

# Les marqueurs des lignées uniparentales

Vocabulaire essentiel : ADN mitochondrial, lignée maternelle, lignée féminine, séquence de référence, chromosome Y, lignée paternelle, lignée masculine, haplotypes, haplogroupes

Ce chapitre traite des structures génétiques permettant de suivre les lignées maternelles et paternelles (lignées uniparentales) au sein des généalogies.

L'ADN mitochondrial et le chromosome Y sont des structures haploïdes, où chaque locus ne présente qu'un allèle, contrairement aux autosomes, qui sont des structures diploïdes, où chaque locus possède deux allèles (sur les deux chromosomes homologues). Si la variabilité autosomale (des SNP ou des STR) est donc représentée par des génotypes (deux allèles par marqueur), la variabilité mitochondriale ou du chromosome Y est représentée par des haplotypes (un allèle par marqueur).

La structure haploïde de ces deux systèmes de transmission impose une adaptation des modèles mutationnels et de leur utilisation. En l'absence de recombinaison, la variabilité observée ne peut être due qu'à des mutations, qui sont identifiées au sein d'une phylogénie étendue, et pas seulement au sein d'une famille. L'identification de la diversité des haplotypes conduit donc à la construction de réseaux caractéristiques avec des temps de divergence plus longs que ceux observés au sein d'une généalogie.

Les lignées féminines et masculines sont, dans une certaine mesure, spécifiques de l'origine biogéographique des individus. Les profils uniparentaux permettent donc de confirmer ou d'infirmer des analyses de parenté basées sur des marqueurs autosomaux, mais aussi d'estimer la diversité ethnique d'une population ou d'associer un individu dont l'identité est inconnue à une population dont les caractéristiques sont connues.

---

## STRUCTURE ET POLYMORPHISME DE L'ADN MITOCHONDRIAL

### Structure de l'ADN mitochondrial

L'**ADN mitochondrial** humain est une molécule circulaire présente dans les mitochondries humaines en plusieurs copies, ces mitochondries étant elles-mêmes des organites présents en de nombreux exemplaires au sein des cellules humaines. Il existe en moyenne quelques centaines à quelques milliers de molécules d'ADN mitochondrial par cellule.

Toutes les molécules d'ADN mitochondrial d'un individu partagent la même séquence de 16569 paires de bases (en l'absence de mutations) et ne sont donc pas seulement homologues, mais identiques.

L'ADN mitochondrial est donc une structure haploïde qu'on peut assimiler, pour les besoins de l'analyse, à une suite de 16569 nucléotides.

### Fonction, transmission et hérédité

Contrairement aux STR et à de nombreuses régions du génome nucléaire, le génome mitochondrial humain est en grande partie fonctionnel ; il comporte 37 gènes [Figure 1], qui codent pour des protéines ou sont transcrits en ARN. Les mitochondries sont des organites indispensables au fonctionnement des cellules et les mutations peuvent donc se révéler délétères si elles interviennent au sein de ces gènes. La molécule comprend également une région non-codante appelée Région de Contrôle, composée de trois sous-régions hypervariables HVR-1, HVR-2 et HVR-3 (« *Hyper Variable Region* »), qui présentent une variabilité beaucoup plus importante que le reste de la molécule d'ADN mitochondrial.

La séquence mitochondriale est strictement transmise par une mère à tous ses enfants [Figure 2]. Le père ne transmet pas son ADN mitochondrial. Toute mutation de l'ADN mitochondrial présente chez la mère est donc héritée par tous les enfants, et aucune mutation présente chez le père. Les marqueurs mitochondriaux sont donc caractéristiques de **lignées féminines** (aussi appelées « **maternelles** »).

