

Tous parents, tous différents : la génétique des populations humaines

Autor(en): **Langaney, André / Cobut, Gérard**

Objektyp: **Article**

Zeitschrift: **Actes de la Société jurassienne d'émulation**

Band (Jahr): **99 (1996)**

PDF erstellt am: **28.06.2024**

Persistenter Link: <https://doi.org/10.5169/seals-684442>

Nutzungsbedingungen

Die ETH-Bibliothek ist Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Inhalten der Zeitschriften. Die Rechte liegen in der Regel bei den Herausgebern.

Die auf der Plattform e-periodica veröffentlichten Dokumente stehen für nicht-kommerzielle Zwecke in Lehre und Forschung sowie für die private Nutzung frei zur Verfügung. Einzelne Dateien oder Ausdrucke aus diesem Angebot können zusammen mit diesen Nutzungsbedingungen und den korrekten Herkunftsbezeichnungen weitergegeben werden.

Das Veröffentlichen von Bildern in Print- und Online-Publikationen ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Die systematische Speicherung von Teilen des elektronischen Angebots auf anderen Servern bedarf ebenfalls des schriftlichen Einverständnisses der Rechteinhaber.

Haftungsausschluss

Alle Angaben erfolgen ohne Gewähr für Vollständigkeit oder Richtigkeit. Es wird keine Haftung übernommen für Schäden durch die Verwendung von Informationen aus diesem Online-Angebot oder durch das Fehlen von Informationen. Dies gilt auch für Inhalte Dritter, die über dieses Angebot zugänglich sind.

Tous parents, tous différents

André Langaney

(Propos recueillis et transcrits par Gérard COBUT, IRSCNB)

INTRODUCTION

En 1990, des chercheurs font état d'une découverte frappante concernant la répartition de gènes organisateurs de l'anatomie chez la drosophile. Ils soulignent un parallèle entre la séquence de ces gènes et les territoires concernés dans l'organisme. Les premiers gènes de la séquence interviennent dans la partie antérieure de la drosophile, les suivants dans une zone plus postérieure, ensuite dans la partie médiane (thorax), enfin les derniers dans l'abdomen. Le monde scientifique, d'abord sceptique, se prend à rechercher le même type de gènes chez d'autres êtres. Il les trouve en effet, même chez des vertébrés. Ainsi chez l'embryon de souris, on rencontre la même organisation, d'avant en arrière.

Cette découverte amène une double révolution dans les concepts. Les invertébrés et les vertébrés partagent désormais des séquences semblables d'ADN (clusters de gènes) organisées à peu près de la même façon. Ensuite un dogme s'effondre, selon lequel il n'y avait aucune similitude entre la répartition des gènes sur le chromosome et celle de leurs effets sur l'être dont l'organisation est codée par ces gènes.

LE MATÉRIEL GÉNÉTIQUE HUMAIN : PAS TRÈS ORIGINAL

Si l'on trouve ces structures semblables, d'avant en arrière, chez la drosophile comme chez l'embryon de souris, on peut s'attendre à ce qu'elles soient présentes chez d'autres espèces, humains compris. Cette caractéristique commune remonterait obligatoirement aux ancêtres

communs des vertébrés et des invertébrés, il y a peut-être un milliard d'années. Dans cette perspective, le matériel génétique humain apparaît avant tout comme du matériel génétique d'être vivant, qui fait fonctionner des cellules. Il comporte bien des portions communes avec les végétaux ou avec les animaux : sa plus grande partie n'est pas originale au sein du monde vivant.

Rappelons d'ailleurs que l'embryon humain, au cours de son développement, passe par un stade qui présente des fentes branchiales et ressemble à un embryon de poisson ; il ressemble ensuite à n'importe quel embryon de mammifère, avant que le fœtus n'évoque celui de n'importe quel primate. Rien d'étonnant, au vu de tout ceci, qu'il y ait beaucoup de gènes en commun entre tous ces êtres...

Si d'aventure on s'intéresse aux chromosomes et qu'on compare ceux de l'Homme et du chimpanzé, on s'aperçoit que, chaque fois que le chimpanzé a un chromosome d'un certain type, l'Homme possède le même. La paire de chromosomes N° 2 de l'Homme fait exception : le chimpanzé a deux chromosomes qui lui correspondent. Or on retrouve sur le bras court du chromosome 2 de l'Homme le même système de bandes que sur un petit chromosome du chimpanzé. Il en est de même entre le bras long du chromosome 2 humain et un autre chromosome du chimpanzé. Cette homologie se retrouve, à quelques petits retournements près, sur l'ensemble de ces matériels génétiques vus sous le microscope. Elle existe également au sein de tous les primates. On retrouve d'ailleurs la même situation entre les carnivores ou parmi l'ensemble de tous les mammifères. En dernière analyse, il est donc possible de retracer l'histoire des chromosomes de tous les vertébrés.

Quand, enfin, on passe au niveau moléculaire et qu'on pratique l'hybridation d'ADN de chimpanzé et d'Homme, la proportion qui ne s'hybride pas est de l'ordre d'un pour mille seulement. La conclusion s'impose : notre matériel génétique n'a rien de très original.

D'OÙ VIENT LE GÉNOME HUMAIN ?

Le matériel génétique humain trouve son origine dans la nuit des temps. Depuis nos ancêtres communs avec les chimpanzés, il y a 5 à 7 millions d'années, il a traversé toute une lignée d'hominidés fossiles – sur lesquels toute information génétique fait défaut. Les premiers Hommes modernes remontent à 100 000 ou 120 000 ans. Ces Hommes modernes fossiles de Omo Kibish (Ethiopie), de Qafzeh (Palestine), de Djebel Irhoud (Maroc), ne présentent pas de différences morphologiques majeures avec l'Homme actuel. Ceci en dépit de reconstitutions parfois très diverses, même si elles sont basées sur des documents paléontologi-

ques authentiques. Ainsi l'Homme de Cro-Magnon a fait l'objet de reconstitutions dont l'apparence varie entre celle d'Arnold Schwarzenegger et d'une vieille sorcière, alors que l'on connaît un vieil homme et une adorable jeune femme de 18-19 ans.

Les hommes modernes ne datent donc que de 1000 siècles, ce qui est extrêmement peu : l'histoire de notre espèce est courte. Au plan de la génétique des populations, de nouveaux gènes apparus par mutation ont très peu de chances d'avoir atteint des fréquences élevées en aussi peu de temps, à moins qu'ils ne confèrent des avantages sélectifs considérables.

