

L'ADN mitochondrial, le chromosome Y et l'histoire des populations humaines

Lluís Quintana-Murci
Reiner Veitia
Silvana Santachiara-Benerecetti
Ken McElreavey
Marc Fellous
Thomas Bourgeron

De nombreuses disciplines étudient les origines de l'espèce humaine. En particulier, l'archéologie, l'anthropologie, la linguistique et la paléontologie tentent de retracer les grandes étapes de notre passé. Plus récemment, la génétique et la biologie moléculaire ont permis de préciser l'origine des différentes populations humaines en étudiant les variations génétiques entre les individus. Parmi tous les marqueurs génétiques utilisés, les plus intéressants sont les marqueurs uniparentaux comme l'ADN mitochondrial, hérité de la mère, et le chromosome Y, hérité du père car ils échappent tous deux à la recombinaison méiotique. Ainsi, contrairement aux autosomes ou au chromosome X, tous les ADN mitochondriaux présents dans les populations humaines peuvent provenir d'un ancêtre maternel commun et les chromosomes Y, d'un ancêtre paternel commun. Ces deux marqueurs ont donc été utilisés pour préciser géographiquement et historiquement l'origine de nos ancêtres communs les plus récents et les interactions entre les différentes populations humaines.

ADRESSES

L. Quintana-Murci : *docteur ès sciences, chercheur postdoctoral*. Laboratoire d'immunogénétique humaine, Inserm U. 276, Institut Pasteur, 25, rue du Docteur-Roux, 75724 Paris Cedex 15, France et Department of genetics and microbiology, University of Pavia, Via Abbiategrasso, 207, 27100 Pavia, Italie. R. Veitia : *docteur ès sciences, chercheur postdoctoral, Institut Pasteur, Universidad de la Habana, Cuba*. Inserm U. 276, Institut Pasteur. S. Santachiara-Benerecetti : *professeur à l'Université de Pavia, Italie*. Department of genetics and microbiology, University of Pavia. K. McElreavey : *docteur ès sciences, chargé de recherche à l'Institut Pasteur*. M. Fellous : *professeur à l'Université Paris 7*. Bourgeron : *maître de conférences à l'Université Paris 7*. Inserm U. 276, Institut Pasteur.

La question des origines, principalement celle de l'espèce humaine, sera probablement toujours un sujet suscitant polémiques et rebondissements. Depuis toujours, un grand nombre de réponses très différentes les unes des autres ont été proposées, parfois même imposées. Actuellement, de nombreuses disciplines scientifiques sont impliquées dans l'étude des origines de l'espèce humaine comme l'archéologie, la linguistique, la paléontologie et,

plus récemment, la génétique et la biologie moléculaire. Ces dernières disciplines ont donné des résultats très intéressants sur l'origine et l'âge de nos ancêtres communs comme sur l'histoire des différentes populations humaines. Malgré les très fortes critiques portant sur la découverte et l'utilisation des termes d'« Adam et Ève africains », il est maintenant difficile de vouloir étudier l'évolution de l'espèce humaine sans prendre en compte les données génétiques.

Variations génétiques et polymorphismes

Pendant de nombreuses années, les généticiens ont étudié les variations génétiques entre les individus et les populations dans le but de comprendre ces différences et les éventuelles interactions historiques des peuples [1]. Au début du siècle, les premières études ont porté, en particulier, sur les variations des groupes sanguins dans les populations mondiales. Cependant, le faible taux de variation des séquences protéiques ne donnait que des informations limitées sur le génome des différents individus. Ce n'est qu'au début des années 1980, grâce à la biologie moléculaire, qu'il a été possible d'identifier un grand nombre de variations génétiques entre les individus et les populations. En effet, l'ADN que nous avons hérité de nos ancêtres a accumulé des mutations différentes au cours de son évolution. Par conséquent, une séquence d'ADN peut être différente d'un individu à l'autre, comme sa fréquence peut l'être d'une population

à l'autre. L'étude de ces différences génétiques, connues sous le nom de « polymorphismes », peut apporter des informations importantes sur les relations de parenté entre individus et sur l'histoire des différentes populations.

C'est l'analyse des marqueurs polymorphes qui permet d'étudier les différences génétiques entre les individus ou les populations [1]. En effet, quand il existe plus d'un « allèle » d'un gène, on peut calculer les proportions respectives dans les populations et obtenir ainsi une représentation géographique des différentes variations génétiques. Enfin, si l'on étudie plusieurs gènes distincts, il est possible d'associer à un individu un ensemble de polymorphismes sur le même chromosome que l'on nomme « haplotype ». L'étude de la diversité génétique des individus permet donc de regrouper les haplotypes les plus voisins (« haplogroupes ») afin d'identifier les haplotypes ancestraux, de préciser les relations entre les différentes populations et de dater les différentes branches de l'arbre phylogénétique. En effet, si

l'on peut estimer le taux de mutation de la région du génome étudiée, il est possible de déterminer une date pour les différents embranchements de l'arbre et ainsi dater « la séparation » entre deux populations [1].

Les marqueurs génétiques uniparentaux : l'ADNmt et le chromosome Y

Parmi les différents marqueurs génétiques disponibles actuellement, les plus intéressants sont les séquences d'ADN héritées d'un seul des parents : l'ADN mitochondrial (ADNmt), hérité de la mère, et le chromosome Y, hérité du père.

L'ADNmt

Les mitochondries sont présentes en nombre variable dans le cytoplasme des cellules (de 75 dans le spermatozoïde à 100 000 dans l'ovocyte) [2]. La fonction principale de cet organe est de produire de l'énergie sous la forme d'ATP à partir de substrats carbonés et en présence d'oxy-

