ...n. Des critères statistiques permettent de retenir le nombre optimal de clusters. Une statistique est alors rendue par personne : elle estime la part de son génome qui proviendrait de chacun de ces clusters : telle proportion venant du cluster 1, une autre venant du cluster 2,...du nième etc. Une telle approche revient à considérer que l'humanité est constituée de groupes distincts, même si il est admis que les gènes d'une personne puissent être attribués à plusieurs de ces groupes. Le concept de continuum est ainsi revisité et laisse la place à une certaine vision classificatoire de l'espèce humaine (finally, the question of the natural divisions of the human species has been resurrected [...] Genetically reified races now present the possibility of « niche marketing » , J. Marks, 2008)

I-2-c) Une autre technique, voisine de la précédente, revient à se fonder non pas sur des populations « statistiques », nécessairement virtuelles même si elles finissent par représenter des clusters dont le sens géographique ou ethnique peut parfois s'interpréter, mais à se fonder sur les populations elles-mêmes, ou plutôt sur les échantillons qui en ont été sélectionné par les généticiens. Des procédures statistiques simples permettent d'attribuer, en probabilité, les gènes de chaque personne à tel ou tel échantillon de « vrai » population (Cf par ex. l'article de Bryc (PNAS 2010 sur l'Ouest africain). La distinction avec la procédure précédente est nette, quant au fond (la population n'est pas « statistique », mais bien réelle) même si cette différence échappe souvent.

I-2-d) Une façon d'évacuer la question des origines, pourtant indispensable pour élaborer une stratégie raisonnée d'échantillonnage, est de se reporter à la déclaration personnelle d'appartenance que peut faire chaque personne de sa « *self identified ethnic or racial group* ». Cette procédure est bien connue pour être entachée de biais phénoménaux que certains généticiens pensent minimiser grâce aux vertus de la théorie des grands nombres qui permet de gommer les erreurs d'attribution, supposées rares, et les effets perturbateurs dus à la présence de quelques « ouliers »... Si cette approche est confortable pour le généticien, elle n'est pas sans inconvénients. Car la race ou l'ethnie ne se définit plus que par ce que les gens veulent bien en penser, si bien que le scientifique ne s'appuyant plus que sur une définition « vulgaire » de la race finit, dans une certaine mesure, par la conforter (cf l'analyse du numéro spécial de Nature Genetics (Octobre 2004) sur le thème « Genetics for the Human Race »)

En conclusion, on peut souligner un évident changement de paradigme : si les approches définies en I-1 tendaient à nier la pertinence de la génétique pour différencier les populations ou les groupes, remettant en cause la possibilité d'attribuer un statut raciale à la personne, en revanche les approches en I-2 montrent que la génétique a les moyens de le faire : le fait que certains gènes seraient des marqueurs de population d'une part et que les données génétiques permettent de construire des groupes distincts de populations ouvrent la voie à un certain retour sur la race qui mérite d'être discuté.

## **II – La génétique généalogique personnalisée et ses effets**

II-1) Quelle origine ? Actuellement, la plupart des sociétés commerciales proposent des tests d'origine dont il n'est pas certains qu'ils répondent réellement à la demande de leurs clients (DTC) dont les motivations peuvent se révéler extrêmement diverses. L'origine peut faire référence aux lieux géographiques d'où proviennent leurs ancêtres, à leur identité culturelle ou leur background ethnologique, ou à la recherche d'une possible ascendance célèbre, valorisante ou pas (Christophe Colomb, Jesse James, Gengis Khan...). Elle peut concerner les générations généalogiquement les plus proches ou au contraire les origines lointaines. Il est à craindre que ces sociétés ne répondent pas avec toute la rigueur souhaitée à toute la diversité de ces demandes, dans la mesure où lorsqu'elles parlent d' « origine », d'« ancestry » ou de « roots », elles ne font jamais que de comparer les gènes de leur client à ceux répertoriés dans diverses banques de données, comme si les populations actuelles avaient la même géo-référence depuis toujours et étaient génétiquement immuables depuis de nombreuses générations. Pour le généticien, le terme « origine » est pourtant à la fois clair et restreint : il fait référence à la descendance issue de populations ancestrales qui ne sont qu'*inférées*, sur la base de marqueurs précis et grâce à l'application d'algorithmes spécifiques d'estimation, et en référence à des populations particulières actuelles (Lee et al., PNAS, 2009,325,p35-6). Ce sont là des limites honnêtement proclamées, mais qui risquent de ne pas satisfaire véritablement la curiosité des consommateurs, ce qui peut amener les fournisseurs de tests à en flouter le message pour en minimiser l'étendue et ne pas risquer de restreindre leur marché potentiel.