ÉTUDIER MILLE SIÈCLES D'HISTOIRE

On peut d'abord songer à étudier les gènes présents (ou absents) à travers toutes les populations du globe dans les systèmes génétiques classiques. Pour l'immense majorité d'entre eux, le répertoire des gènes connus dans les populations est à peu de chose près le même partout, si l'on prend des échantillons suffisamment grands. Or, lorsque les premiers généticiens des populations ont lancé des recherches en collaboration avec des centres de transfusion sanguine, ils cherchaient – selon les paradigmes de l'époque – les gènes des « noirs », des « jaunes », des « blancs »... De plus, ces enquêtes étaient faites sur de très petits échantillons : des donneurs de sang toulousains, quelques étudiants vietnamiens de la cité universitaire, l'un ou l'autre éboueur malien, etc. Dans ces conditions, les chercheurs ont effectivement décrit les « gènes des noirs, des jaunes et des blancs ». Plus tard, j'ai eu l'occasion de mener les premières enquêtes de génétique à grande échelle en Afrique, sur le terrain. Au Sénégal oriental, j'ai récolté 900 tubes de sang et non plus 20. « On » m'a d'abord expliqué que mon travail ne valait rien : « mes » noirs avaient plein de gènes jaunes et des tas de gènes blancs. J'ai répliqué par une boutade : « Je ne savais pas que les gènes avaient une couleur ! ». L'important est de se rendre compte qu'un petit échantillon « rate » statistiquement les gènes qui ont une fréquence faible, d'où l'importance d'une taille suffisante des échantillons.

Ce type d'étude a maintenant été fait un peu partout dans le monde et l'on s'est en effet aperçu que, pour la plupart des systèmes génétiques compliqués, le répertoire des gènes était le même partout. Par contre, ce sont leurs fréquences qui varient entre différentes populations : tel gène plutôt fréquent en Afrique centrale peut être rare en Asie du Sud-Est, etc. Une exception se marque toutefois : les populations « périphériques » ou celles qui trouvent leur origine dans un petit effectif ont

tendance à perdre des variantes génétiques par rapport aux grandes populations continentales.

PAS DE MARQUEURS ABSOLUS, DES DISTANCES

En génétique des populations, le principal problème – pour les systèmes que l'on connaît à ce jour – c'est l'absence de marqueurs génétiques absolus des populations. On a cherché en vain de ces gènes qui seraient à la fois présents chez tous les individus d'une population et absents de tous les individus de toutes les autres populations. Il existe évidemment des gènes rares, qui sont, lorsqu'on les rencontre, l'indice du fait que l'échantillon vient de tel coin du monde. Ainsi certaine variante du système GM n'est présente que chez les Aïnous. Mais ces gènes ne sont présents que chez une minorité de la population en question et ne constituent donc en rien des « marqueurs absolus ».

En résumé, pour comparer des populations entre elles, on ne peut comparer que des fréquences de gènes. Cette comparaison mène au calcul d'une « distance génétique ». Imaginons un exemple où deux populations, I et J, sont envisagées pour un système où il n'y a que trois gènes allèles, A, B, C. Les fréquences trouvées pour ces trois allèles sont données au tableau ci-dessous. Ce qui est commun entre ces deux populations, c'est bien la somme des fréquences partagées pour chacun des allèles.

Allèle	Population I	Population J	Fréquence partagée
A	0,56	0,52	0,52
B	0,22	0,36	0,22
C	0,22	0,12	0,12
Total :			0,86

Il y a donc dans ce cas 86 % de fréquences de gènes partagées entre les deux populations. Or, le complément à 100 de ce total possède les propriétés mathématiques de ce que l'on appelle une distance, qui peut donc servir à mesurer la « distance génétique » entre ces deux populations. On peut répéter ce calcul pour toute une série de populations et pour d'autres systèmes génétiques, puis en retirer diverses représentations graphiques. Celles-ci illustrent les distances génétiques entre les

populations considérées. Ces graphiques sont la base d'une série d'interprétations intéressantes mais, si l'on veut aller plus loin en matière de recherche, il faut évidemment dépasser ce stade des « beaux graphiques » pour se consacrer aux analyses et tests statistiques.

