

## Mise au point

# Génétique et évolution du génome mitochondrial des Métazoaires

P. BOURSOT et F. BONHOMME

*Institut des Sciences de l'Evolution (U.A. 327 du C.N.R.S.)  
Université des Sciences et Techniques du Languedoc,  
place E.-Bataillon, F 34060 Montpellier Cedex*

## Résumé

Dans cette revue nous faisons tout d'abord le point sur les connaissances actuelles concernant la génétique de la transmission et le mode d'évolution moléculaire du génome mitochondrial animal. Nous illustrons ensuite comment les caractéristiques propres à ce génome en font un marqueur génétique original et prometteur pour la biologie évolutive : il est remarquablement polymorphe dans les populations naturelles, et les techniques de biologie moléculaire (en particulier l'analyse de polymorphisme de longueur de fragments de restriction) permettent de révéler aisément les variations de sa séquence d'acides nucléiques ; sa transmission maternelle et son évolution clonale en font un marqueur direct et non ambigu de la généalogie maternelle et de la structuration géographique au sein d'une espèce, ainsi que des échanges génétiques entre populations, entre sous-espèces. Son évolution rapide, en raison d'un taux de mutation particulièrement élevé, autorise son utilisation dans des comparaisons génétiques et des reconstitutions phylogénétiques entre taxons proches, et même de niveau infraspécifique. Toutefois dans un certain nombre de cas déjà rapportés, la phylogénie mitochondriale s'avère en désaccord avec la phylogénie des espèces préalablement établie sur d'autres critères. Nous discutons ici les causes possibles de telles discordances et analysons les rapports possibles entre la phylogénie des espèces et celle des lignées mitochondriales.

*Mots clés : ADN mitochondrial, phylogénie moléculaire, génétique des populations, Métazoaires.*

## Summary

### *Genetics and evolution of the metazoan mitochondrial genome*

In this review we first survey current knowledge of mitochondrial transmission genetics and mode of evolution in animals. We then illustrate how certain characteristics proper to this genome make it an original and promising genetical marker in evolutionary biology. It is remarkably polymorphic in natural populations and, using the techniques of molecular biology, such as restriction fragment length polymorphism, the variations in its nucleic acid sequence are easily detected. Its maternal transmission and clonal evolution make it a straightforward and unambiguous marker of the matriarcal genealogy and geographical structure among conspecifics. It can also reveal genetical exchanges between populations, subspecies and even exchanges between

species. Because of its rapid evolution, due to an exceptionally high mutation rate, it is a suitable marker for the genetical comparison and phylogenetic inference between closely related taxa, down to the infraspecific level. Yet in a certain number of documented cases, mitochondrial phylogeny does not fit with the phylogeny of the species established using other criteria. We discuss here the possible causes of such discrepancies and analyse possible relationships between the phylogeny of the species and that of mitochondrial lineages.

*Key words* : Mitochondrial DNA, molecular evolution, population genetic, Metazoa.

### I. Organisation moléculaire et variation

Les données de la séquence complète de l'ADN mitochondrial (ADNmt) d'une souris (BIBB *et al.*, 1981), d'un homme (ANDERSON *et al.*, 1981) et d'un bœuf (ANDERSON *et al.*, 1982) ont permis d'identifier 13 gènes protéiques, ainsi que les gènes des 22 ARNt et des ARNr spécifiques de la mitochondrie (fig. 1). L'ADNmt est une molécule circulaire, double brin. Sa longueur totale est respectivement de 16 295, 16 569 et 16 338 paires de bases chez la souris, l'homme et le bœuf étudiés par ces auteurs. Il apparaît une grande économie d'organisation de ce génome : les gènes sont quasi-contigus sur le chromosome et ne contiennent aucun intron. Les parties non codantes sont limitées à de très courtes séquences et à une région plus longue correspondant à l'origine de réplication du brin lourd (région de formation de la « D-loop »). Le code génétique mitochondrial présente quelques différences avec celui du noyau (BARRELL *et al.*, 1979, 1980).

L'organisation générale du génome mitochondrial (taille de la molécule, nombre et nature des gènes) semble remarquablement conservée dans l'ensemble du règne animal (voir par exemple BROWN, 1983 pour une revue de ce problème). En particulier, chez les 3 espèces de mammifères où il a été séquencé, le nombre de gènes, leur nature et leur position relative sur la molécule sont identiques. Chez les drosophiles, où l'ADNmt a été particulièrement séquencé, l'ordre des gènes sur la molécule est différent de celui trouvé chez les mammifères, bien que leur nombre et leur nature semblent très voisins (voir WOLSTENHOLME & CLARY, 1985 ; DE BRUIJN, 1983 et références incluses). Remarquons cependant que d'autres groupes d'organismes montrent une organisation très différente et très variée de l'ADNmt (GRIVELL, 1983), les différences portant aussi bien sur la taille (jusqu'à plusieurs milliers de kilobases chez certaines plantes), sur l'organisation générale (linéaire par exemple chez *Tetrahymena*), sur la présence d'introns (cas de la levure) ou sur l'utilisation du code génétique. Nous nous limiterons ici au cas des animaux, parmi lesquels les mammifères ont été les plus étudiés jusqu'à présent en ce qui concerne l'ADNmt.

Une aussi bonne conservation générale semble s'opposer à la constatation de la très forte variabilité de ce génome dans les populations naturelles, et à son évolution rapide, qui ont été constatées chez plusieurs espèces de mammifères, comme nous le verrons plus loin.

En fait, l'analyse de cette variabilité a montré que l'essentiel des transformations jusqu'à présent observées pouvait être attribué à des substitutions nucléotidiques. Il n'a pas été décrit d'événements de remaniement importants. Chez les animaux supérieurs, on n'a jamais mis en évidence de recombinaison de l'ADN mitochondrial. On sait qu'il peut se produire des événements d'addition ou délétion d'une ou plusieurs paires de bases. Ces événements passent inaperçus dans les analyses courantes de polymorphisme



de longueur de fragments de restriction (PLFR) mais peuvent être révélés par des techniques plus résolutive (CANN & WILSON, 1983 ; HAUSWIRTH *et al.*, 1984) ou par séquençage (GREENBERG *et al.*, 1983 ; FORT *et al.*, 1984). Ils semblent limités aux parties non codantes, dont la région de l'origine de réplication et de transcription. Chez les drosophiles, des variations de longueur plus importantes ont été décrites entre différentes espèces (SHAH & LANGLEY, 1979 : 18 600 paires de bases pour *D. melanogaster* et *D. simulans* contre 15 800 pour *D. virilis*) mais aussi au sein d'une même espèce (SOLIGNAC *et al.*, 1983, chez *D. mauritiana*). Ces modifications de taille semblent restreintes à une région non codante riche en A-T.

Les données de PLFR ainsi que celles de séquençage donnent accès à une estimation du taux de substitutions nucléotidiques de l'ADNmt. Chez les mammifères, où il a pu être estimé précisément, ce taux est beaucoup plus élevé que celui décrit pour les séquences uniques nucléaires, et ce d'un facteur multiplicatif 5 ou 10 (voir BROWN *et al.*, 1982 pour une étude détaillée chez les primates). Ce taux est variable suivant les régions du génome considérées (CANN *et al.*, 1984 ; AQUADRO & GREENBERG, 1983 ; BROWN *et al.*, 1982 chez l'homme et les primates ; FERRIS *et al.*, 1983 b chez la souris) mais reste globalement élevé. Dans les régions codant pour des protéines, BROWN *et al.*, (1982), constate une large prédominance de substitutions silencieuses et estiment leur taux à 10 p. 100 par million d'années, valeur dix fois supérieure à la valeur correspondante pour les gènes nucléaires. On constate de plus un biais très important en faveur des événements mutationnels de type transition (remplacement d'une purine par l'autre ou d'une pyrimidine par l'autre) par rapport à ceux de type transversion (voir toutefois FORT *et al.*, 1984 pour une exception). Ce biais favorise donc les événements de substitutions multiples par transitions à un même site, et doit être pris en compte dans l'estimation des taux de substitutions nucléotidiques (BROWN *et al.*, 1982).

Plusieurs types d'explications peuvent être invoqués pour expliquer le fort taux de mutation de l'ADNmt constaté chez les mammifères. On peut citer : une réplication peu fidèle (selon KUNKEL & LOEB (1981) la polymérase mitochondriale serait moins fidèle que celle du noyau), une déficience ou une absence des systèmes de correction et de réparation, et l'absence apparente de recombinaison, ainsi qu'un taux de renouvellement et donc de réplication plus important que celui de l'ADN nucléaire (BROWN *et al.*, 1982). En ce qui concerne l'appareil de traduction (ARNt et ARNr) on a pu émettre d'autres hypothèses d'ordre fonctionnel (BROWN *et al.*, 1979, 1982 ; CANN *et al.*, 1984).

La question reste ouverte de savoir si cette grande vitesse d'évolution vaut aussi pour d'autres organismes que les mammifères. BROWN (1983) suggère que cela pourrait ne pas être le cas chez les drosophiles, dont la polymérase mitochondriale différerait de celle des mammifères.

## II. Biologie et génétique de la transmission

Une cellule somatique d'eucaryote contient de multiples copies de la molécule d'ADNmt. BOGENHAGEN & CLAYTON (1974) ont estimé leur nombre à 900 et 1 100 ( $\pm 250$ ) respectivement dans 2 lignées de cellules L de souris. La manière dont ces différentes copies sont réparties dans les différentes mitochondries (indépendance physiologique entre organelles, nombre de molécules dans chacune d'entre elles) n'est pas connue

à l'heure actuelle. Dans ce qui suit, nous opérerons un raccourci en considérant une population de molécules et non pas d'organelles, ce qui permet de décrire *a posteriori* la transmission de ce génome sans en connaître les mécanismes cellulaires fins.

La compréhension du mode d'évolution du génome mitochondrial passe donc par deux niveaux. Outre celui de la génétique des populations d'individus par l'intermédiaire desquels il se transmet d'une génération à la suivante, il faut considérer celui de la dynamique de la population de molécules mitochondriales contenue dans chaque individu. En effet, d'une part l'ADNmt se réplique plus rapidement que le génome nucléaire, et d'autre part ce sont des centaines de copies d'ADNmt qui, à chaque division cellulaire, se partagent entre les cellules filles.

Cette dynamique propre de l'ADNmt est déterminante pour les phénomènes que nous observons ; tout d'abord au sein de la lignée cellulaire somatique, puisque c'est par elle que nous décrivons les variations et l'évolution du génome cytoplasmique, et ensuite au sein de la lignée germinale puisque, représentant le lien d'une génération à la suivante, elle est le siège des transformations observées. Les modèles élaborés dans ce domaine (voir par exemple CHAPMAN *et al.*, 1982 ; BIRKY *et al.*, 1983 ; TAKAHATA & MARUYAMA, 1981 ; BIRKY, 1983) soulignent l'importance théorique de paramètres tels que le nombre de générations cellulaires dans la lignée germinale, la taille efficace des populations de mitochondries, leur mode de ségrégation durant la division cellulaire. Il existe peu de données expérimentales dans ce domaine, et en tout cas aucune observation directe de ces phénomènes.