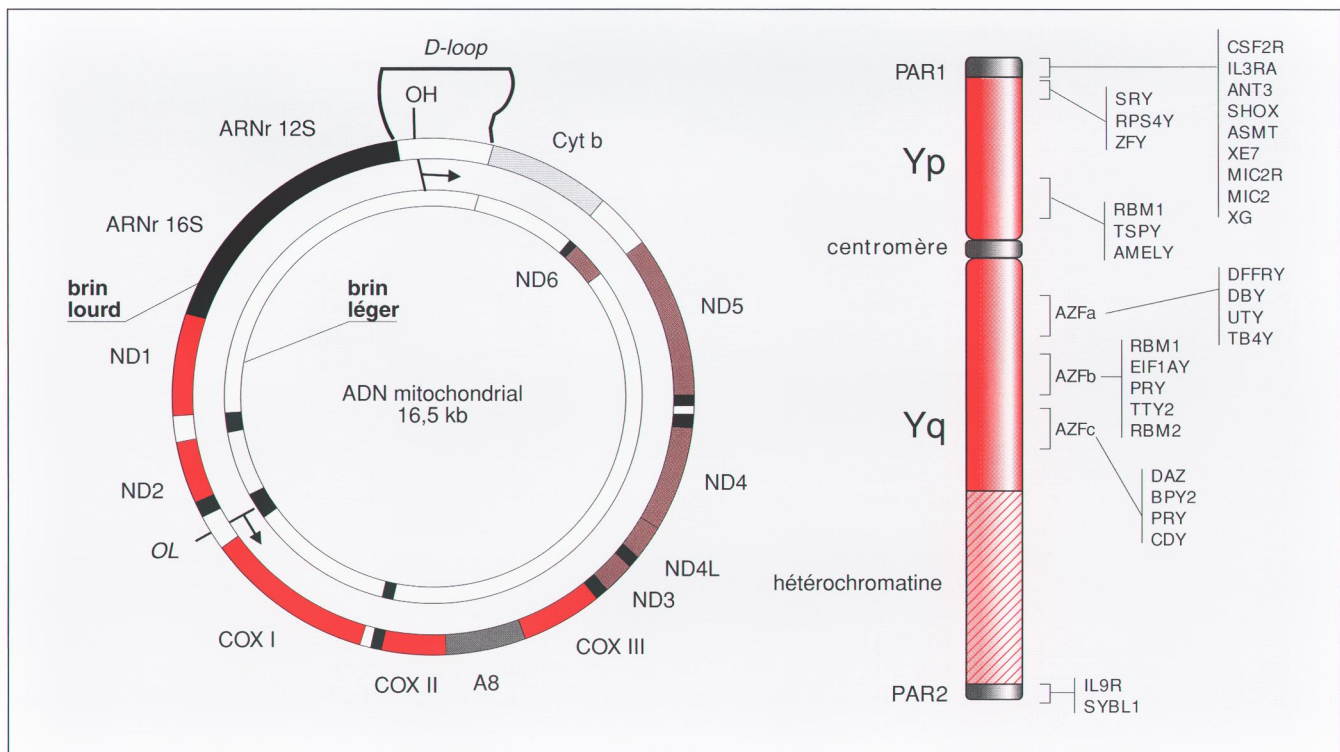


Figure 1. L'ADN mitochondrial (ADNmt) et le chromosome Y humain. A. L'ADNmt. Cytb : cytochrome b ; ND : NADH déshydrogénase ; COX : cytochrome c oxydase ; A : ATP synthétase. B. Le chromosome Y. PAR : région pseudo-autosomique ; AZF : azoospermia factor.

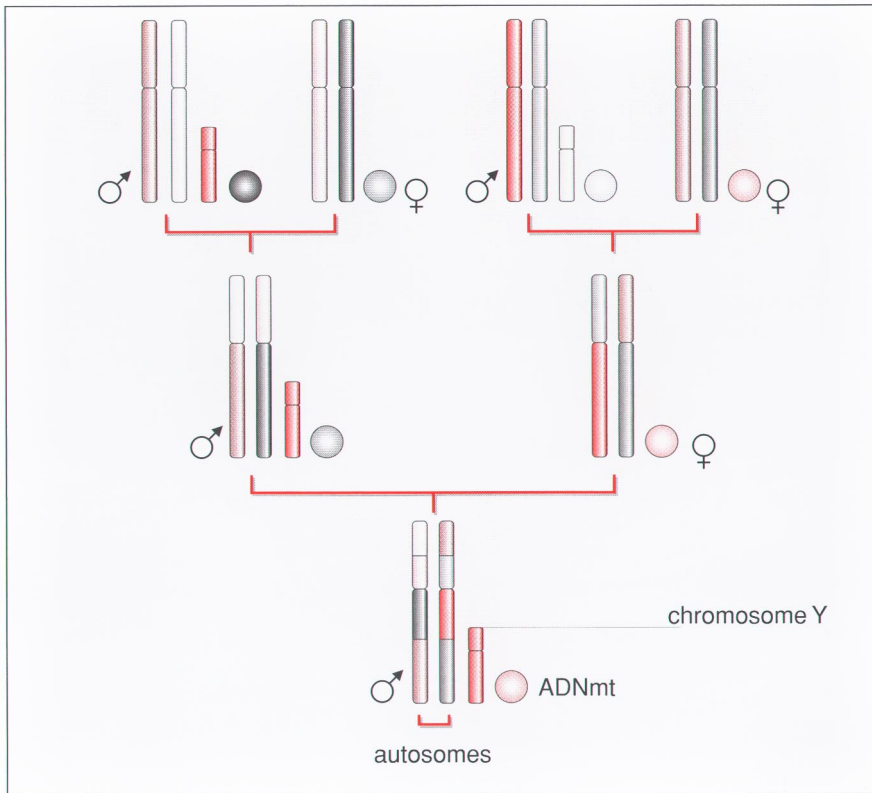


Figure 2. **Ségrégation des autosomes, du chromosome Y et de l'ADNmt au cours des générations.** Pendant la gamétogenèse, les séquences d'ADN portées par les autosomes subissent la recombinaison méiotique. Ainsi, chacun des autosomes porte des séquences d'ADN provenant de différents ancêtres. Au contraire, l'ADNmt et le chromosome Y ne subissent pas la recombinaison méiotique. Ils sont transmis sans modification d'une génération à l'autre. Les seuls changements qui peuvent intervenir sont de nouvelles mutations dans les cellules germinales.

gène. Elles possèdent plusieurs copies d'ADNmt dont la séquence est connue depuis 1981 (figure 1A). L'ADNmt mute environ 10 fois plus que les gènes nucléaires et cette particularité permet d'observer une quantité plus importante de variations de séquences polymorphes chez les individus. Dans l'espèce humaine, c'est uniquement la mère qui donne ses mitochondries à ses enfants (figure 2). Les mitochondries du spermatozoïde rentrent dans l'ovocyte mais sont très rapidement dégradées et ne participent donc pas au patrimoine génétique de l'embryon. Comme nous le verrons plus loin, cette hérédité maternelle, particulière aux mitochondries, a permis de retracer historiquement et géographiquement l'évolution de l'espèce humaine du côté de la femme [3].