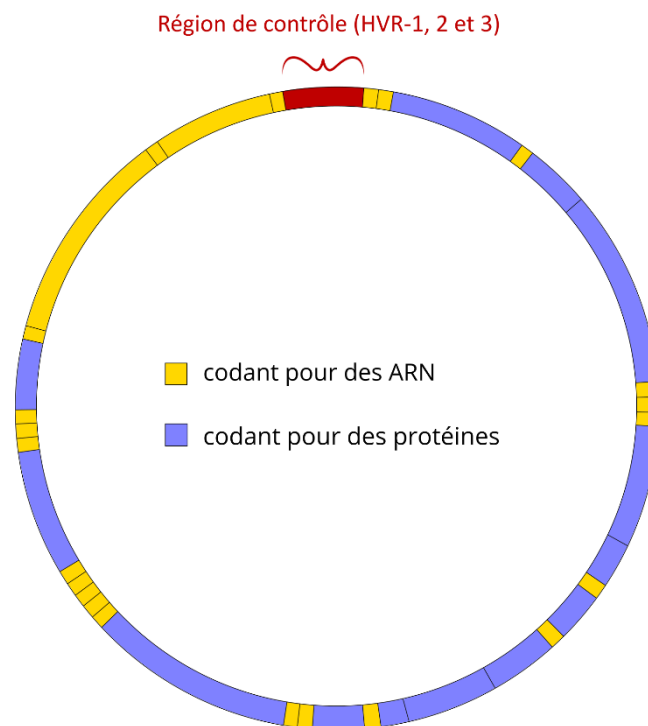


Figure 1 - Structure simplifiée de l'ADN mitochondrial

Adapté depuis [https://commons.wikimedia.org/wiki/File:Mitochondrial\\_DNA\\_it.png](https://commons.wikimedia.org/wiki/File:Mitochondrial_DNA_it.png). On note la présence de la région de contrôle hypervariable et la circularité de la molécule.

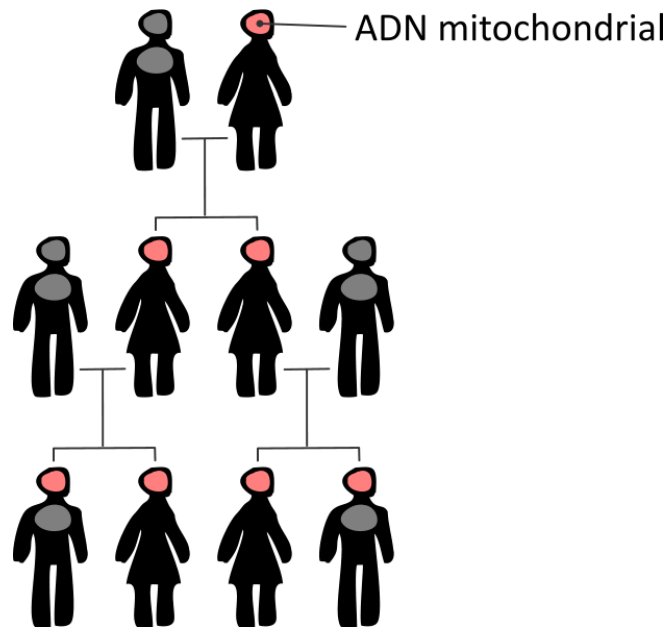


Figure 2 - Transmission féminine

Partage d'une lignée mitochondriale (en rose) au sein d'une famille étendue. Tous les descendants de la fondatrice partagent un haplotype, mais celui-ci n'est transmis que par les femmes de la généalogie.

### SNP et haplotypes mitochondriaux

Il n'existe pas de STR mitochondriaux. Les mutations de l'ADN mitochondrial sont donc des polymorphismes ponctuels, ou SNP. Ceux-ci sont définis, comme pour l'ADN autosomal, par rapport à une **séquence de référence**. Contrairement à l'ADN autosomal cependant, la nature haploïde du génome mitochondrial implique des profils SNP où chaque marqueur présente soit l'allèle ancestral, soit l'allèle muté. Les profils mitochondriaux sont donc des **haplotypes** [Figure 3].

Position du SNP mitochondrial	pos. 16129	pos. 16223	pos. 16298
Séquence de référence	... A A C T G ...	... T G G T A ...	... T C C C T ...
Séquence de l'individu A	... A A A T G ...	... T G G T A ...	... T C T C T ...
Séquence de l'individu B	... A A C T G ...	... T G T T A ...	... T C T C T ...
Haplotype de l'individu A	C 16129 A	—	C 16298 T
Haplotype de l'individu B	—	G 16223 T	C 16298 T

Figure 3 - Haplotypes mitochondriaux

Méthode de détermination des haplotypes mitochondriaux. Les SNP identifiés par rapport à la séquence de référence sont notés par leur position sur cette même séquence de référence et l'ensemble de ces SNP constitue l'haplotype mitochondrial de l'individu.