Toutefois, indépendamment des mécanismes qui en sont responsables, quelques traits de l'évolution de l'ADNmt des animaux semblent établis dans leurs grandes lignes.

#### A. Homogénéité intra-individuelle de la séquence des différentes copies d'ADNmt

La possibilité d'utiliser le génome mitochondrial comme marqueur génétique suppose que les différentes copies d'ADNmt contenues dans un individu présentent une certaine homologie de séquence, c'est-à-dire que l'hétérogénéité intra-individuelle (ou hétéroplasmie) soit faible par rapport à l'hétérogénéité interindividuelle. L'étude de la variabilité de l'ADN mitochondrial sur plusieurs espèces de mammifères (en particulier l'homme, la souris, le rat et les rongeurs américains des groupes *Peromyscus* et *Geomys*) ainsi que sur plusieurs espèces de drosophiles, qui a mis en jeu l'analyse de centaines d'individus, n'a généralement pas mis en évidence l'existence d'individus hétéroplasmiques, c'est-à-dire possédant plusieurs types d'ADNmt différant par leur séquence d'acides nucléiques. Dans tous les cas, chaque individu peut être caractérisé par un génome cytoplasmique unique et, lorsque cela a pu être vérifié, une femelle transmet ce même génome à ses descendants.

Ceci ne signifie pas que l'hétéroplasmie n'existe pas. Pour qu'il y ait apparition et fixation de variants nouveaux — et nous avons vu que cette évolution est très rapide pour l'ADNmt — il y a nécessairement passage par un état hétéroplasmique. Cependant pour que l'évolution se fasse, il suffit que cet état existe au sein de la lignée germinale (voir discussion dans la troisième section de ce chapitre).

Il faut toutefois noter que dans les analyses de PLFR de routine, une forme moléculaire distinguable des autres mais d'abondance réduite (5 à 10 p. 100) peut

passer inaperçue. Ajoutons que l'analyse par les endonucléases ne révèle qu'une partie des substitutions nucléotidiques affectant l'ensemble du génome mitochondrial. Quant au séquençage d'ADN, il implique généralement une étape de clonage, et donc l'isolement d'une molécule particulière. Le clonage peut ne pas être nécessaire (séquençage direct d'un ADNc, FORT *et al.*, 1984), mais la séquence obtenue n'est qu'une séquence consensus qui, à chaque position nucléotidique, révèle la base majoritairement représentée. Les variants minoritaires passent inaperçus.

Quelques études ont été spécifiquement conçues en vue de rechercher l'hétéroplasmie en comparant les ADNmt obtenus à partir de différents tissus d'un même individu (FRANCISCO *et al.*, 1979 ; POTTER *et al.*, 1975). Il n'a été trouvé aucune différence entre tissus. COOTE *et al.* (1979) signalent des différences entre tissus d'un même bœuf, sans qu'il soit toutefois possible d'exclure la possibilité d'une digestion incomplète par les endonucléases.

HAUSWIRTH & LAIPIS (1982) ont trouvé, en étudiant des familles de vaches *Holstein* de généalogie connue, des descendants d'une même mère qui possédaient des ADNmt différents. La mère pourrait donc avoir transmis 2 types différents, mais il n'a pas été trouvé d'individu portant simultanément les 2 types.

HAUSWIRTH *et al.* (1984) mettent en évidence une hétérogénéité intra-individuelle d'une région de la « D-loop », toujours chez des vaches *Holstein*. Ils montrent qu'il s'agit d'une région hypervariable de la molécule, où se produisent très fréquemment des événements d'addition ou de délétion de quelques paires de bases.

Chez *Drosophila mauritiana*, SOLIGNAC *et al.* (1983) repèrent dans une lignée isofemelle la présence de 2 formes mitochondriales de longueurs différentes (500 paires de bases de différence) et démontrent l'existence d'individus hétéroplasmiques. Cette hétéroplasmie semble assez stable puisqu'à la 30<sup>e</sup> génération la moitié des individus (sur 60 analysés) présentent encore cette caractéristique.

La rareté des cas d'hétéroplasmie attestée (\*) se comprend si on remarque que la possibilité de les observer résulte en fait de coïncidences peu probables :

— les différences entre les formes moléculaires simultanément présentes dans un même individu doivent être repérables par les techniques employées. Nous avons vu plus haut les limites de ces techniques, aussi bien pour l'analyse des fragments de restriction que pour la séquence nucléotidique. La probabilité de détection de variants mitochondriaux dans un même individu est encore diminuée par le fait que ces variants ont toutes chances d'être très peu différents (en termes de taux de substitution nucléotidiques, c'est-à-dire de temps de divergence). Si cela n'était pas le cas, cela signifierait que des lignées mitochondriales distinctes peuvent coexister pendant longtemps au sein d'une même lignée matriarcale. On comprendrait mal alors la rareté des cas d'hétéroplasmie ;

— il doit coexister au moins deux formes moléculaires d'abondances non négligeables, qui doivent coïncider avec celles différenciables par ces techniques.

La découverte de tels cas est cependant capitale pour la compréhension de la dynamique des populations de mitochondries de la lignée germinale.

(\*) Depuis la mise sous presse de cet article, quelques cas nouveaux d'hétéroplasmie ont été décrits dans la littérature, chez des Insectes et des Vertébrés.

### B. Transmission maternelle de l'ADNmt

On admet que chez les animaux supérieurs la transmission du génome cytoplasmique d'une génération à l'autre se fait uniquement par voie maternelle, par le cytoplasme de l'ovule.

On a constaté en effet, chez plusieurs espèces (*Peromyscus* : AVISE *et al.*, 1979 b, *Rattus* : FRANCISCO *et al.*, 1979 ; HAYASHI *et al.*, 1978, *Xenopus* : DAWID & BLACKER, 1972 et chez l'homme : GILES *et al.*, 1980) que les descendants de 2 parents dont les génomes cytoplasmiques sont distinguables possèdent toujours celui de la mère.

Toutefois l'éventualité de l'intégration dans l'œuf des mitochondries paternelles (contenues dans la pièce intermédiaire du spermatozoïde) en faible proportion ne peut pas être exclue puisque, comme nous l'avons vu plus haut, la détection des variants minoritaires est délicate du point de vue expérimental. Chez la souris, certaines observations (revues par FISCHER-LINDAHL, 1985) tendraient à montrer que les mitochondries du spermatozoïde pénètrent dans l'œuf lors de la fécondation mais y sont ensuite désintégrées.

En admettant que toutes les mitochondries du spermatozoïde se retrouvent dans l'œuf, elles resteraient de toute manière minoritaires. L'ovule est une grosse cellule contenant des dizaines de milliers de mitochondries alors que le spermatozoïde n'en apporte qu'une centaine environ. PIKO & MATSUMATO (1976), par observation directe, évaluent à environ 100 000 le nombre de mitochondries dans une cellule œuf de souris. MICHAELS *et al.* (1982) trouvent dans les oocytes de la vache environ 260 000 molécules d'ADNmt.

Deux expériences approfondies ont visé à rechercher l'existence d'une transmission paternelle. La première porte sur des insectes lépidoptères (*Héliotis*). Les auteurs (LANSMAN *et al.*, 1983 b) partent d'une lignée femelle fertile issue de l'hybridation entre 2 espèces possédant des génomes cytoplasmiques distincts. Cette lignée est croisée en retour avec des mâles de l'espèce père de l'hybride pendant 91 générations. Malgré les techniques très sensibles employées, on ne détecte pas trace du génome mitochondrial d'origine paternelle chez ces individus de 91<sup>e</sup> génération. La deuxième expérience (GYLLENSTEN *et al.*, 1985) utilise le même principe de croisements en retour sur des hybrides interspécifiques, avec 2 espèces de souris (*M. m. domesticus* et *M. spretus*). Ils ne mettent pas en évidence de transmission paternelle et estiment, en fonction de la précision des méthodes employées, à moins de un pour mille l'importance d'une éventuelle contribution paternelle non détectée.

### C. Relations entre variabilités intra- et interindividuelles : une vision critique

Nous verrons plus bas qu'il existe une forte variabilité interindividuelle (polymorphisme) de l'ADNmt dans les populations naturelles. Ce rapport entre hétérogénéités intra- et interindividuelles est sous la dépendance des mécanismes de la réplication de l'ADNmt et de la partition des différentes copies lors de la division cellulaire. Un simple modèle aléatoire peut-il rendre compte des phénomènes observés ?

Remarquons déjà que l'ensemble des cellules constitutives d'un organisme sont issues de la multiplication du zygote. La dérive au cours de cette multiplication conduit à une réduction de l'hétérogénéité de l'ensemble des molécules d'ADNmt dans l'adulte par rapport à celle existant dans le zygote (on entend ici par hétérogénéité la

probabilité que 2 copies prises au hasard soient différentes, BIRKY *et al.*, 1983). Toutefois cette différence n'est substantielle que si l'effectif efficace de molécules d'ADNmt par cellule est faible.

CHAPMAN *et al.* (1982) ont simulé l'échantillonnage au hasard des mitochondries dans la lignée germinale. Dans leurs simulations, l'hétérogénéité interindividuelle (probabilité que 2 copies d'ADN mitochondrial prises au hasard dans 2 individus quelconques soient différentes) oscille autour d'une valeur qui correspond à l'équilibre entre la mutation et la dérive. Parallèlement l'hétérogénéité intra-individuelle peut atteindre des niveaux très bas à condition que la taille efficace de la population de mitochondries de la lignée germinale soit faible (10-100). On peut définir ici la taille efficace comme le nombre théorique d'unités indépendantes qui, si elles ségrégeaient au hasard d'une génération cellulaire à la suivante, donneraient l'hétérogénéité observée. Le nombre de mitochondries de l'ovule (plusieurs dizaines de milliers) peut ne pas donner une bonne idée de la valeur réelle de ce paramètre, soit que l'échantillonnage des mitochondries au cours des générations cellulaires ne se fasse pas au hasard, soit plus probablement parce qu'il y aurait augmentation du nombre de mitochondries au sein même de l'ovule, dont le cytoplasme est exceptionnellement volumineux. Nous avons vu que des cellules plus petites telles que les cellules somatiques contiennent 10 à 100 fois moins de mitochondries, ce qui ramène à un ordre de grandeur plus raisonnable pour les ovocytes.