Le chromosome Y

Le chromosome Y est présent uniquement chez les mâles. Il possède au niveau des télomères deux régions homologues du chromosome X (région pseudo-autosomique), et une grande partie centrale non recombinante qui lui est spécifique (figure 1B) [4]. La grande majorité des gènes portés par la région spécifique du chromosome Y possède une fonction strictement masculine. En particulier, le gène *SRY* (*sex region Y chromosome*) est un des gènes du déterminisme sexuel, engagé dans la formation du testicule [5]. Il existe aussi plusieurs gènes impliqués dans la spermatogenèse et certaines délétions du chromosome Y peuvent entraîner une infertilité masculine. Le reste du chromosome Y semble être constitué de séquences répétées,

de pseudogènes et d'une grande partie d'hétérochromatine polymorphe ne contenant vraisemblablement pas de gènes. Le chromosome Y est transmis uniquement de père en fils. De nouveau, c'est cette particularité qui a été utilisée pour étudier l'origine des populations humaines, cette fois-ci, du côté des hommes [6, 7].

L'utilité de ces marqueurs génétiques hérités d'un seul des parents tient au fait que les séquences d'ADN portées par ces chromosomes échappent à la recombinaison méiotique. En effet, l'ADN mitochondrial (ADNmt) et la région spécifique du chromosome Y ne recombinent pas lors de la méiose et, par conséquent, sont transmis sans modifications à la génération suivante. Les seules modifications qui apparaissent proviennent de nouvelles mutations touchant les cellules germinales. A l'inverse, les séquences d'ADN portées par le chromosome X ou les autosomes peuvent être le résultat d'un mélange de séquences de plusieurs ancêtres au moment de la recombinaison méiotique (figure 2). Pour simplifier, on peut dire que tous les ADNmt présents dans les populations humaines pourraient provenir d'un ancêtre maternel commun, et les chromosomes Y, d'un ancêtre paternel commun.

Les modèles d'évolution de l'espèce humaine

Depuis toujours, un point très controversé de l'évolution de l'espèce humaine porte sur la relation entre les premiers hommes modernes et les plus anciens hominidés [8]. Nous présenterons dans cet article les deux hypothèses les plus défendues actuellement. Ces deux modèles de l'évolution de l'espèce humaine ont plusieurs points communs, provenant en grande partie des données paléontologiques (figure 3). L'espèce humaine a un ancêtre commun avec les chimpanzés il y a 5-6 millions d'années. Puis, en Afrique de l'Est et du Sud on trouve les Australopithèques qui vivaient il y a 4 ou 5 millions d'années et les plus récents il y a 1-1,5 million d'années. L'Australopithèque la plus connue est Lucy, découverte en 1974 à Hadar en Éthiopie, qui marchait debout contrairement aux singes. Plus proche de nous, plusieurs catégories

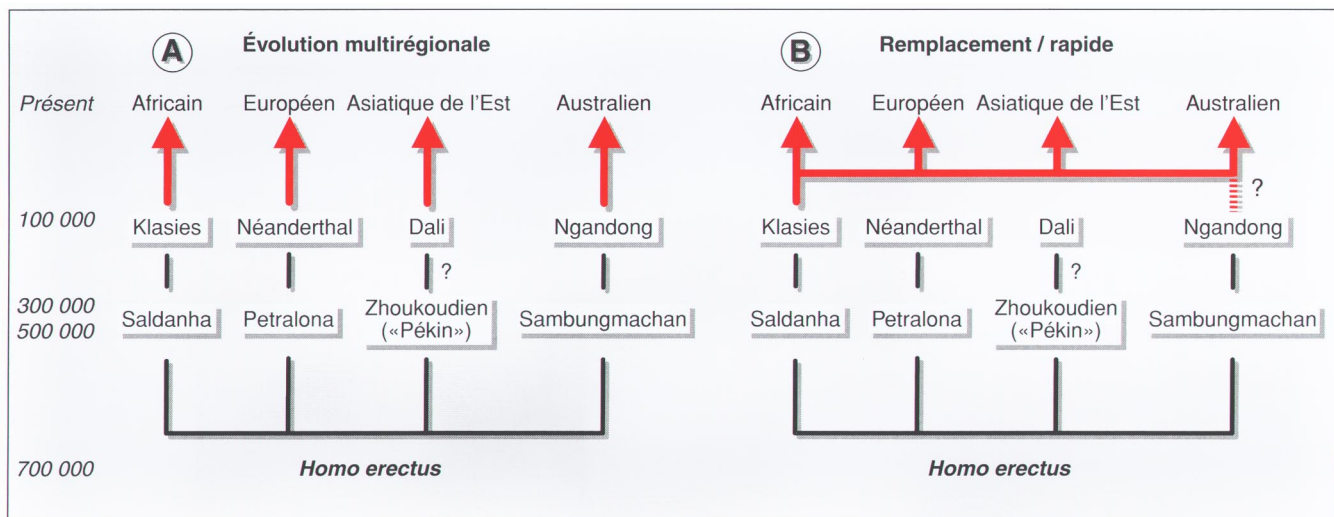


Figure 3. **Les modèles de l'évolution de l'espèce humaine : « évolution multirégionale » ou « remplacement rapide ».** **A.** Le modèle d'« évolution multirégionale » implique une évolution séparée et parallèle de l'espèce humaine à partir de plusieurs ancêtres *Homo erectus*. **B.** Le modèle de « remplacement rapide » considère que toutes les populations modernes dérivent d'un ancêtre commun récent (100 000-200 000 ans) d'origine africaine. Dans ce modèle, les autres lignées se sont éteintes sans laisser de descendants. Les noms entourés font références aux lieux où ont été trouvés les différents fossiles.

de fossiles humains ont été découvertes : *L'Homo habilis*, apparu en Afrique il y a 2 millions d'années, est probablement l'inventeur des premiers outils connus. *L'Homo erectus*, apparu il y a environ 1,8 million d'années, représente la première espèce de notre lignée trouvée en dehors de l'Afrique. Présent en Afrique de l'Est, du Sud et du Nord, il l'était aussi en Chine, à Java et au Proche-Orient. Les derniers fossiles d'*Homo erectus* datent de 150 000 à 300 000 ans. A l'*Homo erectus*, succèdent les *Homo sapiens* archaïques présents, entre -500 000 et -120 000 ans, en Europe et dans les mêmes régions qu'*Homo erectus*. Enfin, les *Homo sapiens neandertalensis* vivaient, entre -120 000 et -35 000 ans, dans une région plus restreinte qui couvre l'Europe et l'Asie du Sud-Ouest. Les premiers fossiles d'hommes anatomiquement modernes ou *Homo sapiens sapiens* sont datés de -100 000 ans. Les polémiques portent principalement sur les derniers chaînons et particulièrement sur les relations de parenté entre les hommes modernes et nos plus proches ancêtres : *Homo erectus*, *Homo sapiens* archaïques, *Homo sapiens neandertalensis*. La grande différence entre les deux modèles touche principalement l'identification de notre ancêtre com-