GÉNÉTIQUE ET GÉOGRAPHIE

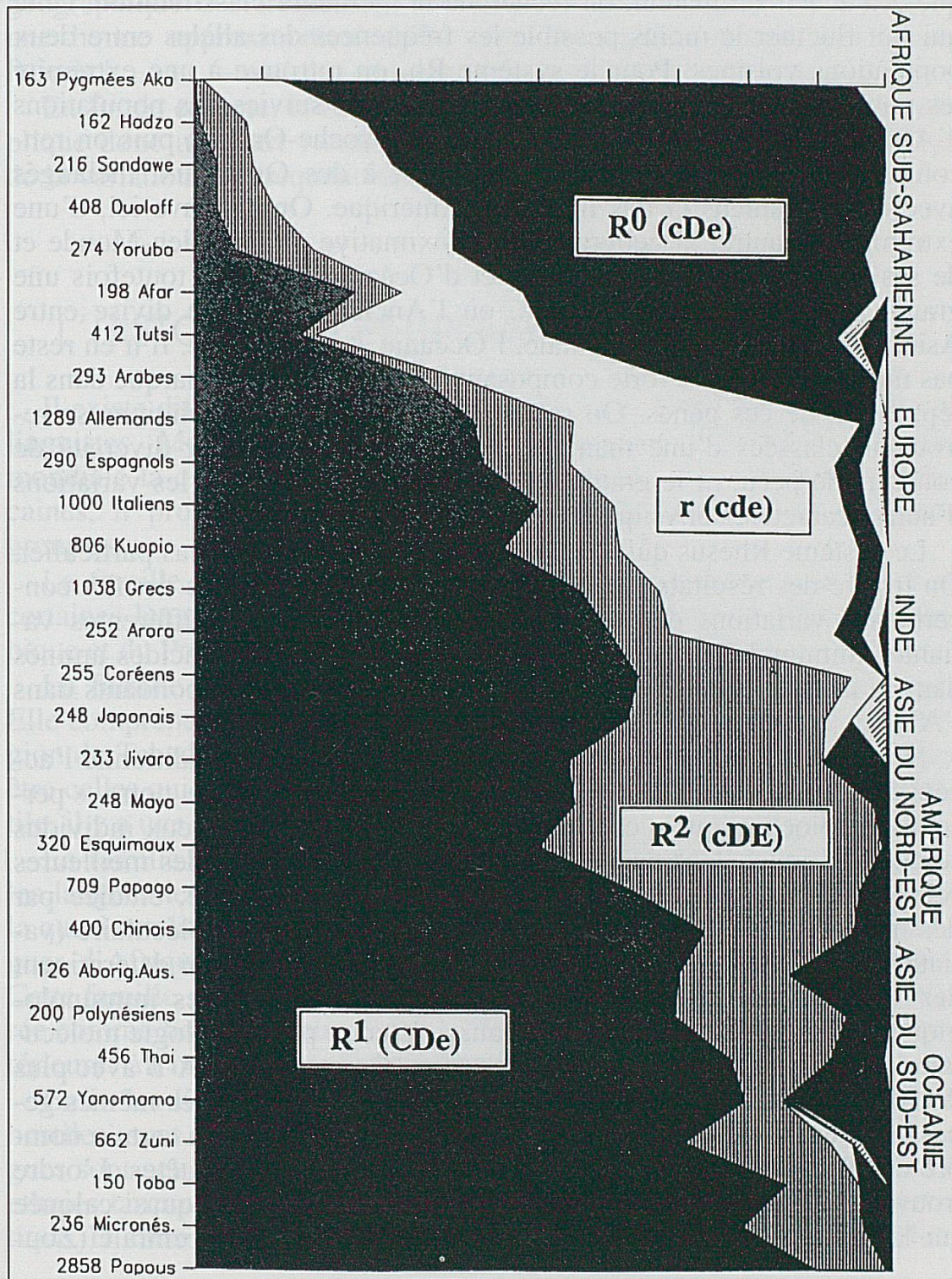


Fig. 1 : Représentation des distances génétiques entre 15 populations.

La figure 1 représente les fréquences et les distances génétiques entre des populations diverses du monde, pour le système Rhésus. Dans ce dernier – dans sa version simplifiée – chaque individu possède 2 allèles parmi 8 possibles. Chaque population est représentée par un segment de droite de longueur 1 ; la fréquence de chaque gène (R_0, r, R_2, R_1, \dots) est représentée pour chaque population par un segment de longueur correspondante. On place les populations voisines selon leur distance génétique 2 à 2. Par ordinateur, on recherche la meilleure classification, celle qui fait fluctuer le moins possible les fréquences des allèles entre deux populations voisines. Pour le système Rh, on retrouve à une extrémité les populations du sud et de l'est de l'Afrique, suivies des populations d'Afrique de l'Est, d'Afrique du Nord, du Proche-Orient ; puis on rencontre des Indo-Européens et l'on aboutit à des Orientaux, mélangés avec des Océaniens et des Indiens d'Amérique. On retrouve ici, d'une extrémité à l'autre, la géographie approximative de l'Ancien-Monde et de ses prolongements d'Amérique et d'Océanie. On note toutefois une grande confusion aux extrémités, où l'Ancien-Monde se divise entre Asie nord-orientale, sud-orientale, l'Océanie et l'Amérique. Il n'en reste pas moins qu'une très forte composante géographique se marque dans la répartition de ces gènes. On remarque aussi que les populations se retrouvent classées d'une manière qui ne respecte pas leur diversité de couleurs de peau, ni le gradient de celles-ci, pas plus que les variations d'autres caractères physiques.

Le système Rhésus qui vient d'être étudié n'est pas un cas particulier. On trouve des résultats similaires pour le système AG. Ce dernier concerne les variations de l'apolipoprotéine ; il comporte différentes variantes immunologiques dont on connaît la séquence des acides aminés dans la protéine, tout comme celle des nucléotides correspondants dans l'ADN.

Nous pouvons aussi envisager le système HLA, qui conditionne l'acceptation ou le rejet des greffes d'organes. Il concrétise notre « personnalité biochimique » car la probabilité est infime que deux individus soient identiques dans ce système. Il est aujourd'hui une des meilleures sources d'information sur la diversité génétique humaine, étudiée par des méthodes à la fois immunologiques et de biologie moléculaire (variations de l'ADN). Il comporte 3 locus conventionnels qui déterminent des variantes immunologiques ; il y a aussi des variantes immunologiques dites de classe 2 qu'on peut aussi détecter par la biologie moléculaire. Le locus le plus complexe dans ce système est le HLA-B avec plus de 60 variantes connues. On y retrouve pourtant souvent les mêmes gènes à travers toutes les populations du monde – même si un certain nombre d'allèles rares peuvent être « ratés » au cours des enquêtes. L'ordre trouvé par l'ordinateur montre à nouveau une succession quasi calquée sur la géographie : à une extrémité, l'Afrique australe et centrale (Zou-

lous, Bantous, Pygmées), puis l'Afrique de l'Est et l'Afrique du Nord. Suivent les Indo-Européens, puis des Orientaux (Vietnamiens, Coréens), quelques populations à la classification douteuse et enfin un mélange d'Indiens d'Amérique et d'Océaniens.

En élargissant l'étude à la combinaison des fréquences des gènes des locus HLA-A et HLA-B, l'ordinateur fournit – toujours en utilisant exclusivement les données génétiques – le même type de succession similitudino-géographique des fréquences des gènes et permet même de retrouver une bonne approximation de la répartition géographique des populations !

En fin de compte, la distance géographique entre des populations explique donc une part variable mais toujours importante et significative de la variation des fréquences de gènes entre elles.

L'HISTOIRE DU PEUPEMENT DE L'AFRIQUE

Il existe différentes versions de cette histoire. Voici d'abord celle des linguistes. Merritt Ruhlen, linguiste californien, a étudié à l'échelle mondiale la classification des langues humaines. Pour les langues africaines, il propose avec J. Greenberg une classification en 4 grands groupes.

La famille afro-asiatique englobe les langues d'Afrique du Nord et certaines langues d'Afrique de l'Est, plus d'autres de la péninsule arabe et du Proche-Orient (notamment les langues sémites).

La famille nilo-saharienne pourrait être dérivée de l'afro-asiatique. Elle comprend des langues est-africaines, plus d'autres langues de la région du Tchad et de la boucle du Niger. Ces populations sont supposées être celles qui occupaient la région saharienne au temps où elle était fertile, il y a une dizaine de milliers d'années.

La famille Niger-Kordofanien, dite aussi « Niger-Congo », renferme les langues les plus nombreuses de l'Afrique subsaharienne. Elle réunit la quasi-totalité des langues ouest-africaines et une bonne partie des langues d'Afrique centrale et du sud (en particulier les langues bantoues). Cette famille, beaucoup plus homogène que les précédentes, est considérée comme beaucoup plus jeune : elle correspond à des expansions postérieures à l'invention de l'agriculture.