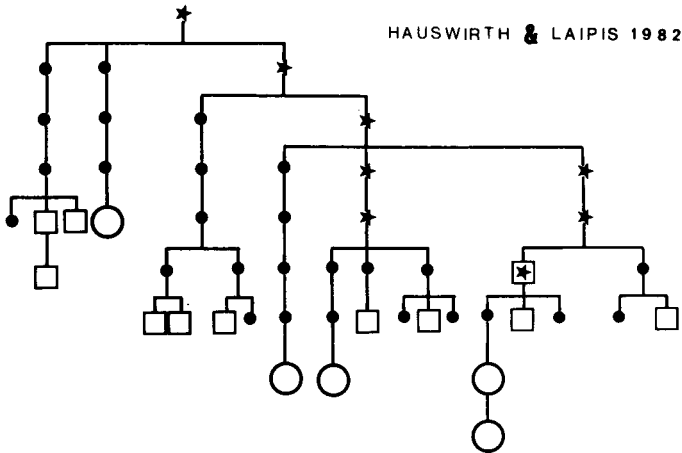


FIG. 2

*Ségrégation de 2 variants mitochondriaux dans une lignée maternelle de vaches Holstein (d'après HAUSWIRTH & LAIPIS, 1982).*

*Segregation of 2 mitochondrial variants in a maternal lineage of Holstein Cows (from HAUSWIRTH & LAIPIS, 1982).*

Sur cet arbre généalogique, chaque symbole représente un individu. Un certain nombre d'individus ont été analysés pour leur génome mitochondrial. Deux variants mitochondriaux ont été trouvés, et ces individus sont représentés par un carré ou un rond blanc suivant qu'ils présentent l'un ou l'autre des 2 variants mitochondriaux. On a marqué d'une étoile les femelles dans la descendance desquelles on trouve les 2 cytoplasmes ségrégeant.

On this genealogy every symbol represents one individual. Some of the individuals were typed for mitochondrial DNA, and two variants were found. These individuals are represented by either an open circle or an open square, depending on the mitochondrial variant they show. We indicate with a star the females in whose descendants the 2 mitochondrial variants segregate.

Ce paramètre de taille efficace a pu être évalué indirectement chez la drosophile (*D. mauritiana*, SOLIGNAC *et al.*, 1984) grâce à la mise en évidence d'une hétéroplasmie facilement détectable (présence de 2 formes S et L différant par leurs longueurs, et en fréquences comparables). L'étude de la ségrégation de ces 2 formes au cours des générations d'individus dans plusieurs lignées maternelles issues de femelles hétéroplasmiques montre une bonne compatibilité avec un modèle de ségrégation aléatoire des mitochondries au cours des générations cellulaires. Le nombre d'unités indépendantes ségrégeant ainsi est estimé à 400.

HAUSWIRTH & LAIPIS (1982) trouvent au sein d'une même lignée maternelle de vaches la ségrégation de 2 morphes (fig. 2). Ceci pourrait suggérer que la femelle ancêtre, ainsi qu'un certain nombre de ses descendantes, étaient hétéroplasmiques. Or les individus analysés (16 individus de la quatrième à la huitième génération) ne sont jamais hétéroplasmiques mais possèdent uniquement l'un des 2 types mitochondriaux. La répartition sur l'arbre généalogique des 2 types trouvés (fig. 2) suggère qu'il existe des femelles homoplasmiques pour l'un des 2 morphes et qui présentent cependant les 2 dans leur descendance. Le problème le plus aigu ici nous semble être les bases expérimentales sur lesquelles on affirme qu'il y a ou non hétéroplasmie somatique. L'hétéroplasmie somatique peut signifier que toutes les cellules contiennent plusieurs types de mitochondries, ou que différentes cellules en contiennent différents types. Dans ce deuxième cas, il se pourrait qu'il y ait des différences entre organes d'un même individu. Comme nous l'avons signalé ci-dessus, des comparaisons entre différents organes d'un même individu ont été faites (POTTER *et al.*, 1975 ; COOTE *et al.*, 1979 ; FRANCISCO *et al.*, 1979) mais dans des cas où l'on n'avait pas *a priori* de raison de penser qu'il pouvait exister une hétéroplasmie détectable par les moyens employés. Par contre, dans les cas où l'on soupçonne l'hétéroplasmie et où l'on a les moyens de la mettre en évidence (cas des vaches *Holstein*), de telles comparaisons entre organes manquent. Si elles révélaient des différences entre tissus, il resterait à expliquer comment les cellules germinales pourraient rester hétérogènes alors que différents types mitochondriaux se fixeraient dans différents tissus. On pourrait le comprendre en remarquant d'une part que chaque organe est issu d'une lignée cellulaire qui lui est propre et d'autre part que les cellules somatiques continuent à se diviser durant toute la vie de l'animal. Il y aura donc une dérive beaucoup plus importante dans la lignée somatique que dans la lignée germinale qui, au moins chez la femelle, arrête très tôt sa multiplication.

Un paramètre important dans l'évolution du génome mitochondrial est le temps de rétention de l'hétérogénéité dans une même lignée germinale maternelle. Chez la drosophile (SOLIGNAC *et al.*, 1984), après 30 générations, la moitié des descendants d'une femelle hétéroplasmique possèdent encore cette caractéristique. Chez les vaches *Holstein*, on peut seulement dire d'après les données de HAUSWIRTH & LAIPIS (1982) que l'hétérogénéité de la lignée germinale s'est maintenue pendant au moins 5 générations (fig. 2).

Tous les modèles de biologie des populations du génome mitochondrial élaborés jusqu'à présent (voir par exemple une revue par TAKAHATA, sous presse) partent de l'hypothèse qu'il existe une dérive très forte pour les génomes extranucléaires lors de la division cellulaire, de sorte que la fixation d'un seul type mitochondrial dans chaque cellule est très rapide (quasi instantanée). Cette hypothèse repose principalement sur des observations effectuées chez les levures et certaines plantes, revues par BIRKY (1983). Il existe peu de données dans ce domaine chez les animaux ; dans le cas des drosophiles (SOLIGNAC *et al.*, 1984) cette hypothèse n'est manifestement pas vérifiée.

### III. Variabilité intra-spécifique de l'ADNmt et génétique des populations

#### A. Ampleur de la variation intra-spécifique

On a maintenant une connaissance assez détaillée de la variabilité de l'ADNmt de plusieurs espèces de mammifères. Il ressort de telles études qu'il existe dans la nature un fort polymorphisme de ce génome. Chez la souris domestique (*M. m. domesticus*), FERRIS *et al.* (1983 b), sur 113 animaux d'origines géographiques variées, trouvent un minimum de 31 types mitochondriaux différents. AVISE *et al.* (1979 a), analysant 87 rongeurs américains (*Geomys pinetis*), révèlent un minimum de 23 génomes cytoplasmiques alors que, dans l'espèce *Peromyscus maniculatus*, LANSMAN *et al.* (1983 a) en trouvent 61 sur 135 animaux analysés.

Ces données, basées sur des analyses de PLFR (polymorphisme de longueur de fragments de restriction) peuvent permettre une quantification des différences obtenues. Le simple décompte du nombre d'allèles observés constitue une mesure peu utilisable de la diversité puisqu'elle est fortement dépendante à la fois de la taille de l'échantillon et de la finesse de la méthode de détection des variants (nombre et nature des endonucléases utilisées). La mesure la plus simple de la similarité entre 2 individus est donnée par la proportion de fragments en commun dans les 2 profils de digestion obtenus :

$$F = 2N_{xy}/(N_x + N_y)$$

où  $N_x$  et  $N_y$  sont les nombres de fragments trouvés respectivement chez les individus x et y, et  $N_{xy}$  le nombre de fragments communs aux deux.

On peut aussi quantifier la différence entre 2 variants par la proportion de nucléotides qui diffèrent entre leurs séquences. Moyennant certaines hypothèses, il est possible d'estimer ce pourcentage de substitutions nucléotidiques (que nous désignerons par p dans la suite) entre 2 génomes caractérisés par analyse de PLFR. Des formules reliant la proportion de fragments communs (F) ou le nombre de sites d'endonucléases communs et p ont pu être établies (NEI & LI, 1979 ; GOTOH *et al.*, 1979).

Les valeurs de p obtenues pour diverses espèces sont de l'ordre de 1 à quelques p. 100. Les exemples les mieux documentés sont la souris (*M. m. domesticus*)  $p = 0,77$  p. 100 (FERRIS *et al.*, 1983 b), *R. rattus* de 0,4 à 9,6 p. 100 et *R. norvegicus* de 0,4 à 1,8 p. 100 (BROWN & SIMPSON, 1981), *Geomys pinetis*  $p = 2$  p. 100 (AVISE *et al.*, 1979 a), *Peromyscus maniculatus* de 3 à 5 p. 100 (LANSMAN *et al.*, 1983 a). Par contre, l'homme semble beaucoup moins polymorphe que les autres espèces de mammifères du point de vue de l'ADNmt ( $p = 0,36$  p. 100, BROWN, 1980).

Remarquons toutefois que la validité de ces estimations de pourcentage de substitutions n'est pas à toute épreuve. L'établissement des formules qui les permettent repose sur plusieurs hypothèses. (1) Le polymorphisme observé résulte uniquement de substitutions nucléotidiques. (2) L'ADNmt est assimilé à une séquence aléatoire vis-à-vis de la distribution des sites de coupure des endonucléases et des mutations qui les font apparaître ou disparaître. (3) Les homologies de fragments de restriction ou de sites de coupures sont correctement établies.

Nous avons vu dans le premier chapitre que la première hypothèse est vérifiée dans ses grandes lignes. On sait que la deuxième est fautive. Les études de PLFR ont

montré que les sites d'endonucléases qui révèlent des variations ne sont pas répartis au hasard sur la molécule mitochondriale (FERRIS *et al.*, 1983 b ; BROWN & SIMPSON, 1981 ; CANN *et al.*, 1984). ADAMS & ROTHMAN (1982) montrent que, pour beaucoup de séquences d'ADN, la distribution même de ces sites (indépendamment de leur variabilité) n'est pas aléatoire, et qu'elle peut varier suivant l'enzyme de restriction considérée. La troisième hypothèse est d'autant plus vérifiée que l'enzyme de restriction employée engendre peu de fragments, alors que la précision de l'estimation de  $p$  augmente avec le nombre de fragments analysés (NEI & LI, 1979).

La précision de ces estimations est vérifiable à partir des données de séquences nucléotidiques. AQUADRO & GREENBERG (1983) ont séquencé un fragment de 900 paires de bases de l'ADNmt humain chez 7 individus. Ils montrent que l'analyse par les enzymes de restriction aurait donné des valeurs de  $p$  très similaires aux valeurs réelles qu'ils déterminent. Toutefois, ils montrent aussi que ces valeurs sont très supérieures à celles obtenues par les analyses de restriction portant sur l'ensemble du chromosome mitochondrial. Ceci est dû à des différences de variabilité entre les différentes régions de ce génome. TAJIMA (1983) étudie les propriétés statistiques du nombre moyen de substitutions nucléotidiques entre 2 « nucléomorphes » pris au hasard. Il montre que le nombre de différences nucléotidiques est sujet à une variance stochastique très importante, même beaucoup plus importante que la variance d'échantillonnage.