mun le plus récent. Le premier modèle « d'évolution multirégionale » propose que l'homme moderne aurait émergé de façon parallèle et indépendante de plusieurs populations d'*Homo erectus*, dans différentes régions du monde (figure 3). Les différentes populations actuelles auraient donc eu un ancêtre commun il y a environ 1 million d'années, probablement en Afrique. Viendraient ensuite 4 grandes lignées indépendantes à l'origine des hommes modernes en Indonésie, en Asie, en Afrique et en Europe (*Homo sapiens neandertalensis*) (figure 3). Plusieurs variantes de ce modèle existent faisant intervenir des croisements entre les différentes lignées pour obtenir les hommes modernes. Le deuxième modèle – de « remplacement rapide » ou *Out of Africa* – considère que notre ancêtre commun est beaucoup plus récent (100 000-200 000 ans) et originaire d'Afrique. Il y aurait donc eu une sortie d'Afrique subsaharienne vers le reste du monde et un remplacement des populations autochtones (figure 3). Ainsi, l'homme de Java, l'homme de Pékin et l'homme de Neandertal n'auraient plus de descendants actuellement. D'un point de vue génétique, l'ADN de ces individus et de ces populations ne fait pas

partie de l'ensemble des gènes actuels.

On peut alors se poser la question : comment la génétique peut-elle donner des arguments en faveur de l'un ou de l'autre modèle ? Deux grandes stratégies ont été utilisées : l'étude de l'ADN des fossiles et l'étude de la diversité génétique des différentes populations actuelles.

■ L'étude de l'ADN ancien

L'isolement de l'ADN des fossiles est particulièrement difficile mais, récemment, l'équipe de Svante Pääbo (Université de Munich, Allemagne), a réussi à isoler des fragments de molécules d'ADNmt à partir d'un humérus provenant d'un fossile d'*Homo sapiens neandertalensis* découvert en 1856 dans la vallée de Neander [9, 10]. Ces chercheurs ont choisi d'étudier l'ADNmt pour deux raisons : la première est l'abondance relative des molécules d'ADNmt (>1 000/cellules) par rapport au nombre de copies d'un gène nucléaire (2/cellules). La deuxième est le nombre important de résultats portant sur l'ADNmt des populations actuelles. Après avoir vérifié que l'os pouvait réellement contenir de l'ADN, ils ont montré que les séquences d'ADNmt obtenues

n'étaient ni une contamination par de l'ADNmt humain, ni des séquences possédant des erreurs introduites par la technique d'amplification de l'ADN. La comparaison de la séquence d'ADNmt d'*Homo sapiens neandertalensis* et des séquences humaines montre que la moyenne des différences entre les hommes modernes et *Homo sapiens neandertalensis* est de $27,2 \pm 2,2$. Or les différences observées entre les populations humaines actuelles est seulement de $8,0 \pm 3,1$. L'analyse phylogénétique incluant la séquence mitochondriale du chimpanzé place la séquence d'*Homo sapiens neandertalensis* très nettement en dehors des hommes modernes. En considérant la date de divergence entre l'homme et les chimpanzés, l'équipe de Svante Pääbo (Munich, Allemagne) date la divergence entre l'*Homo sapiens neandertalensis* et l'homme moderne entre 550 000 et 690 000 ans, époque plus ancienne que les dates mentionnées précédemment (250 000-300 000). La conclusion de ce travail tend donc à prouver que l'*Homo sapiens neandertalensis* n'a pas participé au patrimoine génétique des hommes modernes. Néanmoins, les conclusions de ce travail ont été extrêmement discutées et plusieurs critiques ont été faites. Une première porte sur le fait que les chercheurs n'ont séquencé qu'un seul individu. Une autre s'appuie sur le fait que seul l'ADNmt a été étudié et qu'il est possible que les gènes nucléaires d'*Homo sapiens neandertalensis* soient encore présents dans l'ensemble des gènes portés par les hommes modernes, en particulier en Europe. Malgré les critiques, ces résultats portant sur l'ADN isolé à partir de fossiles sont extrêmement encourageants et devraient permettre de mieux comprendre les relations phylogénétiques des différentes espèces présentes et/ou éteintes. Après l'étude de l'ADN des fossiles, la deuxième stratégie utilisée par les généticiens est de reconstituer le passé à partir des populations d'hommes modernes.

ADNmt et populations humaines

Les travaux portant sur les séquences mitochondriales des différentes populations d'hommes modernes

ont été principalement mis en œuvre par les chercheurs des groupes de Douglas Wallace à Atlanta (GA, USA) et d'Allan Wilson à l'Université de Berkeley (CA, USA) dans les années 1980 [11, 12]. Ils ont ainsi montré que tous les ADNmt actuels dériveraient d'une ancêtre commune vivant en Afrique, il y a environ 100 000 à 200 000 ans.

C'est à partir d'un polymorphisme de restriction Hpa I que l'équipe de Wallace a mis en évidence pour la première fois, qu'une variation de l'ADNmt était corrélée avec une ethnie et une origine géographique particulières [15]. L'étude a montré que plus de 90 % des Africains possédaient le site Hpa I (nt 3592) alors qu'aucun Asiatique ou Européen ne le possédait. Ensuite, l'utilisation d'autres polymorphismes de restriction a confirmé la relation entre la variation de l'ADNmt et l'origine géographique des individus examinés. Ces résultats montraient aussi que les Africains constituaient un groupe ethnique beaucoup plus variable que les autres groupes. Enfin, cette étude était en accord avec l'origine africaine récente des populations humaines car elle datait l'origine des lignées mitochondriales à 100 000 ans. Ce modèle de « remplacement rapide » a été conforté, un peu plus tard, par l'équipe de Wilson [14]. Ces chercheurs ont fait un travail important sur de très nombreuses populations et ils ont montré, d'une part, que c'était, de nouveau, en Afrique que l'on trouvait la plus grande diversité de séquences mitochondriales. D'autre part, l'arbre phylogénétique comportait 2 branches primaires distinctes : la première exclusivement africaine, la seconde possédant des séquences africaines et toutes celles des autres populations. En utilisant, un taux de mutation de 2-4 % par million d'années, les chercheurs ont pu calculer que la séquence ancestrale datait de 200 000 ans. L'existence d'une « Ève africaine » récente a aussi été avancée sur la base des études portant sur la région non codante de l'ADNmt (D-LOOP) qui est 3-4 fois plus polymorphe que les régions codantes [16]. Ces études ont confirmé la plus grande diversité des séquences africaines et l'âge de la première branche entre 166 000 et 249 000 ans. En conclusion, toutes ces études confortaient le modèle de « remplacement

rapide » avançant que l'homme est apparu en Afrique, il y a 100 000-200 000 ans. Par la suite, des populations sont sorties de ce continent pour coloniser de nouvelles terres.