Enfin les Khoisan, très peu nombreux, refoulés dans des régions très hostiles, sont éleveurs (Khoi) ou chasseurs-cueilleurs (San). Leurs langues sont peu nombreuses. Leurs seules affinités connues le sont avec 2 ethnies tanzaniennes, fort différentes à tous égards sauf au plan linguistique. D'autre part, il existe, dans certaines régions de l'Afrique de l'Est,

des restes de peintures rupestres et d'industries lithiques semblables à celles attribuées aux Khoisans.

Les linguistes ont pu suggérer des anciennetés pour ces familles de langues, en fonction notamment de la diversification des différentes langues en leur sein et de liaisons entre le vocabulaire et le calendrier archéologique marqué par des inventions techniques datées (métallurgie, agriculture,...) :

Famille de langues	Afro-asiatique	Nilo-saharienne	Niger-Congo	Khoisan
ordre de grandeur de l'âge estimé (ans)	15 000 ou plus	12 000	6000-8000	?

Certains biochimistes et biologistes moléculaires attribuent pourtant aux populations parlant les langues Niger-Congo un statut de populations anciennes, de « patrimoine génétique ancien de l'humanité » – à partir d'études faites essentiellement sur des Afro-américains descendant d'immigrants d'Afrique occidentale et centrale. Comment arrivent-ils à ces conclusions ? Certainement à cause d'un mauvais usage de l'interdisciplinarité. En effet, certains fossiles jadis jugés très anciens ont été redatés et se sont révélés beaucoup plus récents, tandis que d'autres s'avéraient ne pas appartenir à l'*Homo sapiens sapiens*. On se trouve donc en face d'un préjugé – qui vient à la fois de l'anthropologie traditionnelle et de la paléontologie – qui veut à tout prix que les Africains d'Afrique subsaharienne soient les populations les plus anciennes. C'est ainsi que des chercheurs de l'Université de Californie à Berkeley, à partir d'une étude portant sur 147 placentas humains (étude de l'ADN mitochondrial), ont prétendu avoir démontré que l'origine des Hommes modernes était en Afrique subsaharienne : c'est la « théorie de l'Eve africaine ». L'Afrique noire était représentée, dans leurs échantillons, par des Afro-Américains et, depuis, leurs calculs se sont avérés erronés et leurs conclusions injustifiées.

On possède d'autre part une classification des populations d'Afrique noire basée sur des systèmes génétiques mieux échantillonnés, tels que le système GM. Dans ces représentations, les populations d'Afrique de l'Est se rassemblent avec les Khoisans (ces derniers forment une « branche » sur le buisson). Un autre groupe correspond à toutes les populations de la famille Niger-Congo (Bantous, Africains de l'ouest, Pygmées). Entre ces 2 groupes de populations, une différence est manifeste : certaines variantes géniques sont présentes en Afrique de l'Est et chez certains des Khoisans, mais elles ont disparu ou sont rares chez les

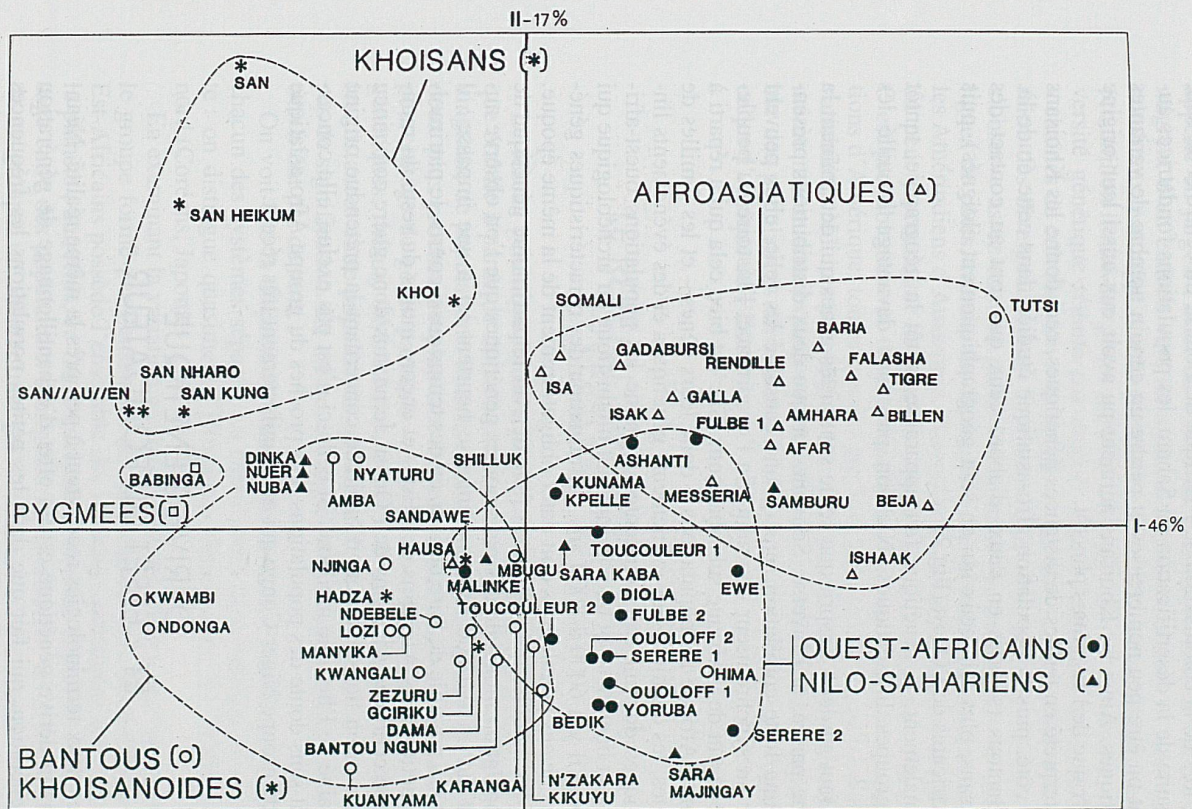


Fig. 2 : Distances génétiques et familles linguistiques en Afrique.

populations du groupe Niger-Congo. Ceci peut s'interpréter comme suit : lors de la colonisation de l'Afrique de l'ouest, du centre et du sud, à la suite de la désertification du Sahara, les populations fondatrices auraient pu être peu nombreuses et perdre un certain nombre de variantes génétiques. Ainsi les Khoisans auraient pu avoir eux aussi leur origine en Afrique de l'Est, mais plus tôt.