Quoiqu'il en soit il reste vrai qu'il existe une forte hétérogénéité interindividuelle de l'ADNmt. Peut-on expliquer ce fait par le simple mode de transmission de ce génome, et par la modification subséquente de paramètres populationnels tels que la taille efficace, par rapport à ce qui vaut pour le génome nucléaire ? Si on suppose que la transmission est strictement maternelle, qu'il y a évolution clonale, que la ségrégation des mitochondries est rapide dans la lignée germinale et que le sex-ratio est de 1, le nombre effectif de génomes cytoplasmiques peut alors être valablement approché par  $N_f$ , nombre de femelles, et est donc 4 fois plus faible que pour le génome nucléaire. Le temps de fixation d'un nouveau mutant sera donc 2 fois plus court que pour un allèle d'un gène nucléaire mais la diversité sera à peu près 4 fois moins grande, ceci dans le cadre d'un modèle neutraliste à nombre infini d'allèles (BIRKY *et al.*, 1983). D'après ce que nous avons vu ci-dessus, cela ne semble pas correspondre à ce qu'on observe dans les populations naturelles. Bien que la modification de certains paramètres (en particulier le sex-ratio et la vitesse de ségrégation des mitochondries dans la lignée germinale) puisse largement influencer, et même inverser ce rapport de diversité entre les 2 génomes, il est plus plausible que la diversité de l'ADNmt résulte en grande partie d'un fort taux de mutation. Nous avons vu que cela semble effectivement être le cas.

### B. Utilisation en Biologie des Populations

L'étude de la variabilité intraspécifique de l'ADNmt est informative tant sur le plan qualitatif que quantitatif. Se basant sur la relativement faible hétérogénéité mitochondriale de l'espèce humaine (voir plus haut), BROWN (1980) émet l'hypothèse que toutes les lignées maternelles humaines actuelles seraient issues d'un ancêtre commun il y a seulement 180 à 360 000 ans. Ceci pourrait signifier, comme le propose l'auteur, que l'espèce a connu à ce moment une forte diminution d'effectif et qu'une seule lignée maternelle a survécu. Toutefois, selon AVISE *et al.* (1984) qui ont effectué des simulations numériques, le processus stochastique d'extinction des lignées maternelles pourrait à lui seul expliquer ce résultat. Une autre hypothèse pourrait être que l'évolution de ce génome est plus lente chez l'homme que chez les autres espèces

analysées (le temps de génération très long pourrait en être une raison. Voir par exemple WU & LI, 1985 pour une discussion de ce problème).

Du point de vue de l'information qualitative, nous donnons quelques exemples de résultats, encore peu nombreux dans ce domaine et proposons quelques interprétations. AVISE *et al.* (1979 b) puis LANSMAN *et al.* (1983 a) ont étudié les relations matriarcales entre 135 individus de l'espèce *Peromyscus maniculatus* échantillonnés sur l'ensemble de l'aire de répartition de l'espèce. Ils montrent une forte structuration géographique du polymorphisme mitochondrial et mettent en évidence 5 assemblages géographiques. Il existe très peu de corrélations avec les variations géographiques d'autres marqueurs morphologiques ou chromosomiques. De plus le polymorphisme des gènes de structure nucléaires, révélé par électrophorèse protéique, s'avère géographiquement homogène.

Quelle est l'origine de telles différences entre les génomes nucléaire et mitochondrial du point de vue des modalités de la variation géographique ? L'ADNmt ne nous renseigne que sur les échanges génétiques par l'intermédiaire des femelles, alors que les flux génétiques nucléaires passent par les 2 sexes. Chez *Peromyscus*, il a été rapporté une plus grande propension des mâles à la migration (LANSMAN *et al.*, 1983 a). Ceci pourrait évidemment expliquer la plus forte structuration géographique du polymorphisme mitochondrial.

Remarquons qu'il peut exister une différence de nature entre les informations concernant ces 2 génomes eu égard aux manières différentes par lesquelles on y accède (PLFR pour l'ADNmt, électrophorèse protéique pour le génome nucléaire) : le nombre de variants protéiques détectés dans une espèce par les techniques biochimiques classiques est généralement faible et l'évolution ainsi visible est lente. Par contre, nous avons vu qu'il en est autrement pour l'ADNmt, à la fois à cause des méthodes d'analyse plus fines employées et du fort taux de mutation de ce génome. On a donc accès à des échelles de temps différentes avec ces 2 types de marqueurs.

Ainsi l'ADNmt, qui évolue vite, peut nous renseigner sur des événements relativement récents : chez la souris domestique (*M. m. domesticus*) largement étudiée sous cet angle par FERRIS *et al.* (1983 b), il n'apparaît aucune structuration géographique du polymorphisme mitochondrial, malgré la large répartition de cette espèce. En d'autres termes il n'y a aucune liaison entre le génotype mitochondrial d'un individu et son origine géographique. Ceci pourrait être mis en rapport avec l'expansion rapide de cette espèce commensale véhiculée par l'homme au moment du développement de l'agriculture en Europe. Par contre chez l'espèce de souris méditerranéenne *M. spretus*, l'analyse de la variabilité mitochondriale (BOURSOT *et al.*, 1985) met en évidence 2 ensembles géographiques phylogénétiquement distincts (voir le chapitre suivant concernant le problème des reconstitutions phylogénétiques). Etant donné que l'expansion de cette espèce dans son aire de répartition actuelle date aussi du néolithique, les auteurs sont amenés à supposer que le peuplement de ces 2 blocs géographiques s'est fait à partir de 2 groupes déjà différenciés.

Pour le marqueur mitochondrial comme pour les autres, on peut se poser la question d'une éventuelle signification adaptative des variations observées. On connaît très peu de cas de variations phénotypiques liées à l'ADNmt (voir FISCHER LINDAHL, 1985, pour une revue chez la souris). Certains arguments pourraient laisser penser que la différenciation fonctionnelle des variants observés soit faible. Si une telle différenciation existait au sein d'une espèce, on pourrait s'attendre à ce qu'elle existe *a fortiori* entre 2 espèces ou sous-espèces. Nous avons cité l'expérience de GYLLENSTEN *et al.* (1985) impliquant la réalisation de croisements en retour répétés d'hybrides entre 2

espèces de souris (*M. m. domesticus* et *M. spretus*) par l'espèce paternelle. On obtient ainsi des individus possédant le génotype nucléaire d'une espèce et le génotype mitochondrial de l'autre. Ils semblent parfaitement viables et fertiles, quel que soit le sens du croisement.

De tels exemples existent dans la nature. L'étude génétique de l'hybridation naturelle entre les 2 semi-espèces européennes de souris (*M. m. domesticus* et *M. m. musculus*) a montré que l'introgression des allèles d'un groupe dans l'autre semblait plus importante pour le génome mitochondrial que pour le génome nucléaire, à tel point qu'on pouvait trouver des populations possédant le génome nucléaire pur d'une des semi-espèces mais ayant totalement fixé le cytoplasme de l'autre (FERRIS *et al.*, 1983 a ; BOURSOT *et al.*, 1984). Entre 2 espèces de drosophile, POWELL (1983) met en évidence des échanges génétiques cytoplasmiques alors que les génomes nucléaires ne révèlent aucune trace d'hybridation.

Deux caractéristiques principales font de l'ADNmt un bon outil de biologie des populations (voir aussi AVISE & LANSMAN, 1983, pour une revue) : sa relative facilité d'étude (par des techniques d'analyse de PLFR, en raison de la facilité de purification et de la multiplicité du nombre de copies) et sa forte variabilité intraspécifique. Il est même possible de mettre en évidence une variabilité intrapopulationnelle, et des analyses plus fines de la variabilité pourraient donner accès à des estimations du nombre minimum de lignées femelles constitutives d'une population, ainsi qu'à une appréciation assez fine de la structure génétique des peuplements naturels et de l'importance des échanges génétiques entre populations (TAKAHATA & PALUMBI, 1985).

Par ailleurs, l'évolution clonale de l'ADNmt et le type de méthodes d'analyse de la variabilité employées autorisent une reconstitution de la généalogie femelle au sein d'une espèce et entre populations. Nous verrons plus bas les techniques possibles de reconstitution de ces phylogénies. On a ainsi accès à une information historique récente sur l'espèce et ses sous-ensembles.

#### IV. ADNmt, phylogénie et spéciation

Quelles sont les méthodes de reconstitution phylogénétique applicables aux données mitochondriales et quels sont les premiers résultats obtenus ?

##### A. Divergence entre taxons, divergence entre morphes mitochondriaux

Pour les comparaisons entre groupes peu distants, la grande diversité des types mitochondriaux observés en général dans les populations naturelles proscrit l'emploi de méthodes visant à déterminer une divergence globale entre taxons. TAKAHATA & NEI (sous presse) montrent qu'une telle estimation (par exemple en utilisant une estimation du nombre net de substitutions nucléotidiques entre taxons, NEI & LI, 1979) entre taxons peu distants est sujette à une variance très importante, qui ne peut être réduite qu'en augmentant considérablement la taille de l'échantillon. Or la nature des données obtenues sur chaque lignée mitochondriale (cartographie de sites de restriction, séquençage) permet le plus souvent de quantifier et de qualifier la divergence entre 2 clones mitochondriaux donnés : il est donc possible, et souhaitable, de reconstituer les phylogénies des molécules avant d'estimer celle des espèces qui les portent.

### B. Méthodes d'inférences phylogénétiques

Plusieurs types de méthodes s'offrent à nous, parmi lesquelles certaines sont certainement mieux adaptées que d'autres aux données mitochondriales. Nous esquissons ci-dessous la présentation de quelques-unes d'entre elles. Une discussion générale de la plupart de ces méthodes est proposée par FELSENSTEIN (1982).

Les plus simples d'entre elles, purement phénétiques, transforment les données brutes par le calcul d'indices de divergence moléculaire, par exemple le pourcentage de substitutions nucléotidiques. Les matrices de distances ainsi obtenues sont ensuite le plus souvent transformées en dendrogrammes par l'un des nombreux algorithmes d'agglomération disponibles (par exemple voir LANCE & WILLIAMS ou SOKAL & SNEATH in LEGENDRE & LEGENDRE, 1981). A l'évidence, il ne faut pas attendre trop de cette approche phénétique pour les données mitochondriales, principalement pour 2 raisons :

— Le calcul d'indices de distance représente nécessairement une perte d'information par rapport aux données brutes. La variance de  $p$  (divergence nucléotidique, que nous avons définie au chapitre précédent) tel qu'il est estimé à partir d'un nombre forcément limité de sites de restriction ou de nucléotides peut être très importante, comme nous l'avons vu plus haut, et rendre ce type de reconstitution statistiquement non fondé.

— Les techniques d'agglomération supposent l'homogénéité des taux d'évolution pour que les phénogrammes obtenus puissent être assimilés à des phylogénies. Néanmoins, certaines techniques de construction d'arbres permettent de s'affranchir de cette dernière hypothèse. Ainsi par exemple les algorithmes proposés par FARRIS (1972) pour la construction d'arbres de longueur minimum (arbres de Wagner) et discutés par SWOFFORD (1981). D'une manière plus générale, l'opportunité et la valeur des méthodes de reconstitution phylogénétiques basées sur l'analyse de matrices de distances est discutée par FARRIS (1981) et FELSENSTEIN (1984). De telles méthodes ont été utilisées pour les données mitochondriales par plusieurs auteurs (AVISE *et al.*, 1979 a ; BROWN & SIMPSON, 1981 ; YONEKAWA *et al.*, 1981, 1982 ; FERRIS *et al.*, 1983 b).