La transmission maternelle, et donc l'absence de recombinaison, implique une histoire évolutive (une phylogénie) pour chaque mutation mitochondriale. La séquence de référence ne peut donc pas dépendre de la population étudiée, car ce n'est pas la fréquence de l'allèle qui porte l'information de parenté ou de proximité génétique, mais sa présence. La référence standard est la Cambridge Reference Sequence (CRS), produite en 1981 et révisée en 1999 (rCRS), appartenant à une femme d'origine européenne. Bien qu'elle ne soit pas à l'origine phylogénétique des lignées mitochondriales, la rCRS permet une notation standard des mutations et la comparaison des haplotypes entre eux.

### Phylogénie et haplogroupes mitochondriaux

Un **haplogroupe mitochondrial** est un ensemble d'haplotypes mitochondriaux proches (c'est-à-dire ayant une origine ancestrale commune récente). Cette notion permet de représenter les divergences majeures dans la phylogénie mitochondriale. Les nœuds de l'arbre phylogénétique sont hiérarchisés en macro-haplogroupes (les branches principales), haplogroupes et sous haplogroupes. La séquence rCRS appartient par exemple au sous-haplogroupe H2a2a1, qui est compris dans l'haplogroupe H2, lui-même compris dans l'haplogroupe H, qui fait partie du macro-haplogroupe N [Figure 4]. Le terme « haplogroupe » est utilisé pour désigner à la fois les macro-haplogroupes, les haplogroupes et les sous-haplogroupes. Il faut noter que la nomenclature et la phylogénie évoluent et sont mises à jour régulièrement, à mesure que des corrections sont apportées et que de nouveaux échantillons sont analysés.

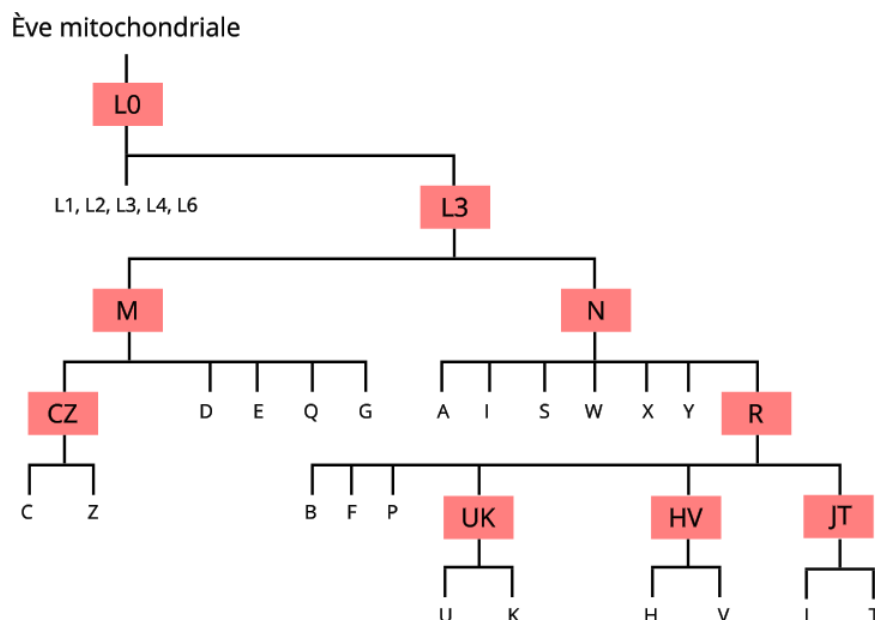


Figure 4 - Phylogénie simplifiée des haplogroupes mitochondriaux

Adapté depuis [en.wikiversity.org/wiki/Biological\\_anthropology#/media/File:MtDNA\\_tree.jpg](https://en.wikiversity.org/wiki/Biological_anthropology#/media/File:MtDNA_tree.jpg). On note la position terminale de l'haplogroupe H, dont provient la séquence de référence rCRS.