Interprété en termes de distances génétiques, ceci donne les Khoisans d'un côté, puis la population afro-asiatique étudiée dans cette étude-là, les Bantous (séparés en deux, d'abord ceux qui sont au contact des Khoisans, plus loin ceux qui en sont géographiquement éloignés), puis les Africains de l'Ouest (Fig.2).

Finalement, la génétique fait apparaître tantôt la géographie, tantôt des groupes linguistiques. N'a-t-on pas ainsi davantage brouillé les pistes ?

Il faut garder à l'esprit que ce ne sont ni les gènes qui déterminent la langue parlée, ni l'inverse. Souvent, lorsque deux distributions présentent une forte corrélation sans qu'une causalité les relie, elles peuvent avoir un synchroniseur commun. En l'occurrence, l'histoire du peuplement récent de l'Afrique subsaharienne. Si c'est bien cela qui a réparti à travers l'Afrique les populations – donc leurs gènes – et les familles de langues, les datations des événements génétiques et des événements linguistiques doivent être cohérentes. Comme les populations ouest-africaines et bantoues ont une origine tant linguistique qu'archéologique qui se situe il y a 6000 à 8000 ans, la formation des caractéristiques génétiques de ces populations doit dater obligatoirement de la même époque. Tous les modèles qui prétendent faire dériver le patrimoine génétique de la population mondiale des patrimoines génétiques que l'on observe aujourd'hui en Afrique subsaharienne se heurtent ici à une impasse : il n'est pas possible de faire dériver, en des temps aussi courts, le patrimoine génétique des Africains de l'Est – et encore moins du reste du monde – de celui des Africains de l'Ouest. Je ne suis donc guère convaincu par ce qu'on nous présente d'habitude concernant la prétendue origine africaine de l'homme moderne. Si celle-ci n'est pas exclue, elle concernerait sans doute des populations plus proches du groupe Afro-asiatique que du groupe Niger-Congo qui est manifestement très récent.

DÉRIVE GÉNÉTIQUE ET « EFFET DU FONDATEUR »

Les deux terminologies recouvrent à peu près la même réalité. Néanmoins la dérive génétique est un effet d'échantillonnage de génération en génération, qui fait que dans les petites populations les fréquences

des gènes varient au hasard, de génération en génération. Ce que l'on appelle en jargon l'effet du fondateur se passe quand des populations migrent d'un endroit vers un autre. Les émigrants sont en général peu nombreux par rapport à la population d'origine et emmènent donc un échantillon limité des gènes présents chez elle. Ils ont en général une diversité génétique réduite par rapport à leur population-mère. Lors de « chaînes de colonisation » de proche en proche, on peut ainsi voir apparaître des divergences dans le répertoire de gènes des populations. Cela s'observe assez souvent parmi les populations disséminées le long des lignes de côtes, comme les Esquimaux, ou le long des rivières, comme les Amérindiens d'Amazonie. Pour peu que les population soient petites, on peut y rencontrer de grandes variations de la fréquence des gènes. C'est probablement ce type de mécanisme qui explique que les populations d'Amérique, toutes passées par le détroit de Béring, présentent des variabilités réduites par rapport aux populations d'Asie continentale. Le même mécanisme explique probablement la variation relativement faible des populations africaines de l'ouest ou des Khoisan par rapport à celles d'Afrique de l'Est.

LE RÉSEAU GÉNÉTIQUE HUMAIN ET SES IMPLICATIONS

Le réseau génétique humain (Fig. 3) est une représentation synthétique établie à partir de données de distances génétiques portant sur une vingtaine de populations. Pour le construire, on relie chacune d'elles aux deux plus proches au plan des distances génétiques – c'est-à-dire du pourcentage de gènes partagés – dans un ensemble de quatre systèmes génétiques, rhésus, GM, HLA-A et HLA-B. Dans cette façon de faire, les distances entre populations qui ne sont pas reliées par un trait n'ont pas de signification. De même, les angles obtenus dans cette représentation n'ont pas de sens.

On voit que la disposition obtenue est similaire à celles fournies par chacun des systèmes séparés. La composante géographique est évidente : on distingue quasiment l'Himalaya, dressé entre les Orientaux du nord (Coréens, Japonais...) et les Indiens !

En examinant la variation génétique dans ce réseau, on s'aperçoit que le groupe formé par l'ensemble des Indo-Européens, Nord-Africains et Est-Africains possède l'ensemble des variantes génétiques présentes sur la planète ; autrefois, ce fait s'interprétait comme la preuve d'un métissage entre des populations d'origines indépendantes. Or, à la périphérie de ce groupe, on observe une perte de diversité par rapport au noyau