Des méthodes plus performantes, prenant en compte les données de façon qualitative et non plus seulement quantitative, ont pu être appliquées avec succès dans quelques cas réels. Quoique non phénétiques, ces méthodes ne sont pas à proprement parler strictement cladistiques : des méthodes purement cladistiques supposeraient en effet de travailler avec des caractères unitaires (c'est-à-dire ici des positions nucléotidiques ou des sites de coupure pris un à un) pour lesquels on puisse reconnaître des formes dérivées et des formes ancestrales. Ne disposant pas d'information sur les séquences des fossiles, la seule façon de déterminer si un site de coupure ou un nucléotide représente l'état ancestral du caractère est de montrer qu'il est présent à la fois dans le groupe d'espèces ou de populations étudiées et dans un groupe frère dont on sait par ailleurs qu'il a divergé de façon plus ancienne. Cette méthode, si elle s'avère puissante pour des caractères ayant une certaine durée de vie et une certaine stabilité évolutive permettant d'en assurer l'homologie avec sécurité dans 2 groupes frères, se montre en général d'emploi très délicat pour les données obtenues sur les acides nucléiques, à cause du petit nombre d'états de chaque caractère (4 pour les nucléotides, 2 pour les sites de restriction) et de la grande réversibilité du processus mutationnel sous-jacent. Ceci est vrai *a fortiori* pour les ADN mitochondriaux, molécules qui, nous l'avons vu, évoluent assez rapidement.

Etant donné donc que nous n'avons pas en général *a priori* de moyen de déterminer les séquences ancestrales, les méthodes présentées ci-dessous permettent de proposer une reconstitution phylogénétique moyennant un surcroît d'hypothèses. En outre, certaines d'entre elles (principes de parcimonie et de compatibilité notamment) permettent de proposer *a posteriori* une reconstitution des assemblages ancestraux de caractères unitaires (séquences ou cartes de restriction).

— *Parcimonie* : On postule que la phylogénie la plus probable est celle qui fait intervenir au total le moins d'événements mutationnels le long des branches de l'arbre phylétique. Ces techniques ont déjà été largement utilisées pour les reconstitutions phylogénétiques à partir de séquences protéiques, et il en existe de nombreuses versions adaptables aux séquences d'ADN ou aux cartes de restrictions. Pour que les résultats obtenus par ces techniques soient consistants (c'est-à-dire qu'ils convergent vers la solution vraie quand la masse des données augmente), il faut au premier chef que les taux d'évolution dans les différentes lignées soient suffisamment faibles pour que l'occurrence d'un événement double dans une branche longue de l'arbre soit beaucoup moins probable que celle d'un événement simple dans une branche courte, c'est-à-dire en quelque sorte que l'homoplasie due à la redondance du processus mutationnel soit faible.

— *Compatibilité* : Introduit plus tardivement, ce type de méthode ne prend en compte que le plus grand sous-ensemble de caractères mutuellement compatibles, c'est-à-dire ceux pour lesquels il ne faut pas faire intervenir de phénomènes d'homoplasie (réversion, convergence, parallélisme) pour expliquer leur distribution conjointe sur l'ensemble des individus. Elle exclut les autres caractères comme étant précisément ceux pour lesquels l'homoplasie a brouillé les cartes, rendant ainsi une reconstitution phylétique hasardeuse. Les hypothèses sous-jacentes à l'emploi de ces méthodes ne diffèrent que peu de celles impliquées dans la parcimonie, et sont discutées en détail par FELSENSTEIN (1978, 1981 b).

— *Maximum de vraisemblance* : Quoique le principe de cette méthode soit connu depuis longtemps en taxinomie numérique, son application est récente à cause des temps de calculs souvent très longs qu'elle requiert. Cette famille de techniques choisit l'arbre phylogénétique qui, si l'on simulait le long de ses branches l'évolution à chaque site selon un modèle stochastique simple, donnerait la distribution de caractères observés comme étant la plus probable. Le fait que les taux de substitution soient supposés constants d'un site à l'autre est l'hypothèse la plus contraignante spécifique à ce modèle paramétrique. Aucune hypothèse n'est cependant faite sur la vitesse d'évolution dans chaque lignée, car c'est la quantité d'évolution qui est estimée dans chaque branche et non pas le temps de divergence. Cette technique a été développée principalement par J. FELSENSTEIN qui en présente une version pour séquences d'acides nucléiques (FELSENSTEIN, 1981 a) et appliquée par FORT *et al.* (1984).

Quelle confiance peut-on accorder à une phylogénie obtenue par l'une ou l'autre des méthodes proposées ci-dessus ? Quelles sont les précautions particulières à prendre dans le cas de l'ADN mitochondrial ?

D'une manière générale, toutes ces techniques supposent l'indépendance de chaque site (nucléotidique ou site de coupure d'endonucléase) vis-à-vis de la mutation. Comme nous l'avons vu plus haut pour les sites d'endonucléases (ADAM & ROTHMANS, 1982), cela n'est sûrement pas toujours vérifié. *A fortiori* pour les nucléotides pris un par un, les interactions fonctionnelles entre certains sites adjacents ou voisins et les contraintes

sur la composition globale de certaines parties de la molécule interdisent de considérer une indépendance totale des sites.

Une autre difficulté qui peut se faire jour est la grande quantité d'homoplasie que l'on peut attendre avec des molécules évoluant vite comme l'ADNmt. L'homoplasie, c'est la dégénérescence de l'information, le « bruit » évolutif qu'aucune méthode ne saurait éliminer totalement. Si les taux de substitution à chaque site sont suffisamment élevés, alors que les événements multiples à un site donné dans 2 lignées indépendantes (parallélisme et convergence) ou dans une même lignée (réversion) peuvent ne pas être rares. Comme nous l'avons vu au chapitre I, la forte prédominance des substitutions de type transition par rapport à celles de type transversion renforce encore ce risque d'événements multiples. Pour les sites d'endonucléases en outre, il existe plusieurs manières de produire des disparitions multiples d'un même site dans plusieurs lignées indépendantes (par mutation de l'un des 4 à 6 nucléotides reconnus par l'enzyme), alors qu'il n'existe qu'une manière de faire apparaître un site donné dans 2 lignées indépendantes. On doit donc s'attendre à ce que l'homoplasie soit plus forte pour les disparitions que pour les apparitions de sites.

Pour les données de PLFR, TEMPLETON (1983) essaie de tenir compte des 2 écueils mentionnés ci-dessus. Cet auteur propose d'appliquer une analyse de parcimonie aux résultats obtenus enzyme par enzyme, et de tester ensuite la compatibilité des différents arbres obtenus. Les endonucléases fournissant des résultats déviants par rapport à la phylogénie majoritaire sont exclues de l'analyse. Cette technique permet d'une part de ne pas mélanger dans une analyse globale des sites ayant des probabilités de mutation très différentes, et d'autre part, d'éliminer les ensembles de caractères possédant l'information la plus dégénérée. TEMPLETON (1983) applique cette méthode aux données de FERRIS *et al.* (1981) sur l'ADNmt des grands singes et de l'homme, et propose un test statistique non paramétrique pour apprécier l'adéquation de la phylogénie obtenue par rapport à une autre supposée. Selon ses résultats, l'hypothèse phylogénétique suivant laquelle les ADNmt du chimpanzé et du gorille auraient divergé postérieurement à la lignée mitochondriale humaine est acceptable au risque de 5 p. 100. Le paragraphe que nous allons maintenant développer incite à la prudence pour extrapoler ces données mitochondriales à la phylogénie des espèces.

### C. Phylogénies mitochondriales

#### 1. Généalogie matriarcale intraspécifique

Dans un petit nombre de cas documentés sur un nombre suffisamment grand d'individus, l'emploi des techniques décrites ci-dessus a permis de reconstituer la généalogie des différents clones ségrégeant dans une population, ce aussi bien pour des données de PLFR (FERRIS *et al.*, 1983 b ; LANSMAN *et al.*, 1983 a ; AVISE *et al.*, 1983) que de séquençage (AQUADRO & GREENBERG, 1983 ; FORT *et al.*, 1984). Les dendrogrammes ainsi produits montrent que l'on peut trouver simultanément dans une même population des clones mitochondriaux dont certains peuvent représenter des stades évolutifs plus anciens que d'autres. Ainsi, un morphe ancestral et ses descendants situés dans plusieurs étapes plus loin dans le chemin mutationnel qui les relie peuvent coexister (au passage, ceci proscrit pour de telles reconstitutions l'emploi de techniques phénétiques supposant que chaque objet analysé est un bourgeon terminal du dendrogramme). Pour l'ensemble des clones présents, on peut reconstituer la séquence

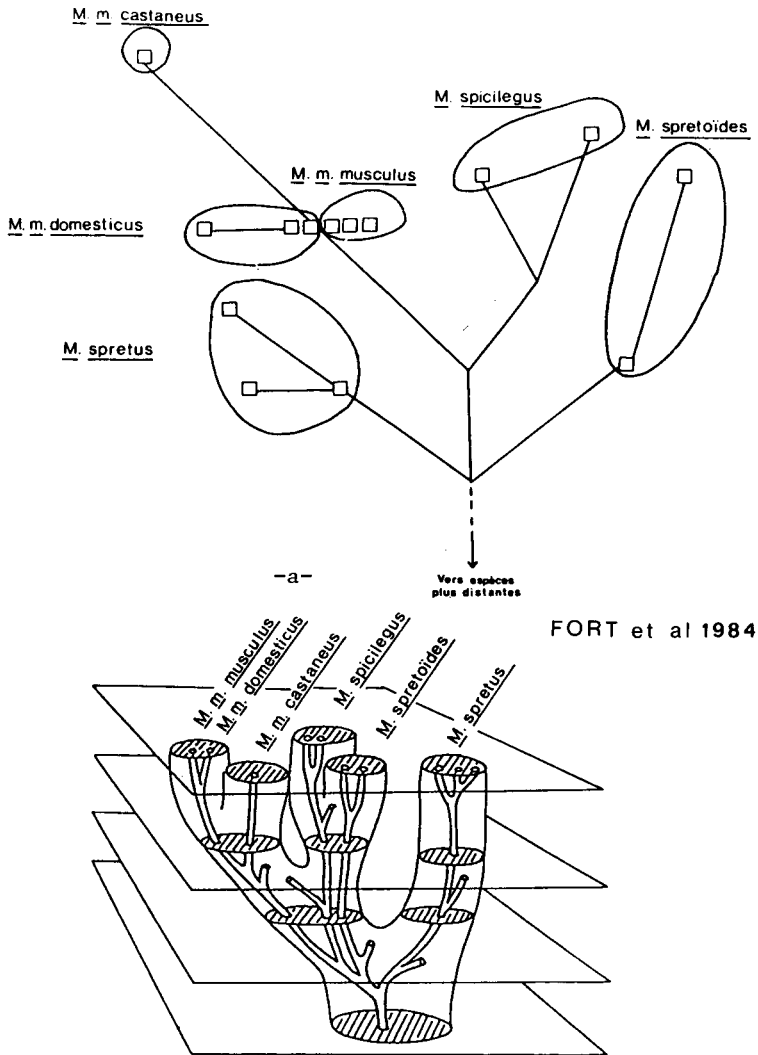


FIG. 3

Variabilité de séquence des 110 nucléotides terminaux de l'extrémité 5' de l'ARNr 16S mitochondrial chez 6 espèces et semi-espèces de souris (*Mus*) (d'après FORT et al., 1984).