De nombreuses polémiques ont porté sur les méthodes employées pour la construction de l'arbre phylogénétique (échantillonnage, qualité des données, estimation de l'horloge moléculaire, nombre d'arbres...) et sur le terme d'« Ève africaine » trop restrictif. Mais, même s'il est difficile de situer convenablement la racine de l'arbre phylogénétique, il semble maintenant que la plus grande partie des données génétiques va dans le sens d'une origine récente en Afrique des populations modernes.

Alors qu'au début de ces études, les techniques utilisées ne permettaient qu'une analyse restreinte des variations de l'ADNmt, il existe maintenant des techniques plus fines pour tenter de répondre aux questions anthropologiques quant à l'âge et à l'origine des différentes populations actuelles. Dans le but de mieux caractériser les populations humaines, une nouvelle méthode d'analyse de l'ADNmt a été utilisée [15]. L'ADNmt est amplifié en 11 fragments chevauchants, digérés par 14 enzymes de restriction différentes. Cette analyse a permis d'étudier plus de 20 % de l'ADNmt permettant de définir différents haplotypes mitochondriaux. Ainsi, les ADNmt possédant des haplotypes similaires sont regroupés en « haplogroupes ». L'étude a porté sur 4 groupes humains : les Africains, les Européens, les Asiatiques et les Indiens d'Amérique.

L'Afrique

Les résultats de cette étude montrent une fois de plus que la population africaine est la plus divergente. Sur 70 haplotypes différents retrouvés dans les populations africaines, 55 appartiennent à l'haplogroupe L « spécifiquement africain », défini aussi par la présence du site Hpa I 3592 cité plus haut. Dans cet haplogroupe, deux sous-groupes, L1 et L2, représentent respectivement 34 % et 42 % de la population africaine. Si les populations d'Afrique de l'Ouest sont dispersées dans les deux sous-groupes, les Pygmées de l'Est et les Pygmées de

l'Ouest montrent une distribution asymétrique entre ces deux groupes. En effet, les Pygmées d'Afrique de l'Est (ex-Zaïre) sont regroupés à 65 % dans une seule sous-lignée de l'haplogroupe L1, alors que les pygmées d'Afrique de l'Ouest (République de Centrafrique), sont regroupés à 54 % dans une seule sous-lignée de l'haplogroupe L2. Un autre groupe d'ADNmt est défini par la perte d'un site Dde I (nt 10394) et représente environ 4 % des populations africaines. Cet haplotype est particulièrement intéressant car il pouvait être à l'origine des ADNmt de plus de la moitié des Européens, et de certaines populations d'Asie et d'Amérique. Enfin, en utilisant un taux de mutation de l'ADNmt de 2-3 %/million d'années, la date d'origine de l'haplogroupe L se situerait entre 98 000 et 130 000 ans et la lignée des ADNmt africains aurait entre 101 000 et 131 000 ans [11, 14].

L'Europe

L'analyse phylogénétique indique que les Européens sont divisés en deux sous-groupes par la présence (25 %) ou l'absence (75 %) du site Dde I (nt 10394). Ainsi, on observe une augmentation très importante de la perte de ce site uniquement retrouvée chez 4 % des Africains. Afin de préciser les variations d'ADNmt présentes en Europe, la même étude a distingué 4 haplogroupes spécifiquement européens H, I, J et K. L'haplogroupe H qui n'a pas le site Dde I (nt 10394) représente 40 % des Européens et les haplogroupes I, J et K possédant le site Dde I en représentent chacun 2-15 %. Les âges approximatifs de ces haplogroupes ont été calculés : l'haplogroupe H serait apparu il y a 31 000-41 000 ans et les haplogroupes I et K, il y a 13 000-19 000 ans. Enfin, les divergences de séquences présentes en Europe indiquent une date de colonisation entre 30 000 et 51 000 ans (figure 4) [11-16].

L'Asie

Les variations d'ADNmt des populations asiatiques peuvent être divisées en 2 grands groupes définis eux aussi par la présence ou l'absence du site Dde I (10394). En outre, un nouveau

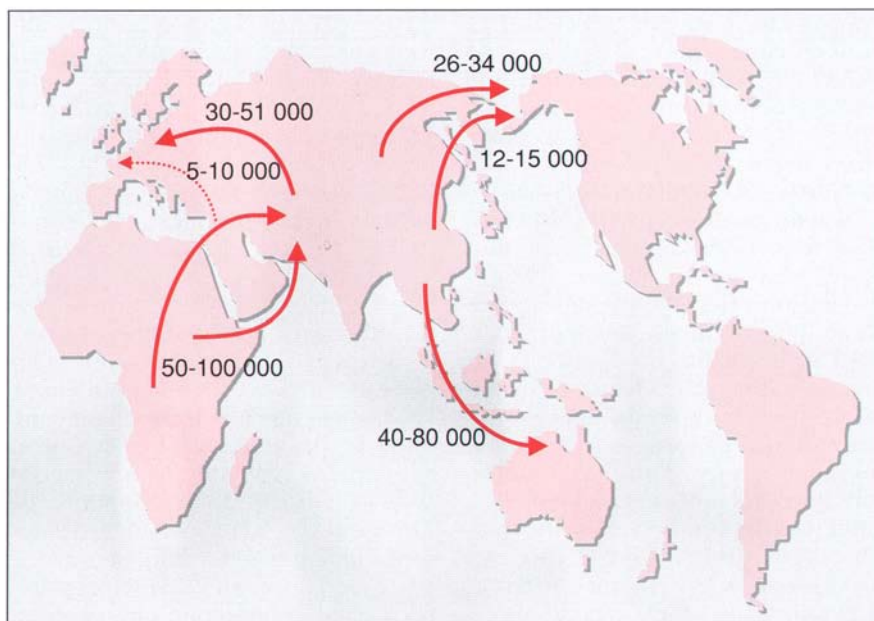


Figure 4. **Les grandes migrations de l'espèce humaine.** Après la sortie d'Afrique, il y a environ 60 000 à 70 000 ans, les homo sapiens sapiens ont migré vers les autres continents. Les trajectoires et les dates de colonisation sont issues des études de l'ADNmt et du chromosome Y. La flèche en pointillé représente l'expansion de l'économie fermière du néolithique, il y a 5 000-10 000 ans.