II-2) Quels produits ? Les nombreuses sociétés commerciales se doivent de faire miroiter les performances et la qualité de leurs tests d'origine. Ainsi, le site « DNA Tribes » propose une « *genetic ancestry analysis* », à partir de STR, « *marqueurs génétiques développés par le FBI pour identifier les personnes* ». Voilà bien une référence solide...! La plus grande limitation des produits qui sont financièrement accessibles au grand public, en l'état actuel, est de ne proposer que des diagnostic d'origine à partir des polymorphismes de l'ADNm ou du ChrY, restant très discret sur le fait que de tels marqueurs ne renseignent que sur un nombre extrêmement faibles d'ancêtres possibles, et seulement sur les lignées maternelles et paternelles, soit deux

ancêtres possibles parmi les  $n^2$  à la  $n$  ième génération (en l'absence de boucles de consanguinité). Autrement dit quasiment rien du point de vue génétique. Mais, pour les sociétés commerciales, la divulgation de cette faiblesse ne reviendrait-elle pas à scier la branche sur laquelle elles font leur profit ? En tout cas, si l'on en croit les réactions que publient leurs clients, la satisfaction semble totale. Par exemple, un homme qui pensait avoir un arrière grand parent venu de l'est se satisfait pourtant fort bien de ses nouvelles origines sont « pur british » ; un allemand découvre que ses ancêtres (au moins un !) sont écossais, expliquant du même coup l'origine étymologique de son nom de famille... Une auvergnate se fait confirmer que ses ascendants sont bien originaires du centre de la France etc... Certaines personnes adoptées retrouvent la trace de leurs ascendants biologiques.

Pour une contribution financière plus importante (de l'ordre de 1000\$) les sociétés commerciales proposent des diagnostics plus précis, avec une estimation de la proportion de gènes venant de telle ou telle population (voir plus haut).

Pour une analyse comparée de quelques produits les plus connus (African Ancestry, Ethnoancestry, 23andme, Family tree DNA, DNA Tribes, Warrior Roots...), voir <http://www.isogg.org/eochart.htm>

II-3) Les implications de ces tests au niveau social, politique et éthique sont nombreuses et méritent d'être analysées.

II-3-a) D'abord le fait d'attribuer une origine, avec toute l'autorité qu'est sensée conférer la Science (la génétique du moins...) peut amener une contradiction avec les croyances ou les certitudes sur les origines, telles qu'elles sont intégrées, souvent depuis des générations, dans la culture d'une personne. Cette contradiction a parfaitement été relevée par Debra Harry, en abordant la question de la domination d'un système de connaissance (la Science) sur un autre (la tradition), reléguant cette dernière dans l'obscurantisme. Par ailleurs, le fait de fixer à une peuple une origine géographique qui n'est pas nécessairement le lieu qu'il occupe présentement comporte des risques, « *Nos droits*, écrivait-elle dans le magazine *New Internationalist* en décembre 2005, *sont fondés sur le fait que nous habitons originellement sur les territoires que nous occupons. Une affirmation qui viendrait contredire l'aboriginalité de certains peuples indigènes pourrait constituer une menace sérieuse. Même si les résultats de ce type de recherche sont spéculatifs, nous n'avons aucun doute sur le fait que ces découvertes seront exploitées comme une arme politique contre nous.* »

II-3-b) Les tests peuvent ne pas être vécus sans conséquences psychologiques ou sociales. Parmi les moindres, rappelons le cas de cet homme dont la déception reste limitée et qui écrit « *Mon plus grand rêve était de savoir si j'avais des racines juives [...] Au début, j'étais un peu déçu du fait que je n'avais pas d'origines juives. Je trouvais les Celtes et les Germains trop normaux. J'attendais quelque chose de plus extraordinaire. Le fait que ma mère ait une lignée Sami m'a réjoui [...] Ce que j'aimerais par la suite c'est pousser mes tests plus loin pour peut-être trouver des origines juives pour pouvoir revendiquer une filiation avec le roi David* ». Rappelons l'exemple rapporté par Sandra Soo-Jin Lee du cas d'un « self-identified » noir américain qui dû reconsidérer son identité ethnique après que la société prestataire lui eut trouvé 57% de gènes européens, 39% de « native Américian », 4% est asiatique, mais 0% d'Africain...

Sur le plan plus politique, certains Afro-américains revendiquent la citoyenneté de pays africains où les tests d'origine auraient localisé leurs ancêtres.