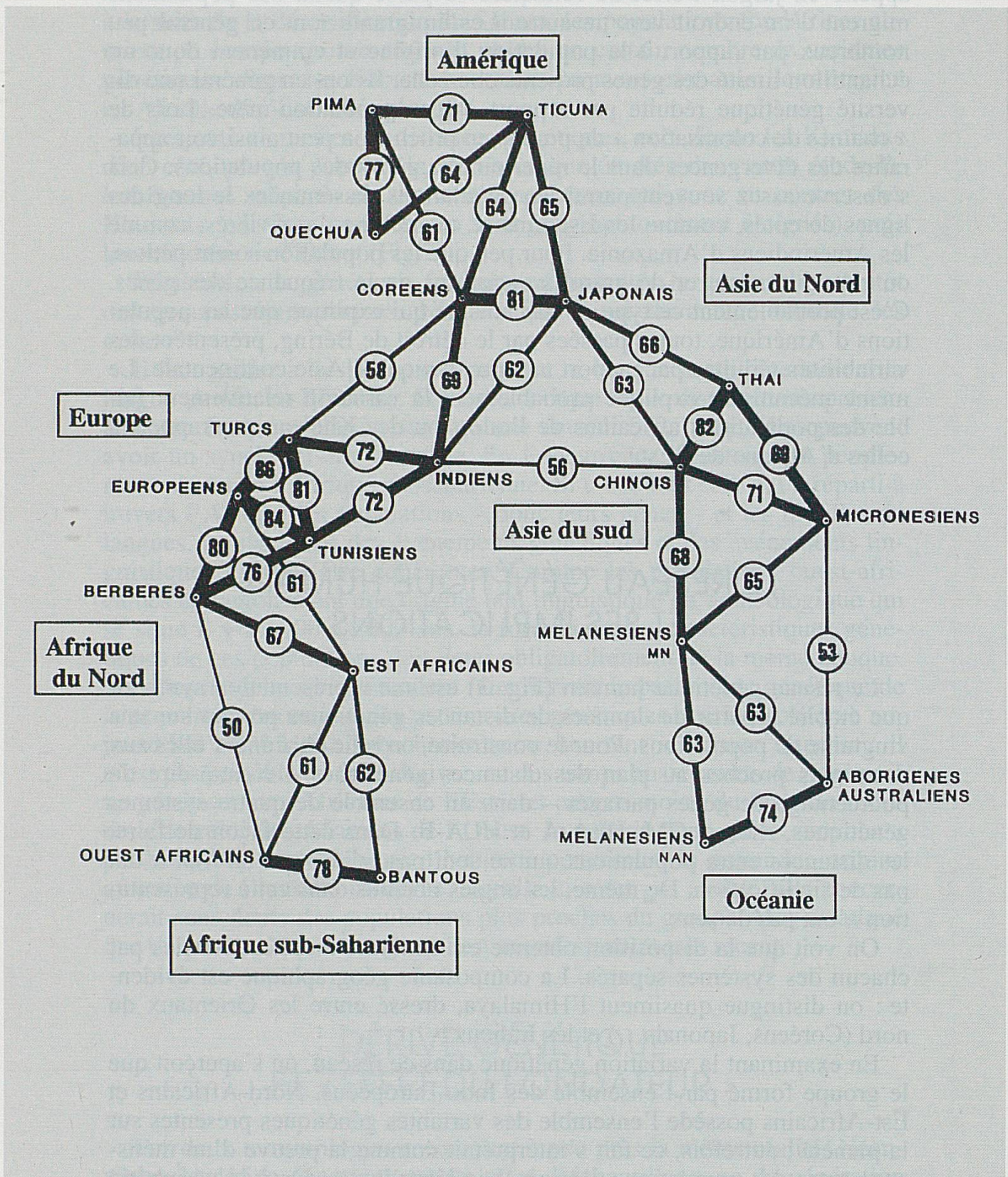


Fig. 3 : Le réseau génétique humain.

central. Compte tenu de l'ancienneté limitée de notre espèce – de l'ordre d'une centaine de milliers d'années pour expliquer toute la diversité génétique humaine – on peut penser qu'il y a un modèle alternatif au modèle « classique » impliquant une rencontre « au nœud du Pamir », avec métissage des populations. Cette alternative, c'est que les populations du noyau central représenteraient les descendants les plus directs de la dernière population fondatrice de l'ensemble des Hommes actuels (si population fondatrice unique il y avait). On expliquerait alors facilement la structure génétique de toutes les populations actuelles par effet du fondateur lors de migrations vers la périphérie, puis par dérive génétique sur place.

TROIS MODÈLES POUR UNE ÉVOLUTION

Comment trancher entre les deux mécanismes que nous venons d'envisager : soit une évolution indépendante depuis très longtemps, soit une population d'origine unique, une colonisation de la planète à partir d'un noyau central ? C'est une question difficile... Trois modèles principaux tentent d'interpréter l'histoire des Hommes modernes.

Le modèle polycentrique

Le premier modèle, polycentrique, estime que les populations humaines actuelles sont des groupes séparés depuis extrêmement longtemps et qui ont toujours évolué sur place. Il se base sur l'argument paléontologique – à mon avis relativement faible – de la continuité des fossiles depuis des périodes très anciennes (500 000 ans et plus, au temps des *Homo erectus*) jusqu'à la période actuelle. Cependant les tenants de ce modèle, avec à leur tête Wolpoff aux Etats-Unis, sont obligés de le compléter par des migrations entre groupes depuis assez longtemps, sans quoi il n'est pas réaliste par rapport aux faits génétiques. Or les généticiens des populations savent pertinemment que même des pourcentages très faibles de migration, c'est-à-dire de faibles échanges de gènes entre populations de la planète entière, auraient eu pour résultat que l'ensemble de l'humanité actuelle offre un pool génétique commun. Et des migrations ont toujours dû exister : n'importe qui (sauf s'il a une automobile !) est capable de marcher 50 km dans sa journée. S'il marche 300 jours dans l'année, cela fait 15 000 km par an ! Au vu des dimensions des continents, l'hypothèse de l'évolution sur place depuis 500 000 ans ne paraît pas très raisonnable... D'autant moins que les climats se sont

déplacés et que les flores et les faunes les ont suivis. Les hommes préhistoriques, qui furent longtemps des nomades suivant leur gibier et les végétaux dont ils dépendaient, ont donc obligatoirement bougé.

Le modèle de Cavalli-Sforza

Luca Cavalli-Sforza, de l'Université de Stanford, suggère une origine des *Homo sapiens sapiens* à partir d'*Homo erectus* il y a environ 100 000 ans (cf. observations de la paléontologie). Dès le départ, il voit une division entre les ancêtres des Africains et ceux des Orientaux. Puis les Indo-Européens seraient issus, il y a 30 000 ans, d'une fusion de populations entre le groupe africain et le groupe oriental. Les Océaniens se seraient séparés de ce dernier un certain temps avant. J'avoue n'avoir qu'un enthousiasme très limité envers ce genre d'hypothèses, parce qu'elles ne semblent viser, en dernière analyse, qu'à expliquer que les fréquences géniques des Indo-Européens, des Proche-Orientaux, des Indiens et des Africains de l'Est sont intermédiaires entre celles des gens d'Afrique subsaharienne et celles des Orientaux. Or ces fréquences géniques intermédiaires s'expliquent beaucoup plus simplement dans un modèle d'isolement génétique par la distance que par des métissages d'il y a 30 000 ans, que l'on ne peut pas démontrer.

Le modèle d'isolement génétique par la distance

Ce dernier modèle, tout aussi invérifiable que les précédents, se base sur le réseau génétique humain, lui-même résultat d'une histoire (Fig. 4). Il suppose aussi un groupe de départ unique, seule hypothèse vraisemblable au vu des données génétiques. Mais il cherche à concilier cette hypothèse avec la nécessité d'arriver au réseau actuel, tout en respectant les données de l'archéologie, de la paléontologie et de la linguistique sur les premiers peuplements connus ou les relations probables entre les populations ancestrales.

Où vivait le groupe de départ ? Nul ne le sait. Seuls les fossiles sont localisés géographiquement : en Afrique et au Proche-Orient, on en trouve qui ont 100 000 à 120 000 ans, puis on n'en trouve quasiment plus avant les premiers ancêtres sûrs des Africains actuels en Afrique subsaharienne, datés de 20 000 à 30 000 ans (peut-être 45 000 ans en Afrique du Nord et de l'Est) : il y a donc une discontinuité presque totale entre les périodes très anciennes et un passé relativement récent. On a donc très peu d'arguments pour affirmer que le peuplement actuel de l'Afrique remonte à très longtemps ; c'est plutôt le contraire qui est étayé provisoirement.

