

# *Rôle et signification du polymorphisme génétique*

## *dans les populations humaines*

---

Josué Feingold

Une des caractéristiques de l'espèce humaine est sa grande diversité : les individus qui la composent diffèrent les uns des autres. Les polymorphismes génétiques permettent d'étudier cette diversité. Actuellement on connaît une centaine d'entre eux qui peuvent être étudiés par l'analyse du phénotype (1) des différents individus, et leur nombre est beaucoup plus considérable si l'on considère les polymorphismes de l'ADN. La découverte de tous ces polymorphismes a conduit les généticiens à réévaluer leur conception concernant la structure du

---

(1) Le phénotype d'un sujet est l'ensemble des caractéristiques d'un individu qu'on peut étudier, certaines nécessitant des investigations complexes. Le génotype est la constitution génétique de l'individu.

génomique des êtres vivants supérieurs et à remettre en cause en partie la théorie néo-Darwinienne de l'évolution.

### ***Définition des polymorphismes***

Si l'on considère un locus donné, on dit qu'il y a *polymorphisme* si la fréquence des hétérozygotes est supérieure à 1 ou 2 % dans la population étudiée. Actuellement on connaît, comme nous l'avons dit, une centaine de polymorphismes qu'on peut répartir en plusieurs classes :

#### ***Les polymorphismes antigéniques***

Ce sont les plus anciennement connus. Rappelons que le premier polymorphisme découvert est celui du groupe ABO. Il l'a été par Landsteiner au début de ce siècle, en même temps que la redécouverte des lois de Mendel. Par la suite, et après un long intervalle, de nombreux polymorphismes concernant les antigènes érythrocytaires ont été découverts. Citons les systèmes MNSs, Rh, Pn, Kk, etc...

Mais les polymorphismes antigéniques ne concernent pas uniquement les globules rouges. En effet, on a décrit les systèmes Gm et Km qui concernent les immunoglobulines et, surtout, le système HLA dont les premiers antigènes ont été découverts par Jean Dausset à Paris. Ce système, qui concerne l'ensemble des cellules nucléées, est extrêmement complexe et très polymorphe. Une dizaine de loci étroitement liés sont actuellement connus et, de plus, de nombreux allèles ont été décrits à chaque locus. Par exemple plus de 30 allèles ont été décrits au locus B.

*Les polymorphismes enzymatiques et protéiques*

C'est grâce à l'électrophorèse qu'ils ont été mis en évidence. Ils concernent le plus souvent les protéines sériques ou les enzymes des éléments figurés du sang.

*Les polymorphismes chromosomiques ou hétéromorphismes*

L'étude du caryotype a montré qu'il existait des différences morphologiques entre les chromosomes d'une même paire. Ces variations liées à l'échéromatine se transmettent d'une génération à l'autre et peuvent être considérées comme des polymorphismes.

*Les polymorphismes pbarmacogénétiques*

Le métabolisme de certains médicaments est sous la dépendance de systèmes génétiques simples qu'on peut mettre en évidence par l'étude de dose traceuse du produit. Citons à titre d'exemple le polymorphisme de l'acétyl-transferase hépatique celui de l'hydroxylation de la debriso-quine et de la pseudo-cholinestérase sérique.

*Maladies héréditaires et polymorphismes*

La fréquence de certaines maladies récessives autosomiques ou liées au sexe est si élevée que la fréquence des hétérozygotes dépassent 2 %. Citons la drépanocytose en Afrique, la  $\beta$ -thalassémie dans certains pays méditerranéens, la mucoviscidose en Europe, le déficit en G6-PD en Afrique et en Asie.

### *Les polymorphismes de l'ADN*

La biologie moléculaire a montré qu'au niveau de l'ADN non codant il existait des variations entre les individus toutes les 100 à 300 bases. Rappelons que l'ensemble du génome haploïde est formé de  $3 \times 10^9$  paires\* de bases et que l'ADN non codant représente plus de 90 % du génome. Cependant, une infime partie de ces polymorphismes peuvent être étudiés car leur mise en évidence n'est possible que s'ils font partie d'un site de restriction et que l'on possède une sonde correspondant à la région du site.