site de restriction Alu I (nt 10397), absent chez les Africains et les Européens, indique que cette mutation est apparue très peu de temps après ou avant l'arrivée des femmes en Asie. Ainsi, la présence du site Dde I et du site Alu I définit l'haplogroupe M considéré comme un marqueur spécifique des populations de l'Asie du Sud-Est. En fait, très récemment, l'haplogroupe M a aussi été retrouvé dans 20 % de la population éthiopienne [17]. Ce résultat est extrêmement intéressant car l'Éthiopie représente un lieu géographiquement stratégique pour comprendre la sortie des populations d'Afrique. En effet, il a été proposé que des *Homo sapiens sapiens* ont quitté l'Afrique soit par l'isthme de Suez, soit le long d'une route passant par l'Éthiopie et l'Inde de l'Ouest (figure 4). Il est donc probable que l'haplogroupe M retrouvé, d'une part, en Asie et, d'autre part, en Éthiopie ait une origine commune. Deux hypothèses sont donc possibles : ce marqueur a été acquis par les Éthiopiens à partir des Asiatiques ou, au contraire, il était déjà présent dans les populations anciennes d'Éthiopie qui ont migré hors d'Afrique et colonisé l'Asie. Si la deuxième hypothèse

est juste, ces résultats et l'absence de ce marqueur au Moyen-Orient conforteraient l'hypothèse d'une sortie récente d'*Homo sapiens sapiens* par l'Éthiopie vers l'Asie de l'Ouest [17, 18]. D'autres haplogroupes A-G ont pu être définis et l'étude de leur distribution géographique a été effectuée. Par exemple, l'haplogroupe F est surtout représenté dans les populations de l'Asie du Sud comme au Vietnam ou en Malaisie mais pratiquement absent en Sibérie. Au contraire, les haplogroupes A, C, D, E et G sont absents en Asie du Sud mais très présents dans les populations du Tibet, de Corée, de Chine et de Sibérie. Enfin, l'haplogroupe B est défini par la présence d'une délétion de 9 pb. Présent en Asie du Sud, en Asie centrale, sur les côtes et dans certaines îles du Pacifique, il est complètement absent en Sibérie. L'étude des variations génétiques asiatiques a permis de calculer la date de colonisation de l'Asie entre 56 000 et 73 000 ans [11].

L'Amérique

Les Indiens d'Amérique du Nord, du Centre ou du Sud appartiennent pratiquement tous aux quatre haplo-

groupes asiatiques A, B, C et D. Des études linguistiques ont permis de diviser les langues des Indiens d'Amérique en trois grands groupes [19] : le NaDéné est parlé par certaines tribus d'Indiens d'Amérique du Nord, du Canada et de l'Alaska. L'Eskalet est parlé par les Eskimos. Et surtout, l'Amérind regroupe une grande diversité de langues parlées par les premiers colonisateurs de l'Amérique occupant une partie du Nord et la totalité du Centre et du Sud de l'Amérique. Il est probable que ces trois groupes de langues proviennent de différentes vagues de migrations. Ainsi, l'étude des variations entre les séquences asiatiques et celles des Indiens d'Amérique parlant l'Amérind, a révélé que ces populations d'Amérindiens dérivent de deux vagues successives de migration passée par le détroit de Behring (figure 4). La première vague de migration (haplogroupes A, C et D) proviendrait d'Asie centrale en passant par la Sibérie il y a 26 000 à 34 000 ans, la deuxième vague de migration, beaucoup plus récente (haplogroupe B), serait passée par les côtes asiatiques, il y a seulement 12 000-15 000 ans. Enfin, l'âge de la migration des populations parlant le NaDéné est estimé à 7 200-9 000 ans [11-20].

L'Océanie

Par rapport aux autres populations, les populations océaniques ont été moins étudiées. Néanmoins, on retrouve dans les populations d'Australie, de Nouvelle-Guinée et de Mélanésie, l'haplogroupe M retrouvé principalement dans les populations d'Asie du Sud-Est. Le calcul de la divergence des séquences confirme l'âge relativement ancien des populations océaniques entre 40 250 et 80 500 ans.

Chromosome Y et populations humaines

L'étude des séquences spécifiques du chromosome Y a tout d'abord montré que les chromosomes Y présents dans les différentes populations humaines étaient très peu différents [6, 7]. Par exemple, aucune différence n'a pu être mise en évidence dans une séquence intronique

du gène *ZFY* de 729 pb parmi les différentes populations testées [21]. D'autres études ont identifié seulement 3 sites polymorphes dans 2,6 kb [22] et 3 substitutions de bases dans une séquence de plus 15 kb [23]. Enfin, récemment, notre équipe a montré qu'une duplication récente (55 000-200 000 ans) des gènes *DAZ* sur le chromosome Y était présente dans la très grande majorité des populations actuelles suggérant l'existence d'un ancêtre mâle commun récent à toutes les populations [24]. L'âge de cet ancêtre mâle commun le plus récent est évalué à 270 000 ans, 188 000 ans ou, encore plus récent, à 37 000-49 000 ans. Alors que les deux premières dates sont similaires à la date obtenue à partir de l'étude de l'ADNmt, la dernière semble beaucoup plus récente. De nouveau, beaucoup de critiques ont été faites à ces travaux portant principalement sur le taux de mutation, la taille et le mouvement des populations d'origine, et sur le nombre de données récoltées. Contrairement à l'ADNmt, les études portant sur les variations de séquences du chromosome Y ont progressé beaucoup plus lentement car très peu de polymorphismes informatifs avaient été mis en évidence. Cependant, ces dernières années, de nombreux marqueurs polymorphes spécifiques du chromosome Y ont été décrits et sont disponibles pour l'étude des populations [25, 26].

Une étude de deux marqueurs polymorphes du chromosome Y a permis de retracer les migrations récentes des populations du Moyen-Orient en Europe [27]. Comme nous l'avons déjà indiqué, l'Europe a été colonisée par *Homo sapiens sapiens* il y a environ 40 000-50 000 ans. Au cours du paléolithique, les hommes étaient principalement des chasseurs-cueilleurs. Puis, il y a environ 10 000 ans, l'agriculture s'est développée dans plusieurs régions du monde. L'un des sites les plus importants était situé au Moyen-Orient. Ces fermiers du néolithique se sont ensuite déplacés vers l'Europe, l'Arabie, l'Afrique du Nord et l'Inde apportant l'agriculture et leurs langues. Deux modèles ont été proposés pour comprendre cette expansion du néolithique : le « modèle

