

Diversité du génome humain et populations en voie d'extinction

Le programme Génome Humain soulève depuis plusieurs années enthousiasme et discussions. Son objectif ultime reste, à terme, le séquençage total de l'ADN, soit 3×10^9 pb. L'ampleur de la tâche ne laisse pas place à un autre aspect de la génétique, la diversité des populations et des individus. Or apprécier cette diversité est capital pour la connaissance de l'évolution : origine des populations actuelles, migrations préhistoriques, sélection naturelle, fréquence des types de mutations. Chaque population, dans la mesure où elle diffère des autres, ouvre ainsi une « fenêtre sur le passé ». De plus, des retombées pratiques sont à espérer : certaines ethnies sont moins sensibles que d'autres à des anomalies endogènes (diabète par exemple) ou exogènes, et l'analyse de leurs gènes peut fournir un éclairage sur les raisons de ces différences. Or, dans la période même où les outils de la biologie moléculaire commencent à permettre l'exploration de génomes individuels, nombre de populations sont en voie de disparition, par extinction, ou noyées par mélange avec d'autres. Il y a donc urgence, s'il ne s'avère pas possible de les sauver, à recueillir du moins leur mémoire. C'est la raison d'un plaidoyer passionné, qui réunit deux groupes scientifiquement rivaux : celui de Luigi Cavalli-Sforza (Stanford, CA, USA) qui mène depuis des années une véritable croisade sur le thème de la diversité humaine, et celui d'Allan Wilson* (Berkeley, CA), un anthropologiste. Ils proposent, dans deux articles parus dans *Genomics* [1, 2], relayés par des éditoriaux dans *Nature* [3] et dans *Science* [4], de lever des fonds pour étudier les populations en voie d'extinction avant qu'il ne soit trop tard.

A. La diversité humaine au niveau moléculaire

Elle ne peut encore être évaluée que très approximativement. On connaît à l'heure actuelle chez l'homme environ 2 300 polymorphismes, et, d'après la taille des zones explorées, on estime qu'il pourrait en exister 1 000 fois plus. La vitesse de mutation par nucléotide serait de 1.10^{-9} par génération. Le génome comporte une variation toutes les 100 à 1 000 bases, en moyenne environ toutes les 500 bases, avec des variations considérables suivant les régions. Kimura (*in* [2]) estime que deux individus non apparentés doivent différer entre eux par plusieurs millions de nucléotides, mais aussi que des altérations dans 95 % du génome doivent rester sans effet nocif.

Les polymorphismes actuellement connus sont probablement ethniquement biaisés car ils proviennent surtout d'études sur des sujets d'origine européenne. Dans le cas des mutations ponctuelles, la comparaison avec le chimpanzé montre que la plupart d'entre elles ne se sont produites qu'une fois : en général, le locus polymorphe humain est monomorphe chez le chimpanzé, et l'allèle identique à celui du chimpanzé est vraisemblablement l'allèle humain d'origine. Toutefois des mutations avantageuses peuvent être récurrentes ; l'exemple le plus connu est celui de la drépanocytose, qui est survenue au moins quatre fois indépendamment dans des ethnies différentes, comme le montre la différence des haplotypes dont les mutants sont porteurs (*m/s* n° 1, vol. 3, p. 54).

* Le professeur Allan Wilson, un des plus grands anthropologistes de notre temps, animateur de l'équipe ayant utilisé l'ADN mitochondrial pour dater la vie de la « mère commune » de l'espèce *Homo sapiens sapiens* (*m/s* n° 5, vol. 2, p. 278 et n° 8, vol. 7, p. 869) vient de décéder des suites d'une leucémie, aux États-Unis.

B. Stratégie

Elle comporte un objectif ethnologique, une définition des techniques envisagées, et des impératifs financiers.