3a : Arbre mutationnel minimal représentant la généalogie des 14 clones mitochondriaux.

Les longueurs des branches sont proportionnelles au nombre de mutations.

3b : Modèle représentant les phylogénies emboîtées des clones mitochondriaux

(ramifications internes) et du génome nucléaire (branches principales de l'arbre, obtenues par l'analyse électrophorétique de protéines codées par 42 locus chromosomiques).

Sequence variability of the 110 nucleotides of the 5' terminus of the 16S mitochondrial rRNA in 6 species and semi-species of the genus *Mus* (from FORT et al., 1984).

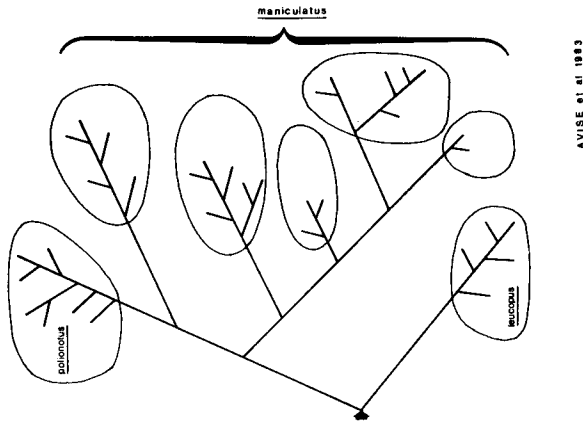
3a : Minimal mutation tree representing the genealogy of the 14 mitochondrial clones.

Branch lengths are proportional to the number of mutations.

3b : Model representing the embedding of the phylogenies of mitochondria (inner tree) and of the nuclear genomes (outer tree, obtained by protein electrophoresis at 42 chromosome loci).

Il apparaît ainsi que *Mus spicilegus* et *Mus spretoides* sont les 2 espèces les plus proches que l'on connaisse du point de vue de leurs gènes de structure, alors que leurs ADN mitochondriaux montrent une divergence fort ancienne.

It appears that *Mus spicilegus* and *Mus spretoides* are the closest species we know from the viewpoint of structural genes, while their mitochondrial DNAs show a very ancient divergence.



AVISE et al. 1983

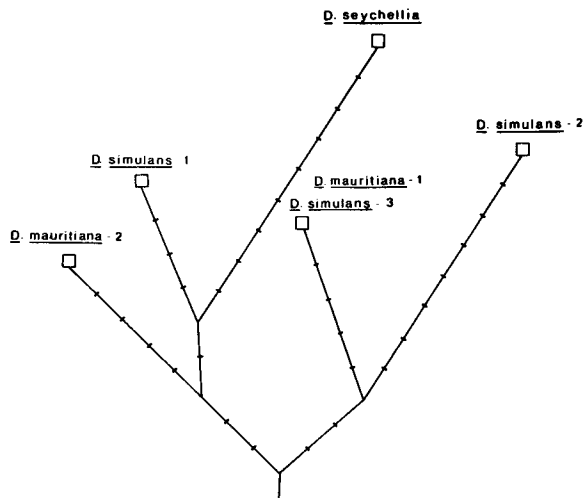
FIG. 4

Reconstitution de la phylogénie mitochondriale chez 3 espèces de *Peromyscus* (d'après AVISE et al., 1983), illustrant un cas de paraphylétisme.

Reconstitution of mitochondrial phylogeny among 3 species of *Peromyscus* (from AVISE et al., 1983), illustrating a case of paraphyletism.

Ici, *P. maniculatus* constitue un groupe paraphylétique par rapport à *polionotus*, du moins en ce qui concerne les généalogies mitochondriales.

Here, *P. maniculatus* appears as a paraphyletic group with *polionotus*, at least in respect of mitochondrial genealogy.



SOLIGNAC, pers com

FIG. 5

Dendrogramme montrant les relations phylogénétiques existant entre 6 ADN mitochondriaux de 3 espèces de *Drosophila*, analysés par 12 enzymes de restriction (arbre recalculé d'après les données de SOLIGNAC, communication personnelle).

Dendrogram showing phylogenetic relationships existing between 6 mitochondrial DNA's in 3 species of *Drosophila*, analysed with 12 restriction enzymes (tree was recalculated from the data of SOLIGNAC, personal communication).

Les tirets le long des branches indiquent les événements de gain ou de perte d'un fragment de restriction.

The segments along the branches represent the events of gain or loss of a restriction site.

ancestrale commune, qu'elle soit ou non encore présente dans la population. En supposant que l'ADNmt évolue sous le seul effet de la mutation et de la dérive, le temps nécessaire à l'élaboration des morphes les plus évolués à partir du morphe ancestral est directement lié à l'histoire démographique de l'espèce au cours de ce laps de temps, AVISE *et al.* (1984) ont modélisé la probabilité de survie de lignées matriarcales indépendantes en fonction de ces paramètres démographiques et montré que, en l'absence de nouvelles mutations, la fixation d'un morphe unique était hautement probable au bout d'un nombre de générations supérieur ou égal à 4 fois le nombre de femelles dans des populations d'effectif stable. Dans des populations en expansion, par contre, ce processus de fixation est considérablement ralenti.

Ainsi donc la rétention de polymorphismes mitochondriaux anciens apparaît comme fortement conditionnée par les événements démographiques connus par l'espèce, l'âge du morphe ancestral (si tant est qu'on puisse l'estimer par la mesure du taux de substitution nucléotidique) donnant une idée de la durée minimale pendant laquelle la population n'a pas subi de goulot d'étranglement majeur.

## 2. *Phylogénies interspécifiques*

Comment les phénomènes de spéciation (terme pris ici au sens de l'acquisition de l'isolement reproductif, notion s'appliquant donc strictement à des populations d'organismes sexués) affectent-ils l'évolution d'éléments génétiques non recombinants à reproduction clonale, et par conséquent strictement isolés les uns des autres tels que l'ADNmt ? En retour, en quoi les généalogies mitochondriales nous renseignent-elles sur la phylogénie des espèces, qu'elles soient proches ou éloignées ?

Le paragraphe précédent fournit tout de suite des éléments de réponse à ces questions : dans la mesure où il est fréquent d'avoir rétention de polymorphisme, les ADN mitochondriaux, qui ne semblent pas recombiner entre eux, gardent la trace d'événements mutationnels anciens mis en évidence par la reconstitution du morphe ancestral de chaque groupe. Par quelque mécanisme qu'elle se réalise, la partition d'une espèce polymorphe va engendrer 2 ensembles dont les généalogies matriarcales peuvent être alors reliées par un ancêtre commun remontant à une date bien plus ancienne que l'acquisition de l'isolement reproducteur lui-même. FORT *et al.*, (1984) obtiennent dans le genre *Mus* des résultats suggérant un tel phénomène et proposent un modèle rendant compte de l'antériorité de la divergence des lignées mitochondriales par rapport à celle des espèces qui les portent (fig. 3). AVISE *et al.*, (1983) précisent les cas théoriques qui peuvent se présenter pour l'emboîtement des phylogénies mitochondriales et nucléaires et donnent un exemple de paraphylétisme dans le groupe *Peromyscus* (fig. 4). Chez les drosophiles du groupe *melanogaster*, SOLIGNAC (communication personnelle) montrent une intrication presque complète des généalogies matriarcales pour 3 espèces (fig. 5). En figure 6 nous présentons un schéma illustrant les relations possibles entre phylogénies spécifique et mitochondriale.

Les conséquences pratiques du décalage entre l'âge du morphe ancestral et le début de l'isolement reproducteur sont que la vitesse d'évolution estimée pour l'ADNmt à partir d'espèces proches dont la divergence est connue par d'autres méthodes (paléontologiques par exemple) peut être surestimée, ou réciproquement que les temps de divergence estimés à partir de données mitochondriales peuvent être supérieurs à la réalité.

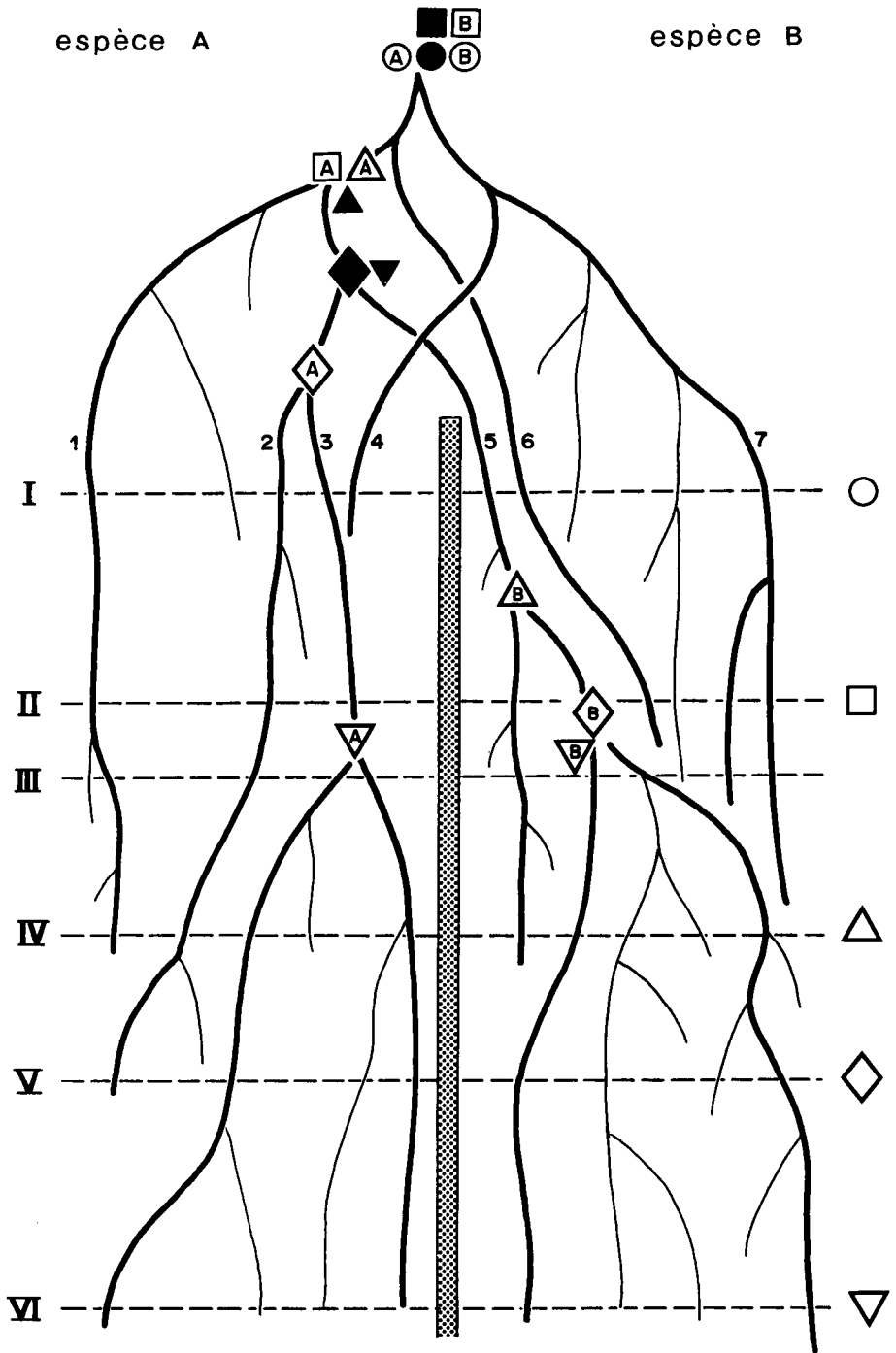


FIG. 6

*Schéma synthétique illustrant les différents types de relations phylogénétiques possibles entre les clones mitochondriaux de 2 espèces A et B.*