culturel » et le « modèle migratoire ». D'après le premier modèle, l'expansion s'est faite par la transmission culturelle des nouvelles technologies sans mouvement de population et donc sans changement de l'ensemble des gènes des populations autochtones. En revanche, le deuxième modèle prétend que l'expansion néolithique de l'agriculture provient des migrations de populations qui se sont progressivement mêlées aux populations de chasseurs-cueilleurs. Contrairement à la première version, ce modèle implique un changement de la fréquence des allèles dans les populations. Ainsi, les allèles spécifiques des populations de fermiers devraient être présents dans les régions colonisées mais leurs fréquences devraient décroître proportionnellement à la distance du lieu d'origine de la migration. Deux polymorphismes Taq I du chromosome Y détectés par deux sondes différentes p49f et p12F2 ont été étudiés chez environ 3 000 individus provenant de différentes populations : Europe (surtout le Bassin méditerranéen), Afrique et Asie [27]. La sonde p12F2 détecte un allèle de 8 kb uniquement présent chez les populations caucasiennes du Moyen-Orient, d'Afrique du Nord et d'Europe. En revanche, cet allèle est absent dans les populations d'Afrique subsaharienne et chez les Orientaux. La deuxième sonde, p49f, détecte un grand nombre de polymorphismes de restriction. En particulier l'haplotype Ht 15 qui n'est retrouvé qu'en Europe ou dans les populations ayant eu des contacts avec les Européens. La distribution de ces deux marqueurs s'est révélée très informative sur l'histoire de ces populations [27]. En effet, on peut observer un gradient de la fréquence de l'allèle 8 kb (p12F2) le long du Bassin méditerranéen. Ce polymorphisme est très présent dans les populations du Moyen-Orient mais sa fréquence diminue vers l'Europe occidentale. En particulier, il est pratiquement absent dans les populations basques. D'un autre côté, l'haplotype Ht 15 (p49f) est distribué contrairement à l'allèle 8 kb de p12F2. Sa fréquence la plus élevée se situe en Europe de l'Ouest. Environ 60 % de la population basque possède cet haplotype. Or, il est connu que la population basque a eu

peu de relations avec les nouveaux venus du néolithique et qu'elle parle une langue différente de l'Indo-européen. Ainsi, l'haplotype Ht 15 peut être considéré comme un haplotype proto-européen qui a été graduellement mélangé aux haplotypes des fermiers du néolithique provenant du Moyen-Orient.

En conclusion, la distribution opposée de ces marqueurs génétiques, d'est en ouest, a permis de bien illustrer la migration des populations du néolithique. A partir de ces résultats, il semble donc que l'expansion de l'économie fermière du néolithique provient d'une migration de populations plutôt que d'une transmission purement culturelle [27].

D'un autre côté, l'allèle 8 kb a été très utile pour comprendre certaines migrations des populations caucasiennes. En effet, l'allèle 8 kb est présent dans des populations ayant eu des liens avec le Moyen-Orient comme en Tunisie et en Algérie, révélant probablement la migration des Phéniciens au Maghreb. Enfin, comme nous le verrons plus loin, ce marqueur est aussi retrouvé en Éthiopie, confirmant les relations entre l'Éthiopie et le Moyen-Orient décrites par les historiens comme par les linguistes car la plupart des Éthiopiens parlent des langues d'origine sémitique.

Si les études du chromosome Y sont complémentaires des études de l'ADNmt, elles peuvent aussi donner des résultats apparemment contradictoires très intéressants. En effet, l'histoire du chromosome Y peut être différente de celle de l'ADNmt. Ainsi, l'étude comparée de ces marqueurs permet d'illustrer les différences liées au sexe lors de certaines pratiques culturelles, des migrations, des guerres et des colonisations.

Un bon exemple de cette différence est illustré par les résultats d'une étude récente des populations éthiopiennes [18]. Nous avons déjà expliqué pourquoi la position géographique de l'Éthiopie est stratégique dans le modèle proposant une sortie récente des populations d'*Homo sapiens sapiens* hors d'Afrique. Ainsi, l'étude parallèle de l'ADNmt et du chromosome Y de la population éthiopienne a été effectuée dans le but d'identifier les origines maternelles et/ou paternelles de ces populations.

Quand on compare les variations de l'ADNmt et du chromosome Y, on remarque que, pour les séquences d'origine caucasienne, on trouve une différence importante : 5,4 % des ADNmt et 25,4 % des chromosomes Y. Il est donc possible que la contribution caucasienne soit principalement d'origine masculine. On peut donc penser que ce sont surtout des hommes provenant du Moyen-Orient qui ont migré vers l'Éthiopie, alors que la contribution subsaharienne des populations éthiopiennes est surtout d'origine féminine (*m/s* 1999, n° 1, p. 126).

Conclusions

A côté de l'étude des origines de l'espèce humaine, il subsiste de nombreux aspects de l'histoire des populations qui restent dans l'ombre. En particulier, il reste à préciser les dates et l'origine géographique des différentes vagues de migrations d'*Homo sapiens sapiens* hors d'Afrique et de retracer les routes utilisées pour coloniser les autres continents. En effet, il est possible qu'il y ait eu une deuxième sortie d'Afrique, il y a 40 000-50 000 ans à l'origine de certaines populations d'Europe et d'Asie du Nord. Enfin, de nombreuses questions restent en suspens qui concernent, d'une part, le peuplement de l'Asie du Sud et de l'Australie (il y a environ 60 000 ans passant vraisemblablement le long des côtes de l'Océan Indien), d'autre part, les variations génétiques entre les anciens groupes africains comme les Pygmées et les Khoisiens. Plus particulièrement en Europe, il reste à préciser l'origine des populations basques et la relation entre l'*Homo sapiens neandertalensis* et les populations d'*Homo sapiens sapiens*. Toutes ces questions sont en cours d'investigation dans les laboratoires de génétique. De plus, les chercheurs sont aidés par de nouvelles technologies permettant par exemple de séquencer très rapidement toute la molécule d'ADNmt d'un individu en quelques heures. Il est donc certain que de nombreux résultats plus précis vont être présentés rapidement qui devront être intégrés dans l'ensemble des résultats provenant d'autres disciplines comme l'anthropologie, l'archéologie, la linguistique et la paléontologie ■

Note ajoutée aux épreuves

Depuis la rédaction de cet article, les auteurs ont identifié un marqueur mitochondrial qui supporte une deuxième sortie d'Afrique des hommes modernes, il y a environ 60 000 ans. La route prise par ces populations partirait de l'Éthiopie vers l'Inde en passant par les côtes de l'Arabie Saoudite. D'après les données archéologiques, il semblerait que la première sortie d'Afrique, il y a 100 000 ans par le Moyen-Orient, n'aurait pas eu de succès. En prenant en compte l'ensemble de ces résultats, il est plus probable que l'origine de la majorité des populations eurasiennes actuelles provienne de cette deuxième sortie plus récente.