1. Les ethnies visées. Ce sont celles qui ont longtemps été isolées, à l'abri de barrières géographiques, ou linguistiques et culturelles. Cavalli-Sforza *et al.* [2] pensent qu'il faudrait en étudier au moins une centaine. Le *Tableau I* présente quelques ethnies que l'on espère particulièrement instructives. Il n'est pas envisagé, en raison de la structure des laboratoires impliqués, d'organiser des descentes dans les villages pour prélever du sang. Au contraire, les auteurs comptent sur des réseaux de chercheurs ou d'ethnologues locaux pour les aider dans leur cueillette. Ils sont d'ailleurs conscients de deux écueils ; l'un est technique : lorsqu'on doit prélever dans une zone isolée, il est difficile d'amener les échantillons là où l'on doit les traiter dans le temps requis (pour, idéalement, immortaliser une lignée ; sinon, au moins, isoler les lymphocytes). L'autre est éthique : on doit éviter ce qui pourrait être considéré comme une exploitation et examiner avec les ethnologues quel bénéfice on pourrait apporter à la population locale, en termes par exemple d'approvisionnement en médicaments.

2. Techniques d'étude. Aux moyens très limités de jadis (études de protéines, caractères immunologiques), s'est naturellement substituée l'analyse de l'ADN. Mais deux grandes options sont possibles [4], qui continuent à opposer les protagonistes de cette recherche, et qui conditionnent des approches différentes : doit-on s'attacher, comme le préconise Cavalli-Sforza, à l'ADN nucléaire, permettant d'obtenir une

Tableau I	
EXEMPLES DE POPULATIONS A EXPLORER	
Ethnies	Pays
Pygmées de l'est	Zaïre
Pygmées de l'ouest	République Centrafricaine
! Kung	Namibie-Botswana
Bushmen	Afrique du Sud
Aïnus	Nord du Japon, îles Kouriles
Iles Andaman	Voisines de la Malaisie
Polynésiens	Iles de la Polynésie
Aborigènes	Australie, Mélanésie
Amérindiens	États-Unis, Amérique latine
Indiens Yanomani	Ethnie particulière d'Amazonie
Kurdes	Turquie de l'Est
Lapons	
Basques	

vue plus profonde, mais laborieuse et obligeant à des choix difficiles, ou, avec Wilson, préférer l'ADN des mitochondries, d'étude plus aisée et qui évolue plus vite ?

a. *ADN nucléaire*. On ne saurait envisager une analyse complète, aussi faut-il imaginer des régions représentatives du génome. Bowcock *et al.* [5] se sont d'abord adressés à des polymorphismes de restriction ; l'examen de 100 polymorphismes distincts a été pratiqué dans des populations originaires de quatre continents, avec pour objectif l'appréciation des rôles respectifs de la dérive génétique et de la sélection dans l'évolution humaine. L'ambition du projet s'est ensuite accrue, et les auteurs [2] proposent maintenant l'analyse de séquences. Leur choix est de systématiser des « strates » d'ADN comportant par exemple : exons, séquences flanquantes, introns, pseudogènes, mais aussi séquences de gènes codant pour des ARN, et des séquences répétitives, Alu, L1, centromères, télomères, et répétitions nucléotidiques. Le nombre total de ces séquences dépendra avant tout des moyens à disposition. Lorsque l'on découvre des allèles différents d'un même gène nucléaire, il est habituel que tous les allèles existent dans toutes les populations. Ce qui varie, c'est la fréquence des allèles, variation qui suivant les cas est faible ou importante. Cette considération dicte les conditions de l'échantillonnage : il faut, dans chaque population, un nombre assez élevé de

sujets, une centaine par exemple, et l'on ne saurait se contenter de coups de sonde. Par suite, le nombre de populations ne peut qu'être limité : nous l'avons dit, l'ambition du groupe est d'atteindre 100 ethnies, avec une extension idéale jusqu'à 500.

b. *ADN mitochondrial (ADNmt)*. Cet ADNmt, avec ses quelques 16 000 pb, est beaucoup plus facile d'accès. Son étude est encore simplifiée du fait que la plupart de la diversité se localise dans la zone dite de la boucle de développement, qui ne comporte pas de gènes de structure, mais qui commande transcription et réplication (figure 1). Cette région de 1 122 bases a été analysée chez 189 individus dont 121 Africains [6]. Les auteurs ont trouvé 201 polymorphismes (179 ponctuels, 22 changements de longueur) définissant 135 types, dont seulement 16 ont été trouvés plus d'une fois. Outre des confirmations concernant l'origine africaine de l'espèce humaine (*m/s n° 5, vol. 2, p. 278*), et l'ancienneté probable de l'ancêtre commun autour de 200 000 ans (*voir cependant pour discussion m/s n° 8, vol. 7, p. 869*), cette diversité considérable explique la méthode préconisée : il est probable qu'un allèle donné ne se rencontre que dans des zones géographiques limitées. Il faut donc couvrir la plus grande surface possible, par exemple prélever tous les 50 ou 100 km tout autour du globe, mais chaque échan-

tillon n'a pas à être très important. Un moyen terme devra être trouvé entre ces deux conceptions.