### *Importance des polymorphismes*

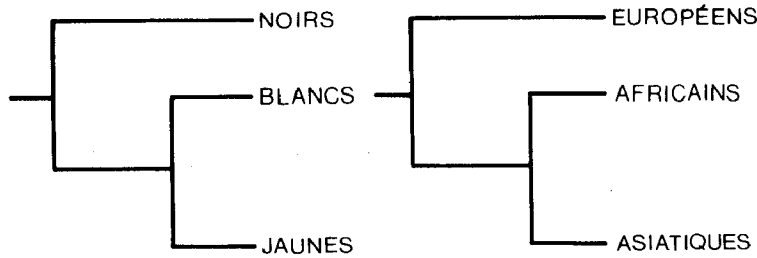
On estime actuellement que le génome humain comprend environ 60.000 gènes (ADN codant). Parmi eux on pense qu'un tiers sont polymorphes ; ce chiffre (1/3) a été calculé par Lewontin à partir de ses recherches sur les polymorphismes de la drosophile. Actuellement on connaît environ 120 polymorphismes soit 0,6 % du total (120/20.000). Le nombre de polymorphismes de l'ADN serait à  $10^7$  ( $3 \times 10^9 / 300$ ). Il s'agit d'un chiffre théorique car, en pratique, le nombre de polymorphismes étudiable est actuellement de l'ordre de 200 à 300.

Ces polymorphismes sont à l'origine de la très grande diversité qu'on note dans les populations humaines.

---

\* Les nombres exprimés en "puissance de 10" ;  $10^n$  indiquent le nombre n de zéros suivant le chiffre indiqué. Ainsi :  $3.10^9$  signifie 3, suivi de neuf zéros, soit 3.000.000.000, 3 milliards.

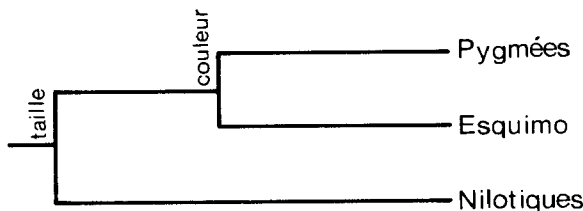
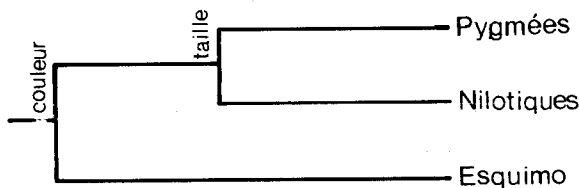
La notion de race correspond à deux phénomènes : l'évidence et le besoin de classer. Mais les choses ne sont pas si simples. Ainsi, si la couleur de la peau est une évidence première, la persistance -ou non persistance de la lactase- (c'est-à-dire la persistance de la digestion du lait après le sevrage) en est une autre. Un classement correspondant à la première évidence semble rapprocher les Jaunes et les Blancs ; un second concernant la persistance de la lactase les éloigne et rapproche Africains et Asiatiques. Comment, dès lors, classer les races ? Cette question n'a pas de réponse. D'autant que bien d'autres critères de classement peuvent entrer en jeu. Ainsi la taille et la couleur, selon que l'on choisit l'ordre ; la couleur, puis la taille, ou l'ordre inverse : c'est le classement par taille qui importe alors, et les regroupements ne sont plus les mêmes.



**Classement des hommes selon...**

la couleur de la peau

la persistance de la lactase



A titre d'exemple, considérons 200 loci polymorphes, chacun étant diallélique (à deux allèles : cf Cahier n° 4 : Mémento de Génétique). Par locus, 3 génotypes sont possibles ; si l'on considère l'ensemble des 200 loci,  $3^{200}$  individus différents sont concevables. Ce dernier nombre est très supérieur au nombre total d'atomes existant dans l'univers. Le nombre est sous-estimé car, d'une part, le nombre de loci polymorphes est supérieur à 200 et, d'autre part, de nombreux loci sont poly-alléliques. Ces variations quasi infinies des individus sont le résultat de la *recombinaison génétique* que permet la sexualité. Chez les êtres inférieurs haploïdes à reproduction non sexuée, le même nombre de mutants conduit à une diversité très nettement plus réduite. Comme le souligne André Langaney\*, la sexualité accélère à l'extrême le rythme autrement très lent de l'évolution.