*Synthetic scheme illustrating different possible phylogenetic relationships between the mitochondrial clones of 2 species A and B.*

L'arbre représente la phylogénie des lignées mitochondriales. Les lignées principales sont numérotées de 1 à 7. La barre verticale symbolise l'isolement reproducteur entre les 2 espèces. Nous allons suivre l'évolution au cours du temps, à 6 stades numérotés de I à VI et correspondant chacun à un des symboles figurés à droite. A chaque stade nous avons porté sur l'arbre : (1) avec un symbole noir, le Plus Jeune Ancêtre Commun (P.J.A.C.) à toutes les lignées existant à cet instant ; (2) avec 2 symboles blancs, les P.J.A.C. à chacune des 2 espèces A et B.

*Stade I :* Le P.J.A.C. à l'ensemble de toutes les lignées matriarcales des 2 groupes est confondu avec le P.J.A.C. de chaque espèce. Chacune des 2 espèces peut alors être considérée comme un groupe paraphylétique pour ce qui est de ses lignées mitochondriales, puisque toute la descendance de chaque P.J.A.C. n'est pas incluse dans un seul groupe.

*Stade II :* Le P.J.A.C. à l'espèce A s'est rajeuni, mais pas celui de l'espèce B, qui forme un groupe paraphylétique relativement à A, qui apparaît comme un rameau divergent dont la phylogénie mitochondriale est incluse dans celle de B.

*Stade III :* Aucun des P.J.A.C. n'a bougé, mais la disparition de la lignée 6 chez B peut faire considérer cette espèce comme un groupe polyphylétique constitué de la lignée 7 et de la lignée 5, qui semble alors provenir de l'ascendance de A.

*Stade IV :* Par perte de la lignée 7, le P.J.A.C. de B s'est considérablement rajeuni. C'est A qui apparaît maintenant comme paraphylétique, et B monophylétique.

*Stade V :* Les lignées matriarcales des 2 espèces sont maintenant bien séparées. Le P.J.A.C. à A est antérieur à la divergence des espèces alors que celui de B est postérieur.

*Stade VI :* Le processus de rajeunissement s'est poursuivi et les 2 P.J.A.C. sont maintenant d'âge postérieur à la divergence.

The tree represents the phylogeny of the mitochondrial lines. The principal lines are numbered from 1 to 7. The vertical bar symbolises reproductive isolation between the 2 groups. We are going to follow evolution with time at 6 steps numbered I to VI, and each corresponding to a particular symbol figuring on the right. For each of these steps we have placed on the tree : (1) a closed symbol indicating the Youngest Common Ancestor (Y.C.A.) of all lines existing at this time ; (2) with 2 empty symbols, the Y.C.A.'s of each of the 2 species A and B.

*Step I :* The Y.C.A. of all maternal lines of the 2 groups is identical to the Y.C.A. of each species. Each of the 2 species can thus be considered as a paraphyletic group in respect of maternal lineages, as the descendants of the Y.C.A. of each species is not totally included in that species.

*Step II :* The Y.C.A. of species A has become younger, but not that of species B, which thus forms a paraphyletic group when compared to A, which appears as a divergent branch whose phylogeny is included in that of B.

*Step III :* None of the Y.C.A.'s has moved, but the extinction of line 6 in B could make us consider this species as a polyphyletic group made up of line 7, and line 5 which seems to be derived from A.

*Step IV :* After extinction of line 7, the Y.C.A. of B has become much younger. A now appears as a paraphyletic group relatively to B, which is now monophyletic.

*Step V :* Maternal lineages of the 2 species are now well separated. The Y.C.A. of A is older than the divergence between the species, while that of B is younger.

*Step VI :* The rejuvenation process has gone on, and the Y.C.A.'s of the 2 species are now younger than the species.

Enfin, un dernier type d'événement peut être la cause de disharmonies importantes entre phylogénies spécifiques et généalogies mitochondriales : les échanges de matériel cytoplasmique entre semi-espèces ou même entre espèces proches semblent être un phénomène relativement fréquent, puisque plusieurs cas ont déjà été attestés dans divers groupes zoologiques, comme nous l'avons signalé au chapitre précédent. De tels phénomènes peuvent également entrer en ligne de compte pour expliquer *pro parte* les résultats proposés en figures 3 et 5.

En fin de compte l'extrapolation de phylogénies entre espèces proches à partir de l'analyse du génome mitochondrial de quelques individus s'avère être une chose délicate. Quoiqu'il en soit, quand cette phylogénie est connue par d'autres moyens, les disharmonies éventuelles sont de précieux indices d'événements importants ayant affecté l'histoire antérieure des espèces.

Reçu le 6 mars 1985.

Accepté le 23 août 1985.

### Références bibliographiques

- ADAMS J., ROTHMAN E.D., 1982. Estimation of phylogenetic relationship from DNA restriction patterns and selection of endonuclease cleavage sites. *Proc. Natl. Acad. Sci. (U.S.A.)*, **79**, 3560-3564.
- ANDERSON S., BANKIER A.T., BARRELL B.G., DE BRUIJN M.H.L., COULSON A.R., DROUIN J., EPERON I.C., NIERLICH D.P., ROE B.A., SANGER F., SCHREIER P.H., SMITH A.J.H., STADEN R., YOUNG I.G., 1981. Sequence and organization of the human mitochondrial genome. *Nature*, **290**, 457-465.
- ANDERSON S., DE BRUIJN M.H.L., COULSON A.R., EPERON I.C., SANGER F., YOUNG I.G., 1982. Complete sequence of bovine mitochondrial DNA. *J. Mol. Biol.*, **156**, 683-717.
- AQUADRO C.F., GREENBERG B.D., 1983. Human mitochondrial DNA variations and evolution : analysis of nucleotide sequences from seven individuals. *Genetics*, **103**, 287-312.
- AVISE J.C., GIBLIN-DAVIDSON C., LAERM J., PATON J.C., LANSMAN R.A., 1979 a. Mitochondrial DNA clones and matriarchal phylogeny within and among geographic populations of pocket gopher, *Geomys pinetis*. *Proc. Natl. Acad. Sci. (U.S.A.)*, **76**, 6694-6698.
- AVISE J.C., LANSMAN R.A., SHADE R.O., 1979 b. The use of restriction endonucleases to measure mitochondrial DNA sequence relatedness in natural populations. I. Population structure and evolution in the genus *Peromyscus*. *Genetics*, **92**, 279-295.
- AVISE J.C., LANSMAN R.A., 1983. Polymorphism of mitochondrial DNA in populations of higher animals. In : NEI M., KOEHN R.K. (eds.), *Evolution of genes and proteins*, 147-164, Sinauer, Sunderland.
- AVISE J.C., NEIGEL J.E., ARNOLD J., 1984. Demographic influences on mitochondrial DNA lineage survivorship in animal populations. *J. Mol. Evol.*, **20**, 99-105.
- AVISE J.C., SHAPIRA J.F., DANIEL S.W., AQUADRO C.F., LANSMAN R.A., 1983. Mitochondrial DNA differentiation during the speciation process in *Peromyscus*. *Mol. Biol. Evol.*, **1**, 38-56.
- BARRELL B.G., ANDERSON S., BANKIER A.T., DE BRUIJN M.H.L., CHEN E., COULSON A.R., DROUIN J., EPERON I.C., NIERLICH D.P., ROE B.A., SANGER F., SCHREIER P.H., SMITH A.J.H., STADEN R., YOUNG I.G., 1980. Different pattern of codon recognition by mammalian mitochondrial tRNAs. *Proc. Natl. Acad. Sci. (U.S.A.)*, **77**, 3164-3166.
- BARRELL B.G., BANKIER A.T., DROUIN J., 1979. A different genetic code in human mitochondria. *Nature*, **282**, 189-194.