3. *Aspects financiers*. Nous n'y ferons qu'une allusion mais ils conditionnent évidemment l'ensemble de l'entreprise. Le coût du programme prévu est d'environ 10 millions de dollars pendant cinq ans. Comme c'est à l'occasion du programme Génome Humain que les auteurs développent leur plaidoyer, ils escomptent que l'organisation HUGO, qui n'est pas par elle-même un bailleur de fonds, leur apportera une aide pour trouver des subventions, ce qui semble en bonne voie.

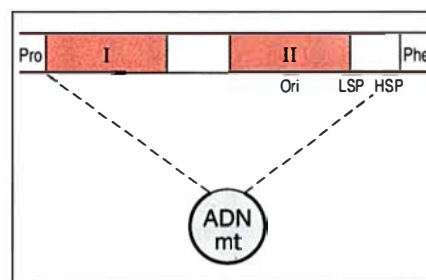


Figure 1. **La région de contrôle de l'ADNmt**. Le cercle représente l'ensemble du génome mt, la ligne fine la zone de contrôle agrandie au-dessus. Ont été indiqués les gènes flanquants d'ARNt Pro et Phe, les deux segments hypervariables séquencés dans ce travail sont hachurés. On voit également l'origine de réplication Ori, et les promoteurs d'initiation de la transcription pour les brins léger (LSP) et lourd (HSP). (D'après [6]).

Bien entendu, les auteurs [1] s'adressent à toutes les sources possibles, nationales et internationales. Ils espèrent en outre que des collègues, intéressés par les mêmes problèmes, viendront se joindre à eux.

C. Conclusions

Dans la discussion de ce projet nous avons délibérément négligé les questions éthiques fondamentales. On ne peut manquer de remarquer que le principal argument de l'urgence est la disparition en cours de nombre des ethnies dont l'étude est envisagée. Il

est amer de constater que rien, ou bien peu, n'est et ne sera fait pour sauvegarder l'existence de tant de populations destinées à être étouffées par la civilisation ou à se fondre en elle, diminuant d'autant la richesse du patrimoine génétique de l'humanité. C'est donc un pis-aller, lui-même déjà difficile à réaliser, que de se borner à perpétuer la mémoire génétique de peuples destinés à une extinction prochaine. On ne peut cependant rendre responsables de cette situation les biologistes qui sont à l'origine de ce projet, qui n'ont aucun moyen d'agir sur un problème purement politique. Aussi peut-on, malgré tout, appuyer leur démarche, en espérant que les résultats seront suffisamment passionnants pour justifier leur entreprise, et même — c'est bien peu probable — pour peut-être inciter les responsables à prêter davantage d'attention au drame que vivent les ethnies dont s'évanouit l'existence, et peut-être, si rien n'est fait, bientôt le souvenir.

J.C.D.

1. Cavalli-Sforza LL, Wilson AC, Cantor CR, Cook-Deegan RM, King MC. Call for a worldwilde survey of human genetic diversity : a vanishing opportunity for the Human Genome project. *Genomics* 1991 ; 11 : 490-1.
2. Bowcock A, Cavalli-Sforza L. The study of variation in the human genome. *Genomics* 1991 ; 11 : 491-8.
3. Diamond JM. A way to world knowledge. *Nature* 1991 ; 352 : 567.
4. Roberts L, Gibbons A. A genetic survey of vanishing peoples. *Science* 1991 ; 252 : 1614-7.
5. Bowcock AM, Kidd JR, Mountain JL, et al. Drift, admixture, and selection in human evolution : a study with DNA polymorphisms. *Proc Natl Acad Sci USA* 1991 ; 88 : 839-43.
6. Vigilant L, Stoneking M, Harpending H, Hawkes K, Wilson AC. African populations and the evolution of human mitochondrial DNA. *Science* 1991 ; 253 : 1503-7.