La phylogénie mitochondriale est, dans une certaine mesure, associée à l'origine géographique des individus : les haplogroupes sont parfois spécifiques de certaines régions du monde ou de certaines périodes historiques ou archéologiques. La présence dans une population archéologique de Sibérie du Sud d'individus porteurs de l'haplogroupe H, associé aux populations européennes, peut donc être l'indication d'une interaction entre ces populations, il y a plusieurs centaines ou milliers d'années. De la même façon, dans une population européenne homogène, la présence d'un haplogroupe typiquement associé aux populations Est-Asiatiques est une information précieuse dans l'identification de cet individu, y compris dans un contexte médico-légal.

### Mutations privées

Les mutations permettant de classer les haplotypes en haplogroupes sont dites « diagnostiques », les autres sont dites « privées ». Les mutations identiques apparaissant chez des individus éloignés généalogiquement et phylogénétiquement sont, par définition, des mutations privées. La distinction entre les mutations diagnostiques et privées est comparable à la distinction entre les caractères partagés et les convergences évolutives, dans le cadre de l'évolution des espèces. Toute nouvelle mutation dans une lignée est d'abord privée avant de se propager ou non dans la population, donnant potentiellement naissance à un nouveau sous-haplogroupe.

Certaines positions du génome mitochondrial sont particulièrement variables, c'est-à-dire que leur taux de mutation est plus élevé que dans le reste de la molécule. Bien que ces positions aient une valeur faible pour l'attribution d'un haplogroupe ou d'une origine géographique, elles permettent de constituer un haplotype mitochondrial très spécifique à l'individu, ce qui peut conduire à son identification.

### Implications méthodologiques dans l'étude des parentés

**Les individus d'une même lignée maternelle partagent le même haplotype mitochondrial.** Cependant, l'apparition de mutations peut conduire à des incohérences entre deux individus dont on connaît la parenté, ou dont la parenté est indiquée par d'autres méthodes. Plus les individus sont généalogiquement éloignés, plus la probabilité d'observer une mutation qui les distingue est importante.

Si la probabilité d'observer une incohérence entre l'haplotype d'une mère et de son enfant peut être calculée à partir des taux de mutation des SNP ou des régions concernées, un tel cas doit, en pratique, être spécifiquement discuté et la confrontation des différentes méthodes doit permettre d'estimer la fiabilité d'une telle parenté, indépendamment du partage de l'haplotype mitochondrial. Dans les cas où plusieurs mutations privées différencient deux individus, on peut considérer qu'il est raisonnable de ne pas conclure. Ici encore, l'interprétation des résultats peut rester impossible dans certains cas, en particulier dans l'étude de l'ADN ancien.

# STRUCTURE ET POLYMORPHISME DU CHROMOSOME Y

## Structure du chromosome Y

Le **chromosome Y** est l'un des plus petits chromosomes humains, avec une taille d'environ 57 millions de nucléotides. Sa structure est comparable à celle des autres chromosomes de l'ADN nucléaire, excepté qu'il ne fait pas partie d'une paire de deux chromosomes homologues, comme c'est le cas pour tous les autosomes, et qu'il n'est présent que chez les individus de sexe masculin, au sein de la paire XY (puisque'il porte les gènes déterminant le sexe masculin).

La majeure partie du chromosome Y ne subit pas de recombinaison avec le chromosome X lors de la gamétogenèse et peut donc être considérée indépendamment. Les profils Y sont donc, comme les profils mitochondriaux, des haplotypes (un allèle par locus) et non des génotypes (deux allèles par locus) [Figure 5].