- BIBB M.J., VAN ETEN R.A., WRIGHT C.T., WALBERG M.W., CLAYTON D.A., 1981. Sequence and gene organization of mouse mitochondrial DNA. *Cell.*, **26**, 167-180.
- BIRKY C.W. Jr., 1983. Relaxed cellular controls and organelle heredity. *Science*, **222**, 468-475.
- BIRKY C.W., MARUYAMA T., FUERST P., 1983. An approach to population and evolutionary genetic theory for genes in mitochondria and chloroplasts, and some results. *Genetics*, **103**, 513-527.
- BOGENHAGEN D., CLAYTON D.A., 1974. The number of mitochondrial deoxyribonucleic acid genomes in Mouse L and Human HeLa cells. *J. Biol. Chem.*, **249**, 7991-7995.
- BOURSOT P., BONHOMME F., BRITTON-DAVIDIAN J., CATALAN J., YONEKAWA H., ORSINI P., GUERASIMOV S., THALER L., 1984. Introgression différentielle des génomes nucléaires et mitochondriaux chez deux semi-espèces de souris. *C.R. Acad. Sci.*, **299**, 365-370.
- BOURSOT P., JACQUART T., BONHOMME F., BRITTON-DAVIDIAN J., THALER L., 1985. Différenciation géographique du génome mitochondrial chez *M. spretus*. *C.R. Acad. Sci.*, **301**, 161-166.
- BROWN G.G., SIMPSON M.V., 1981. Intra- and interspecific variation of the mitochondrial genome in *Rattus Norvegicus* and *Rattus Rattus* : restriction enzyme analysis of variant mitochondrial DNA molecules and their evolutionary relationships. *Genetics*, **97**, 125-143.
- BROWN W.M., 1980. Polymorphism in mitochondrial DNA of humans as revealed by restriction endonuclease analysis. *Proc. Natl. Acad. Sci. (U.S.A.)*, **77**, 3605-3609.
- BROWN W.M., 1983. Evolution of animal mitochondrial DNA. In : NEI M., KOEHN R.K. (eds.), *Evolution of genes and proteins*, 62-88, Sinauer, Sunderland.
- BROWN W.M., GEORGE M., WILSON A.C., 1979. Rapid evolution of animal mitochondrial DNA. *Proc. Natl. Acad. Sci. (U.S.A.)*, **76**, 1967-1971.
- BROWN W.M., PRAGER E.M., WANG A., WILSON A.C., 1982. Mitochondrial DNA sequences of Primates : tempo and mode of evolution. *J. Mol. Evol.*, **18**, 225-239.
- CANN R.L., BROWN W.M., WILSON A.C., 1984. Polymorphic sites and the mechanism of evolution in human mitochondrial DNA. *Genetics*, **106**, 479-499.
- CANN R.L., WILSON A.C., 1983. Length mutations in human mitochondrial DNA. *Genetics*, **104**, 699-711.
- CHAPMAN R.W., STEPHENS J.C., LANSMAN R.A., AVISE J.C., 1982. Models of mitochondrial DNA transmission genetics and evolution in higher eucaryotes. *Genet. Res. Camb.*, **40**, 41-57.
- COOTE J.L., SZABADOS G., WORK T.S., 1979. The heterogeneity of mitochondrial DNA in different tissues from the same animal. *F.E.B.S. Letters*, **99**, 255-260.
- DAWID I.B., BLACKLER A.W., 1972. Maternal and cytoplasmic inheritance of mitochondrial DNA in *Xenopus*. *Develop. Biol.*, **29**, 152-161.
- DE BRUUN M.H.L., 1983. *Drosophila melanogaster* mitochondrial DNA : a novel organization and genetic code. *Nature*, **304**, 234-241.
- FARRIS J.S., 1972. Estimating phylogenetic trees from distance matrices. *Amer. Nat.*, **106**, 645-667.
- FARRIS J.S., 1981. Distance data in phylogenetic analysis. In : FUNK V.A. et BROOKS D.R. (ed.), *Advances in cladistics*, 3-23, Proc. First Meeting Willi Hennig Soc. N.Y. Botanical Garden, New York.
- FELSENSTEIN J., 1978. Cases in which parsimony or compatibility methods will be positively misleading. *Syst. Zool.*, **27**, 401-410.
- FELSENSTEIN J., 1981 a. Evolutionary trees from DNA sequences : a maximum likelihood approach. *J. Mol. Evol.*, **17**, 368-376.
- FELSENSTEIN J., 1981 b. A likelihood approach to character weighting and what it tells us about parsimony and compatibility. *Biol. J. Linn. Soc.*, **16**, 183-196.
- FELSENSTEIN J., 1982. Numerical methods for inferring evolutionary trees. *Quart. Rev. Biology*, **57**, 379-404.
- FELSENSTEIN J., 1984. Distance methods for inferring phylogenies : a justification. *Evolution*, **38**, 16-24.
- FERRIS S.D., SAGE R.D., HUANG C.M., NIELSEN J.T., RITTE U., WILSON A.C., 1983 a. Flow of mitochondrial DNA across a species boundary. *Proc. Natl. Acad. Sci. (U.S.A.)*, **80**, 2290-2294.
- FERRIS S.D., SAGE R.D., PRAGER E.M., RITTE U., WILSON A.C., 1983 b. Mitochondrial DNA evolution in Mice. *Genetics*, **105**, 681-721.

- FERRIS S.D., WILSON A.C., BROWN W.M., 1981. Evolutionary tree for apes and humans based on cleavage maps of mitochondrial DNA. *Proc. Natl. Acad. Sci. (U.S.A.)*, **78**, 2432-2436.
- FISCHER LINDAHL K., 1985. Mitochondrial inheritance in mice. *Trends in Genetics*, **1**, 135-139.
- FORT P., BONHOMME F., DARLU P., PIECHACZYK M., JEANTEUR P., THALER L., 1984. Clonal divergence of mitochondrial DNA versus populational evolution of nuclear genome. *Evol. Theory*, **7**, 81-90.
- FRANCISCO J.F., BROWN G.G., SIMPSON M.V., 1979. Further studies of types A and B rat mitochondrial DNA : cleavage maps and evidence for cytoplasmic inheritance in mammals. *Plasmid*, **2**, 426-436.
- GILES R.E., BLANC H., CANN H.M., DOUGLAS C.W., 1980. Maternal inheritance of human mitochondrial DNA. *Proc. Natl. Acad. Sci. (U.S.A.)*, **77**, 6715-6719.
- GOTOH O., HAYASHI J.I., YONEKAWA H., TAGASHIRA Y., 1979. An improved method for estimating sequence divergence between related DNAs from changes in restriction endonuclease cleavage sites. *J. Mol. Evol.*, **14**, 301-310.
- GREENBERG B.D., NEWBOLD J.E., SUGINO A., 1983. Intraspecific nucleotide sequence variability surrounding the origin of replication in human mitochondrial DNA. *Gene*, **21**, 33-49.
- GRIVELL L.A., 1983. Mitochondrial gene expression 1983. In : *Mitochondria 1983*, 25-45, Gruyter, Berlin.
- GYLLENSTEN U., WHARTON D., WILSON A.C., 1985. Maternal inheritance of mitochondrial DNA during backcrossing of two species of Mice. *J. of Heredity*, **76**, 321-324.
- HAUSWIRTH W.W., LAIPIS P.J., 1982. Mitochondrial DNA polymorphism in a maternal lineage of Holstein cows. *Proc. Natl. Acad. Sci. (U.S.A.)*, **79**, 4686-4690.
- HAUSWIRTH W.W., VAN DE WALLE M.J., LAIPIS P.J., OLIVO P.D., 1984. Heterogeneous mitochondrial DNA D-Loop sequences in bovine tissues. *Cell*, **37**, 1001-1007.
- HAYASHI J.I., YONEKAWA H., GOTOH O., WATANABE J., TAGASHIRA Y., 1978. Strictly maternal inheritance of rat mitochondrial DNA. *Biochem. Biophys. Res. Comm.*, **83**, 1032-1038.
- KUNKEL T.A., LOEB L.A., 1981. Fidelity of mammalian DNA polymerases. *Science*, **213**, 765-767.
- LANSMAN R.A., AVISE J.C., AQUADRO C.F., SHAPIRA J.F., DANIEL S.W., 1983 a. Extensive genetic variation in mitochondrial DNA's among geographic populations of the deer mouse, *Peromyscus maniculatus*. *Evolution*, **37**, 1-16.
- LANSMAN R.A., AVISE J.C., HUETTEL M.D., 1983 b. Critical experimental test of the possibility of « paternal leakage » of mitochondrial DNA. *Proc. Natl. Acad. Sci. (U.S.A.)*, **80**, 1969-1971.
- LEGENDRE L., LEGENDRE P., 1981. Ecologie numérique t. II : *La structure des données écologiques*, 247 pp., Masson, Paris.
- MICHAELS G.S., HAUSWIRTH W.W., LAIPIS P.J., 1982. Mitochondrial DNA copy number in Bovine oocytes and somatic cells. *Dev. Biol.*, **94**, 246-251.
- NEI M., LI W.H., 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc. Natl. Acad. Sci. (U.S.A.)*, **76**, 5269-5273.
- PIKO L., MATSUMOTO L., 1976. Number of mitochondria and some properties of mitochondrial DNA in the mouse egg. *Dev. Biol.*, **49**, 1-10.
- POTTER S.S., NEWBOLD J.E., HUTCHISON III C.A., EDGELL M.H., 1975. Specific cleavage analysis of mammalian mitochondrial DNA. *Proc. Natl. Acad. Sci. (U.S.A.)*, **72**, 4496-4500.
- POWELL J.R., 1983. Interspecific cytoplasmic gene flow in the absence of nuclear gene flow : evidence from *Drosophila*. *Proc. Natl. Acad. Sci. (U.S.A.)*, **80**, 492-495.
- SHAH D.M., LANGLEY C.H., 1979. Inter- and intraspecific variation in restriction maps of *Drosophila* mitochondrial DNA's. *Nature*, **281**, 696-699.
- SOLIGNAC M., GENERMONT J., MONNEROT M., MOUNOLOU J.C., 1984. Genetics of mitochondria in *Drosophila* : mtDNA inheritance in heteroplasmic strains of *D. mauritiana*. *Mol. Gen. Genet.*, **197**, 183-188.
- SOLIGNAC M., MONNEROT M., MOUNOLOU J.C., 1983. Mitochondrial DNA heteroplasmy in *Drosophila Mauritiana*. *Proc. Natl. Acad. Sci. (U.S.A.)*, **80**, 6942-6946.
- SWOFFORD D., 1981. On the utility of the distance Wagner procedure. In : FUNK V.A. & BROOKS D.R. (ed.), *Advances in cladistics*, 25-43, Proc. First Meeting Willi Hennig Soc. N.Y. Botanical Garden, New York.

- TAJIMA F., 1983. Evolutionary relationships of DNA sequences in finite populations. *Genetics*, **105**, 437-460.
- TAKAHATA N., 1985. Population genetics of extranuclear genomes : a model and review. In : OHTA T., AOKI K. (eds.), *Population Genetics and Molecular Evolution*, Japan Scientific Society (in press).
- TAKAHATA N., MARUYAMA T., 1981. A mathematical model of extranuclear genes and the genetic variability maintained in a finite population. *Genet. Res. Camb.*, **37**, 291-302.
- TAKAHATA N., NEI M., 1985. Gene genealogy and variance interpopulational nucleotide differences (in press).
- TAKAHATA N., PALUMBI S.R., 1985. Extranuclear differentiation and gene flow in the finite island model. *Genetics*, **109**, 441-457.
- TEMPLETON A.R., 1983. Phylogenic inference from restriction endonuclease cleavage site maps with particular reference to the evolution of human and apes. *Evolution*, **37**, 221-244.
- WOLSTENHOLME D.R., CLARY D.O., 1985. *Drosophila melanogaster* mitochondrial DNA : a novel organization and genetic code. *Nature*, **304**, 234-241.
- WU C.I., LI W.H., 1985. Evidence for higher rates of nucleotide substitution in rodents than in man. *Proc. Natl. Acad. Sci. (U.S.A.)*, **82**, 1741-1745.
- YONEKAWA H., MORIWAKI K., GOTOH O., HAYASHI J.I., WATANABE J., MIYASHITA N., PETRAS M.L., TAGASHIRA Y., 1981. Evolutionary relationships among five subspecies of *M. musculus* based on restriction enzyme cleavage patterns of mitochondrial DNA. *Genetics*, **98**, 801-816.
- YONEKAWA H., MORIWAKI K., GOTOH O., MIYASHITA N., MIGITA S., BONHOMME F., HJORTH J.P., PETRAS M.L., TAGASHIRA Y., 1982. Origin of laboratory mice deduced from restriction patterns of mitochondrial DNA. *Differentiation*, **22**, 222-226.