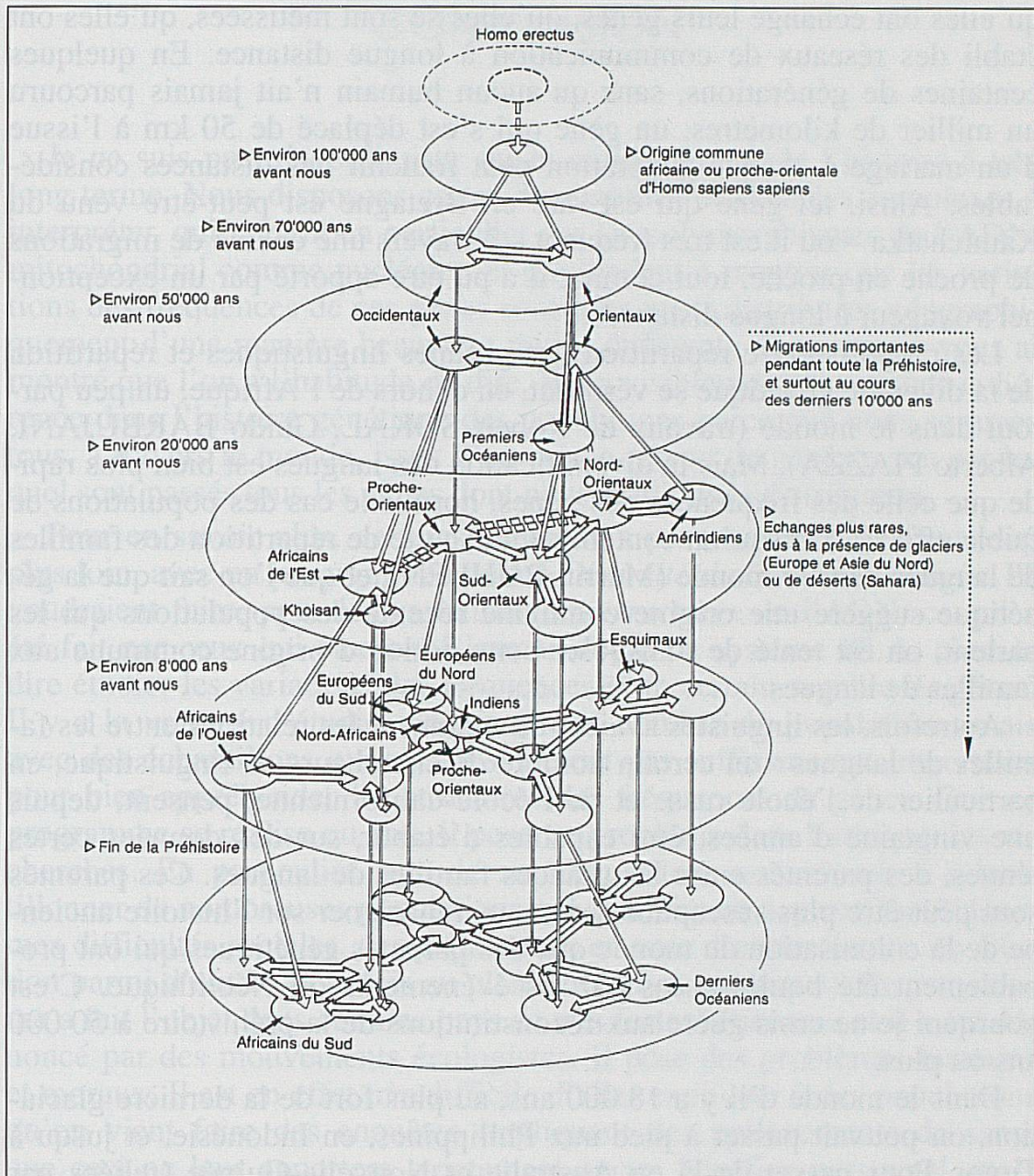


Fig. 4 : Histoire possible du réseau génétique humain.

LES GÈNES VOYAGEURS

Dans l'histoire de l'humanité, certains événements ont eu une importance considérable. Parmi eux, l'invention de l'agriculture, entre 15 000 ans et 9000 ans avant nous, pratiquement simultanée sur 5 continents. Dès ce moment, les densités humaines ont été multipliées par 20 à 30 selon les lieux, des populations sont entrées en contact. « Elles se sont fait la guerre et l'amour », c'est-à-dire qu'elles se sont affrontées et

qu'elles ont échangé leurs gènes, qu'elles se sont métissées, qu'elles ont établi des réseaux de communication à longue distance. En quelques centaines de générations, sans qu'aucun humain n'ait jamais parcouru un millier de kilomètres, un gène qui s'est déplacé de 50 km à l'issue d'un mariage à chaque génération peut franchir des distances considérables. Ainsi, tel gène qui est rare en Bretagne est peut-être venu du Kamtchatka – où il est très fréquent – à travers une chaîne de migrations de proche en proche, tout comme il a pu être apporté par un exceptionnel voyageur à longue distance...

Les rapports entre répartition des groupes linguistiques et répartition de la diversité génétique se vérifient, en dehors de l'Afrique, un peu partout dans le monde (travaux de Robert SOKAL, Guido BARBUJANI, Alberto PIAZZA). Mais la diversification des langues est bien plus rapide que celle des fréquences des gènes, hormis le cas des populations de faible effectif. Lorsqu'on contemple une carte de répartition des familles de langues dans le monde (Merritt RUHLEN), et que l'on sait que la génétique suggère une origine commune récente des populations qui les parlent, on est tenté de transposer cette notion d'origine commune aux familles de langues.

Autrefois, les linguistes avaient un *tabou* sur les relations entre les familles de langues : un certain nombre de chercheurs en linguistique, en particulier de l'école russe et de l'école californienne, pensent, depuis une vingtaine d'années, être capables d'établir, sur des données certes ténues, des parentés entre les grandes familles de langues. Ces parentés sont peut-être plus susceptibles de nous renseigner sur l'histoire ancienne de la colonisation du monde que des parentés génétiques qui ont probablement été bouleversées par les événements du Néolithique. C'est pourquoi je ne crois guère aux reconstitutions de la préhistoire à 50 000 ans ou plus.