### ***Mesure du polymorphisme et de la diversité entre populations***

La fréquence des hétérozygotes (dialléliques) permet de mesurer la contribution du locus considéré à la diversité de la population étudiée. La fréquence des hétérozygotes est celle donnée par la célèbre loi de Hardy - Weinberg qui suppose que, pour le caractère étudié, les mariages se font au hasard (panmixie). Faisons ce calcul pour le système MN qui est un polymorphisme à deux allèles. Si la fréquence du gène M est  $n = 0,60$  et celle du gène N  $q = 0,40$ , la fréquence des hétérozygotes est égale à deux fois le produit des fréquences des gènes soit  $2nq = 0,48$  (la fréquence des 3 génotypes MM, MN et NN est donnée par le développement de  $(n+q)^2$ ). Pour

---

\*"Le sexe et l'innovation" Le Seuil.

un système à deux allèles, la fréquence des hétérozygotes est maximale si les 2 gènes ont même fréquence. La fréquence des hétérozygotes augmente avec le nombre d'allèles du système, et d'autant plus que les fréquences sont égales. Le très grand polymorphisme observé aux différents loci du système HLA explique pourquoi cette région contribue beaucoup à la diversité génétique. Rappelons que si un système est formé de 10 allèles d'égale fréquence, 90 % des individus sont hétérozygotes ; pour un système à 20 allèles, 95 % des individus le sont. Comment se répartit la diversité génétique ? Est-elle inter- ou intra-population ?

Considérons à titre d'exemple, deux situations extrêmes. Dans le premier cas, une population ne possède que le gène M, tous les sujets sont homozygotes MM, tandis qu'une autre ne possède que le gène N. Dans ce cas toute la diversité est inter-populations. Dans le second cas la fréquence des gènes M et N est la même dans les 2 populations : la diversité est tout entière intra-population.

Dans les populations humaines il est rare que l'on soit dans une de ces deux situations extrêmes. Lewontin a analysé la diversité des populations humaines en utilisant un critère dérivé de celui de la fréquence des hétérozygotes. L'étude a porté sur 17 loci et de nombreuses populations. Il a montré que la diversité intra-population est plus grande que celle inter-populations ou inter-races. *Il a estimé que 85 % de toute la diversité humaine se situe entre les individus d'une même population ; le reste se répartissant à peu près également entre les populations d'un même continent et les différentes grandes races.* Ce fait, comme le souligne A. Jacquard, constitue la meilleure preuve de l'inanité des tentatives de définition des races : paradoxalement, ce qui n'empêche pas que ce soit exact scientifiquement, un Auvergnat par exemple peut être plus "proche" génétiquement d'un Kabyle ou d'un Camerounais que de son voisin de village. Cela dépend du critère (parmi une multitude) de comparaison choisi.

### ***Origine de la variabilité et son maintien***

Les mutations germinales, c'est-à-dire celles survenant dans un gamète, sont à l'origine de nouveaux gènes et donc des polymorphismes. Les mutations peuvent être géniques ou chromosomiques. Dans ce paragraphe nous ne considérons que les premières. Dans tous les cas il s'agit d'une modification de structure *irréversible* et qui peut être transmissible d'une génération à l'autre. Le taux de mutation moyen a été estimé à  $3 \times 10^{-7}$  \* par gamète et par génération. Si l'on se souvient que chaque gamète comporte 60.000 gènes, la probabilité pour qu'il soit porteur d'une nouvelle mutation est  $1/200$  ( $6 \times 10^4 \times 3 \times 10^{-7} \approx 0,5 \%$ ).