m/s n° 9, vol. 7, novembre 91

■■■■ BRÈVES ■■■■

■■■■ **Synthèse intestinale d'apolipoprotéine B chez les sujets a- β -lipoprotéïnémiques.** L'a- β -lipoprotéïnémie est un désordre génétique rare, autosomique récessif, caractérisé par l'absence d'apolipoprotéine B dans le plasma, absence liée au défaut de sécrétion de lipoprotéines contenant l'apo B par le foie et l'intestin ; ceux-ci sont le siège d'une rétention lipidique, le tableau clinique associant des signes de malabsorption lipidique et de vitamines liposolubles — notamment de la vitamine E —, une neuropathie périphérique, une rétinite pigmentaire et une acanthocytose des érythrocytes. Chez le sujet normal, le foie synthétise une apo B100 de plus de 4 500 acides aminés, et l'intestin synthétise une apo B48 de 2 152 acides aminés identiques à la séquence aminoterminal de l'apo B100. L'incapacité des sujets a- β -lipoprotéïnémiques à synthétiser l'apo B, notamment intestinale, a été suggérée par de précédents travaux. Un récent travail [1] a étudié quantitativement la synthèse d'apolipoprotéine B et les taux d'ARN messenger de l'apo B dans la muqueuse duodénale de sujets normaux et atteints d'a- β -lipoprotéïnémie. Après incubation *in vitro* en présence de leucine-[H3], les biopsies intestinales de trois des quatre sujets a- β -lipoprotéïnémiques était le siège d'une synthèse d'apo B immunoprécipitable de mobilité identique (en électrophorèse en gel de dodécyl sulfate de sodium) à l'apo B normale. Par radio-immunologie, le contenu intestinal en apo B des malades a été estimé à, environ, 15 % de celui des sujets normaux. En revanche, le taux d'ARN messenger de l'apo B était 3 à 20 fois supérieur à celui trouvé dans une muqueuse normale. La synthèse et les taux d'ARN messenger de deux autres apoprotéines des chylomicrons (apo A I et apo A IV) étaient, quant à eux, réduits de 50 %. Des résultats préliminaires suggèrent une dégradation accrue de l'apo B

synthétisée par les sujets a- β -lipoprotéïnémiques. Ce travail, basé sur des méthodes immunologiques, devrait déboucher sur l'étude des bases moléculaires de l'a- β -lipoprotéïnémie, notamment l'étude structurale de la protéine et l'analyse de l'ARN messenger trouvés chez ces patients.

[1. Glickman RM, et al. *Gastroenterology* 1991 ; 101 : 749-55.]

■■■■ **Effet protecteur de l'apolipoprotéine AI contre l'athérogenèse.** Les études épidémiologiques ont bien mis en évidence la corrélation inverse entre le risque d'athérosclérose et les concentrations plasmatiques en lipoprotéines de haut poids moléculaires (HDL) et apolipoprotéine AI (ApoA-I), sans que pour autant une relation directe de cause à effet ait pu être démontrée. Pour tester cette hypothèse E. Rubin et al. (Berkeley, Université de Californie, CA, USA) ont comparé les effets de régimes athérogènes sur des souris sauvages et des souris transgéniques exprimant de hauts niveaux d'ApoA-I humaines. Ces régimes entraînent chez tous les animaux traités une augmentation du cholestérol plasmatique non lié aux HDL. En revanche, la concentration plasmatique du cholestérol lié aux HDL est deux fois plus élevée chez les animaux exprimant l'ApoA-I. Selon les régimes riches en graisses administrées, les auteurs observent une protection complète, c'est-à-dire l'absence de plaques pré-athéromateuses au niveau de l'aorte, ou l'apparition de plaques beaucoup plus petites chez les souris transgéniques que chez les témoins. Ces résultats suggèrent donc que des concentrations plasmatiques élevées d'apolipoprotéines AI ont un effet inhibiteur direct sur les étapes précoces de l'athérogenèse.

[1. Rubin ME, et al. *Nature* 1991 ; 353 : 265-7.]