Dans le monde d'il y a 18 000 ans, au plus fort de la dernière glaciation, on pouvait passer à pied aux Philippines, en Indonésie, et jusqu'à Timor. Pour passer de là en Australie et Nouvelle-Guinée (reliées par un pont continental), il fallait franchir un bras de mer qui se réduisait à 90 km aux périodes les plus favorables. Au nord, toute une partie du monde était inhabitable. En Afrique du Nord, de l'Ouest et au nord de l'Afrique de l'Est, s'étendait un immense désert – l'agriculture n'avait pas encore été inventée. Il y avait une bande assez étroite de forêt équatoriale, avec deux bandes de savane autour, dans laquelle des populations de chasseurs-cueilleurs de la forêt équatoriale ont certainement été piégées. Peu nombreuses, elles ont dû être le siège de fortes dérives génétiques, de forts effets du fondateur, qui ont dû modifier radicalement les fréquences des gènes de ces populations. Cela pourrait être l'origine de la singularité des fréquences génétiques observées chez certaines populations d'Afrique noire et en particulier dans le groupe Niger-Congo.

CONCLUSION

Je ne suis pas très optimiste sur la reconstitution de l'histoire à très long terme. Nous disposons certes de quantités de données complexes à interpréter, qui portent en particulier sur les polymorphismes de l'ADN, mitochondrial comme nucléaire, et qui tendent à montrer que les variations des fréquences de ces gènes sont elles aussi distribuées géographiquement d'une manière beaucoup moins différente de ce que je vous ai montré que l'on a l'habitude de dire. Il y a visiblement une grande cohérence dans l'histoire génétique des populations parce que nous sommes tous, à travers le monde, issus d'un même réseau de généalogies par lequel sont passés tous les gènes dont nous étudions la distribution.

Pour en savoir plus, certains chercheurs proposent d'aller beaucoup plus loin, avec un programme nommé *Diversité du Génome Humain*. Ils voudraient faire, pour les variations des structures de l'ADN, ce qui a été fait par ceux qui étudient l'immunologie du système HLA, c'est-à-dire étudier les variations des fréquences d'un certain nombre de gènes. Il y a la volonté d'étudier un très grand nombre de populations, mais avec des échantillons qui ne me semblent pas suffisamment importants pour bien appréhender les structures des fréquences de gènes. Que ce programme se réalise ou non, il pose un problème dans l'éthique des recherches. En particulier, une de ses dimensions consisterait à échantillonner de nombreuses populations qui connaissent aujourd'hui d'énormes difficultés sociales, des populations en voie d'extinction ou de dilution parmi d'autres (à la fois au plan culturel et génétique). Ce programme a fait l'objet dans certains pays de très fortes réticences et il a été dénoncé par des mouvements écologistes. Il pose des problèmes humains et moraux. Il est en effet très difficile d'aller expliquer à des populations qu'on vient faire des enquêtes impliquant des prélèvements de sang, que ceci ne leur apportera strictement aucun bénéfice personnel, mais qu'il est dans l'intérêt de l'humanité de savoir quel est leur patrimoine génétique. Mon expérience au Sénégal oriental est lourde d'enseignements à ce sujet...

Ce type d'études se heurte aussi à un problème méthodologique, car les populations relictuelles – que certains voudraient analyser en priorité – ont un effectif réduit. Elles ont subi une dérive génétique importante par rapport à leurs origines et ont perdu une grande partie de leurs variantes génétiques de départ. On a en fait beaucoup moins de chances de trouver le répertoire originel des gènes présents dans les populations ancestrales de l'humanité en allant à la pêche dans les très petites populations qu'en allant, au contraire, là où l'on a fait très peu de recherches, c'est-à-dire dans les grandes populations continentales. Pensons à

l'Afrique de l'Est, qui a peut-être eu un rôle fondamental dans la colonisation du reste du monde : c'est une des régions du monde pour lesquelles on a le moins de données. L'immense groupe linguistique Niger-Congo, comptant des centaines de millions d'individus, est moins connu au plan génétique que les Pygmées ou les Khoisans, qui sont quelques dizaines de milliers.

Enfin, les usages commerciaux des données recueillies lors des enquêtes de génétique des populations ne sont plus une vue de l'esprit, mais un début de réalité qui suscite quelques réserves supplémentaires. En fait, il est assez difficile aujourd'hui de se lancer dans de grandes études de génétique des populations avec l'enthousiasme mais aussi l'inconscience des problèmes éthiques qu'avaient nos illustres prédécesseurs lors de leurs grandes enquêtes d'immunologie comparée.

Un tel projet comporte néanmoins un aspect très positif. Pour l'appréhender, il faut remarquer que les généalogies qui ont conduit les gènes que nous avons étudiés là où ils sont maintenant ont aussi convoyé des gènes détériorants (ceux de maladies génétiques, généralement rares) et des facteurs de risque génétique aux grandes pathologies (diabète, maladies cardio-vasculaires, cancers, etc.), que l'on est en train de découvrir dans les laboratoires. Si les raisons fondamentales d'aller faire des grandes enquêtes aux quatre coins du monde ne sont pas toujours faciles à justifier, nous allons par contre vers des années où il y aura un intérêt épidémiologique considérable à faire à nouveau de grandes enquêtes de génétique des populations, pour mieux connaître la répartition des gènes détériorants à travers le monde et pour connaître la répartition de facteurs de risque génétique aux grandes pathologies. Il est sage de déjà penser à ces recherches, même si nous n'avons pas encore les moyens techniques de les faire directement. Or les études de génétique des populations à l'échelle mondiale, qu'il s'agisse des recherches sur les marqueurs « conventionnels » fournissant, avec le HLA, nos plus grosses banques de données génétiques actuelles, ou des enquêtes plus limitées sur la diversification des marqueurs fins ou des séquences d'ADN sont sur la seule piste actuelle pour l'exploration d'une épidémiologie à la fois génétique et géographique des grandes causes de mortalité.

André Langaney est directeur du Laboratoire de génétique et biométrie à l'Université de Genève et directeur du Laboratoire d'anthropologie biologique au Musée de l'Homme, Paris.

NOTES

¹ Conférence présentée le 15 mars 1995 à Porrentruy ; original publié dans *Probio-Revue*, 18 (1995), N° 3.

POUR EN SAVOIR PLUS

– L.L. Cavalli-Sforza, P. Menozzi et A. Piazza, 1994

HISTORY AND GEOGRAPHY OF THE HUMAN GENES (Princeton University Press)

– A. Langaney, 1988

LES HOMMES (Armand Colin)

– A. Langaney, N. Hubert van Blijenburgh et A. Sanchez-Mazas, 1992

TOUS PARENTS, TOUS DIFFERENTS (Réédition 1995, Laboratoire de Génétique et Biométrie, Université de Genève et Musée de l'Homme, Paris)