Le taux de mutation par base a été estimé à  $10^{-9}$ , l'ADN de chaque gamète étant composé de  $3 \times 10^9$  paires de bases, chacun d'eux porte en moyenne 3 nouvelles mutations situées la plupart du temps au niveau de l'ADN non codant.

Sur le plan expérimental, on a montré que les radiations ou les mutagènes chimiques pouvaient augmenter les taux de mutation. Cependant, il est actuellement impossible de surveiller dans les populations humaines les taux de mutation pour savoir si notre "environnement" devient plus mutagène. En effet, le calcul du taux de mutation nécessite le dépistage des sujets porteurs de néomutations pathologiques ou non, ce qui oblige à étudier un très grand nombre d'individus. Le prix d'un tel dépistage est actuellement très élevé.

---

\*  $(3 \cdot 10^{-7} = \frac{3}{10^7}$  , soit 3 sur 10 millions)



Dans la conception classique du génome, la plupart des loci sont monomorphes, c'est-à-dire ne pouvant être occupés que par un type de gènes ; tous les individus sont par conséquent homozygotes. Les loci polymorphes dans cette conception sont rares et sont maintenus par des mécanismes sélectifs 'simples. Le plus classique est l'avantage sélectif des hétérozygotes. L'exemple le plus connu est celui de la drépanocytose. Il s'agit d'une anémie hémolytique mortelle due à l'hémoglobine anormale S. La fréquence de cette maladie récessive autosomique est supérieure à 1 % chez les jeunes enfants d'Afrique centrale. La fréquence des hétérozygotes est supérieure à 20 % dans cette région. Le maintien d'un gène délétère à une telle fréquence s'explique par un avantage sélectif des hétérozygotes. Sur le plan biologique, selon Allison, les hétérozygotes résisteraient mieux au paludisme à *Plasmodium falciparum* qui est la forme la plus grave de la maladie, que les sujets ayant une hémoglobine normale.



Bédiks du Sénégal oriental

Au nombre d'environ 1.500, les Bédiks du Sénégal présentent la particularité d'avoir une proportion importante de leur population porteuse du gène S, marqueur sanguin de l'anémie faciforme du nom de drépanocytose. Cette drépanocytose présente toutefois une étonnante caractéristique puisque, dans les régions fortement impaludées, elle offre une relative protection contre le paludisme.

Cette conception classique du génome a été remise en cause par la découverte des nombreux polymorphismes dont nous avons donné une rapide description. La théorie néo-darwinienne explique mal le maintien d'un si grand nombre de polymorphismes. Il est, en particulier, impossible d'expliquer ce fait par un avantage sélectif des hétérozygotes.

L'ensemble de ces faits conduit certains généticiens à remettre partiellement en cause la théorie néo-darwinienne de l'évolution. A côté de la sélection naturelle, il faut faire jouer un rôle important au hasard dans le maintien de la diversité. Les aspects théoriques de cette approche ont été développés par Kimura et son école.

En outre, les faits sont sûrement moins simples qu'on ne le croyait. Reprenons l'exemple de la drépanocytose ; dans les populations où le gène est fréquent, on trouve d'autres caractéristiques : citons le gène de l'hémoglobine C, celui de l' $\alpha$ -thalassémie du déficit en glucose-6-phosphate deshydrogénase. Les facteurs sélectifs ont donc vraisemblablement concerné l'ensemble de ces gènes et, de plus, des interactions entre ces systèmes doivent exister. Récemment, on a par exemple montré que la drépanocytose était moins grave chez les sujets  $\alpha$ -thalassémiques. Quoiqu'il en soit, l'ensemble des découvertes de la biologie moléculaire nécessitera une nouvelle théorie de l'évolution.

### ***Conclusion***

*Les progrès de la biologie ont montré que les polymorphismes, loin d'être l'exception comme on le croyait, il y a trente ans, sont fréquents. Ce fait a obligé les généticiens à revoir leur conception du génome et à modifier la théorie de l'évolution. En ce qui concerne les populations humaines, l'étude des polymorphismes a montré que la notion de "race" n'avait aucun fondement.*

**Josué FEINGOLD**  
Directeur de Recherche à l'INSERM