II-3-c) Les tests génétiques d'origines peuvent indirectement servir à confirmer ou infirmer des liens de parenté, ou à proposer une origine à ceux qui ont été adoptés. Rien n'interdit de faire passer des tests génétiques à ses enfants ou à ses parents pour constater la validité des filiations biologiques, en parfaite illégalité, du moins avec le droit français. Certains sites commerciaux, sans aller jusqu'à encourager explicitement ce genre de contournement de la loi, proposent néanmoins la possibilité de mettre en relation des personnes qui auraient le même profil génétique. « *Trouver des cousins génétiques. [...] vous pouvez retrouver les noms et adresses e-mail des personnes qui vous correspondent génétiquement en lignée paternelle et maternelle. La banque de donnée est mise à jour constamment et il y aura donc de plus en plus de personnes vous correspondant génétiquement* » (site web d'Igenea). Et ce témoignage (sur IGenea) : « *Personne dans la banque de données n'avait 36 ou 37 marqueurs en commun avec moi. J'ai perdu 6 mois à tourner en rond, puis j'ai reçu un mail qui m'a coupé le souffle. Le test FT ADN-Y montrait que j'avais 37 marqueurs d'ADN en commun avec quelqu'un. Cela voulait dire qu'avec une probabilité de 90% nous avons un ancêtre en commun dans les 5 dernières générations. J'ai évidemment tout de suite contacté mon cousin au x-ème degré, il est mon seul consanguin. Et moi, le bébé adopté qui a grandi et est devenu un homme sans origines, je suis très heureux. Mon parent consanguin et moi, nous nous écrivons régulièrement des mails. Nous n'avons pas seulement un lien génétique, nous nous entendons aussi très bien* » Si ce témoignage montre l'aspect positif de cette démarche, qu'en est-il des témoignages, jamais relatés, qui ne le sont pas ?

II-3-d) La question de la propriété de l'ADN envoyé par les clients se pose également. En effet, ces échantillons expédiés aux diverses sociétés servent-ils à d'autres propos que simplement satisfaire la demande pour laquelle les « DTC » ont payé ? Ou servent-ils à d'autres fins ? La réponse pourrait ne pas être simple. En particulier, il est probable que ces échantillons servent à enrichir le panel sur lequel se fonde toute la statistique d'attribution d'origine. Pour les compagnies qui font également du diagnostic de risques pour tout un ensemble de maladies, ces échantillons pourraient également être utilisés comme panel témoin, géographiquement localisé. Toutes ces situations sont difficilement contrôlables, si bien que la propriété de l'ADN du client reste problématique. [Cf en détail le contrat de 23andme, par exemple : « *You agree that this license includes a right for 23andMe to make such content available to other companies, organizations, or individuals with whom 23andMe has relationships, and to use such content in connection with the provision of those services* ».

II-4) Autres aspects de la génétique « récréative ». Les possibilités de la génétique pour prédire des traits psychologiques ou comportementaux ont été prônées depuis longtemps. Citons le livre bien connu de Léone Bourdel sur « Sang et Tempérament » qui, malgré son ancienneté (1962), fait encore les beaux jours sur la toile au travers de nombreux épigones, tout comme les ouvrages préconisant des régimes alimentaires en fonction des groupes sanguins (cf. les travaux du Docteur Peter d' Adamo)... De nos jours de tels outils génétiques comme les groupes sanguins paraissent un peu frustes ou obsolètes, complètement dépassés par la puissance triomphale de la génétique moléculaire et du séquençage du génome humain (*the revolutionary tool to complete your ancestry with DNA ; 23andme*)... Mais l'esprit reste. De fait, ces sites proposent des tests génétiques pour « Avoidance of errors », « Measure of Intelligence », « Food preference »..., précisant, toutefois, que ces tests sont basés sur des travaux de haute qualité, certes, mais n'ayant qu'une évidence scientifique limitée dans la mesure où ils ne sont pas encore tous répliqués.

## Quelques conclusions

- Le fait de posséder et partager avec d'autres certains déterminants génétiques peut renforcer le sentiment d'appartenance et favoriser l'illusion d'une identité communautaire de nature « scientifique ». Dans le même temps, l'absence de partage de tels déterminants génétiques peut conduire à construire le statut de « l'autre », de l'étranger.
- Assimilation abusive entre origines génétique/appartenance ethnologique/culturelle ou citoyenneté... et confusions entre « ethnicity », « race », « ancestry »
- Conflit possible entre une certaine tradition culturelle des origines et les connaissances « rationnelles » des « scientifiques »
- Sous couvert de recherche des origines : recherche de paternité, vérification de filiation, risques d'exclusion, hors de tout contrôle juridique
- Réduction de l'être à son « essence » génétique

Pour info : Recommandations de la ASHG

[http://www.ashg.org/pdf/ASHGAncestryTestingStatement\\_FINAL.pdf](http://www.ashg.org/pdf/ASHGAncestryTestingStatement_FINAL.pdf)



(Charlie Hebdo, 6 août 2008)