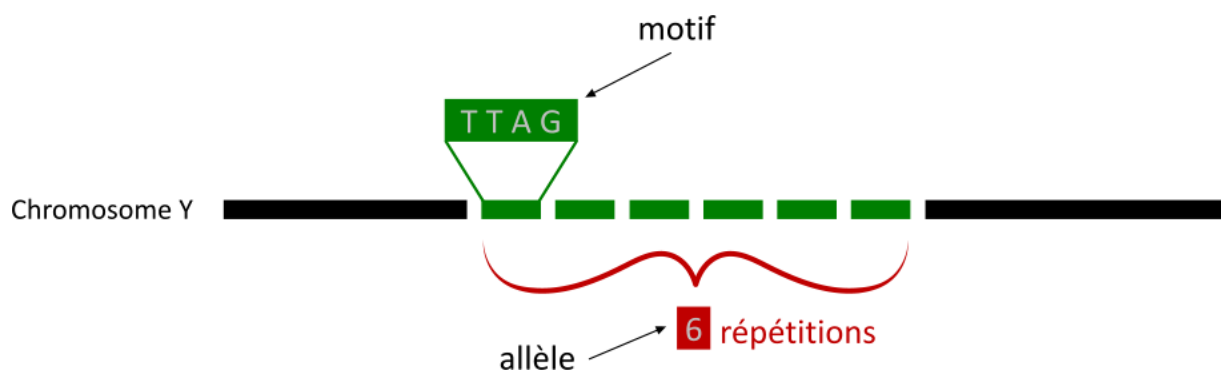


Figure 5 – Structure des STR-Y

*De la même manière que les allèles des STR autosomaux, les STR-Y sont des nombres de répétitions. La structure haploïde implique cependant que les loci STR-Y sont mono-alléliques.*

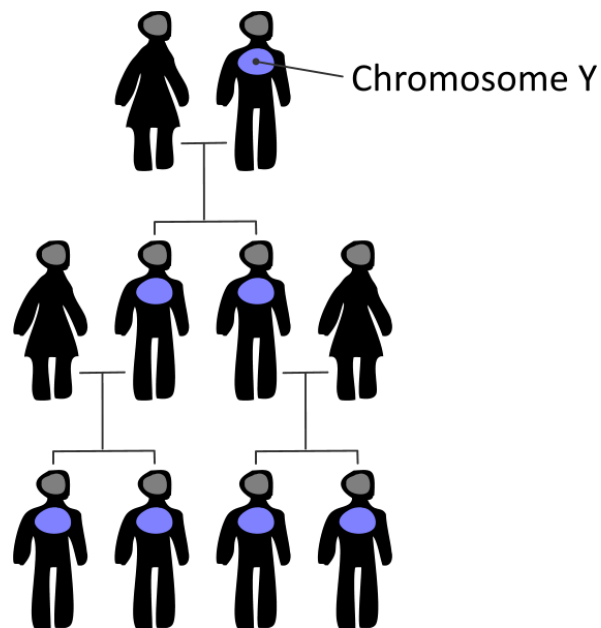
## Fonction, transmission et hérédité

Le chromosome Y comprend un petit nombre de gènes – dix à trente fois moins que des autosomes de même taille. Il porte en particulier des gènes liés à la masculinisation de l'embryon mais son absence chez la moitié des individus de l'espèce humaine (les femmes, qui portent deux copies du chromosome X) ne permet pas que le chromosome Y porte des gènes dont l'absence serait délétère. En conséquence, le chromosome Y est l'une des parties les plus variables du génome nucléaire, même si des différences importantes existent selon le type de marqueur étudié (STR ou SNP).

La séquence du chromosome Y est transmise par un père à tous ses fils [Figure 6], les femmes ne possédant pas de chromosome Y. **Tout polymorphisme du chromosome Y présent chez le père est donc hérité par tous ses fils** et les marqueurs Y sont caractéristiques de **lignées masculines** (aussi appelées « **paternelles** »). Les marqueurs de ces lignées sont sélectionnés dans des régions non-recombinantes du chromosome.

## Haplogroupes Y

Comme la phylogénie mitochondriale, la phylogénie Y est associée à l'origine biogéographique des individus. Les **haplotypes** des individus peuvent également être classés dans des **haplogroupes**.



*Figure 6 – Transmission masculine*

*Partage d'une lignée Y (en bleu) au sein d'une famille étendue. Tous les descendants mâles du fondateur partagent un haplotype Y. Les femmes ne possèdent pas de chromosome Y.*