Remerciements

Nous remercions Marie-Louise Prunier et Stéphane Jamain pour la lecture critique du manuscrit.

RÉFÉRENCES

1. Cavalli-Sforza LL, Menozzi P, Piazza A. *The history and geography of human genes*. Princeton : Princeton University Press, 1994.
2. Tyler D. *The mitochondrion in health and disease*. New York : VCH Publishers, 1992.
3. Stoneking M, Soodyall H. Human evolution and the mitochondrial genome. *Curr Opin Genet Dev* 1996 ; 6 : 731-6.
4. Lahn BT, Page DC. Functional coherence of the human Y chromosome. *Science* 1997 ; 278 : 675-9.
5. Barbaux S, Vilain E, McElreavey K, Fellous M. Le point sur le déterminisme du sexe chez les mammifères. *Med Sci* 1995 ; 11 : 529-36.
6. Jobling M, Tyler-Smith C. Fathers and sons: the Y chromosome and human evolution. *Trends Genet* 1995 ; 11 : 449-55.
7. Mitchell RJ, Hammer MF. Human evolution and the Y chromosome. *Curr Opin Genet Dev* 1996 ; 6 : 737-42.
8. Jones S, Martin R, Pilbeam D. *The Cambridge encyclopedia of human evolution*. Cambridge : Cambridge University Press, 1992.
9. Krings M, Stone A, Schmitz RW, Krainitzki H, Stoneking M, Pääbo S. Neandertal DNA sequences and the origin of modern humans. *Cell* 1997 ; 90 : 19-30.
10. Denamur E, Lecomte G. L'ADN de l'homme de Neandertal. *Med Sci* 1997 ; 13 : 1488-90.
11. Wallace D. Mitochondrial DNA variation in human evolution, degenerative disease, and aging. *Am J Hum Genet* 1995 ; 57 : 201-23.
12. Cann RL, Stoneking M, Wilson AC. Mitochondrial and human evolution. *Nature* 1987 ; 325 : 31-6.
13. Denaro M, Blanc H, Johnson MJ, Chen KH, Wilmsen E, Cavalli-Sforza LL, Wallace DC. Ethnic variation in Hpa I endonuclease cleavage patterns of human mitochondrial DNA. *Proc Natl Acad Sci USA* 1981 ; 78 : 5768-72.

RÉFÉRENCES

14. Vigilant L, Stoneking M, Harpending H, Hawkes K, Wilson AC. African populations and the evolution of human mitochondrial DNA. *Science* 1991; 253: 1503-7.
15. Ballinger SW, Schurr TG, Torroni A, Gan YY, Hodge JA, Hassan K, Chen KH, Wallace DC. Southeast Asian mitochondrial DNA analysis reveals genetic continuity of ancient mongoloid migrations. *Genetics* 1992; 130: 139-52.
16. Torroni A, Huoponen K, Francalacci P, Petrozzi M, Morelli L, Scozzari R, Obinu D, Savontaus ML, Wallace DC. Classification of European mtDNAs from an analysis of three European populations. *Genetics* 1996; 144: 1835-50.
17. Passarino G, Semino O, Bernini LF, Santachiara-Benerecetti AS. Pre-Caucasoid and Caucasoid genetic features of the Indian population, revealed by mtDNA polymorphisms. *Am J Hum Genet* 1996; 59: 927-34.
18. Passarino G, Semino O, Quintana-Murci L, Excoffier L, Hammer M, Santachiara-Benerecetti AS. Different genetics components of the Ethiopian population identified by mtDNA and Y-chromosome polymorphisms. *Am J Hum Genet* 1998; 62: 420-34.
19. Greenberg JH, Turner CGII, Zegura SL. The settlement of the Americas: a comparison of the linguistic, dental, and genetic evidence. *Curr Anthropol* 1986; 27: 477-97.
20. Torroni A, Neel JV, Barrantes R, Schurr TG, Wallace DC. A mitochondrial DNA «clock» for the Amerinds and its implication for timing their entry into North America. *Proc Natl Acad Sci USA* 1994; 91: 1158-62.
21. Dorit RL, Akashi H, Gilbert W. Absence of polymorphism at the ZFY locus on the Y chromosome. *Science* 1995; 268: 1183-5.
22. Hammer MF. A recent common ancestry for human Y chromosomes. *Nature* 1995; 378: 376-8.
23. Whitfield LS, Sulston JE, Goodfellow PN. Sequence variation of the human Y chromosome. *Nature* 1995; 378: 379-80.
24. Groupe 1: Alguinik AI, Zharkikh A, Boettger-Tong H. Groupe 2: Bourgeron T, McElreavey K, Bishop CE. Evolution of the *DAZ* gene family suggests that Y-linked *DAZ* plays little, or a limited, role in spermatogenesis but underlines a recent African origin for human populations. *Hum Mol Genet* 1998; 7: 1371-7.
25. Underhill PA, Jin L, Lin AA, Mehdi SQ, Jenkins T, Vollrath D, Davis RW, Cavalli-Sforza LL, Oefner PJ. Detection of numerous Y chromosome biallelic polymorphisms by denaturing high-performance liquid chromatography. *Genome Res* 1997; 7: 996-1005.
26. Hammer MF, Spurdle AB, Karafet T, et al. The geographic distribution of human Y chromosome variation. *Genetics* 1997; 145: 787-805.
27. Semino O, Passarino G, Brega A, Fellous M, Santachiara-Benerecetti AS. A view of the Neolithic demic diffusion in Europe through two Y chromosome-specific markers. *Am J Hum Genet* 1996; 59: 964-8.

TIRÉS À PART

T. Bourgeron.

Summary

Mitochondrial DNA, Y chromosome and human population history

Many disciplines such as archaeology, anthropology, linguistic and paleontology are involved in the study of human origins. More recently, genetics and molecular biology have been used to shed light on human origins. For many years, geneticists have studied the DNA variations among individuals and populations in order to establish relationships between different populations. Of the DNA markers available, the most interesting are the uniparental inherited markers, which are the maternally inherited mtDNA and the paternally inherited Y-chromosome. They both escape from regular recombinational processes at meiosis and, consequently, these markers are transmitted together as haplotypes preserving a unique record of mutational changes that have occurred in previous generations. While X chromosomes and autosomes each have multiple ancestors because of recombination, all modern mtDNAs could have a single maternal ancestor and Y-chromosomes could have a single paternal ancestor. Here we review the general characteristics of the Y chromosome and the mtDNA and, using specific examples, we show how haplotypes can be used to determine human origins and study different population interactions in historical times.