

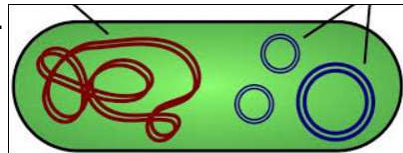
## ORGANISATION DU GÉNOME

### 4ème séance 2016

**Un Génome** est l'ensemble de l'information héréditaire d'un organisme, présente en totalité dans chaque cellule.

#### Caractéristiques des génomes procaryotes:

- Chromosome circulaire unique
- Une origine de réplication
- Présence possible de petites séquences d'ADN circulaires indépendantes : les plasmides.

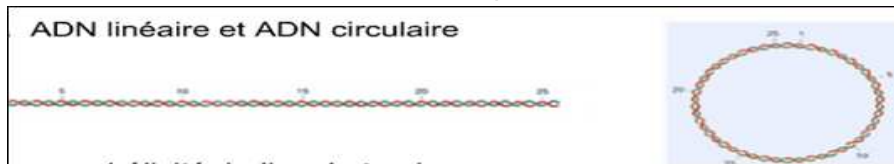


#### Les gènes procaryotes :

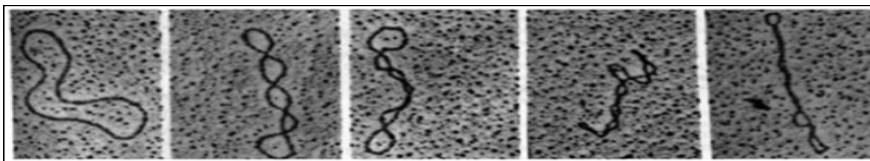
- Fraction codante des génomes élevée: > 90% codant
- Peu de séquences intergéniques : Génome « compact »
- **la séquence des gènes est continue : pas d'intron**
- Gènes organisés en opérons. 600 opérons dans le génome de *Escherichia coli*. (Un opéron est une unité d'ADN fonctionnelle regroupant des gènes qui opèrent sous le signal d'un même promoteur)

### Organisation Des Adn des Procaryotes : Les Topoisomères

- L'ADN peut exister à l'état *relâché* avec une contrainte minimale dans la molécule. C'est la forme la plus stable de la molécule



- L'axe de la double hélice d'ADN peut s'enrouler sur lui-même en formant un **superenroulement positif ou négatif** : rendre l'ADN plus **compact** et diminuer ainsi le volume occupé dans la cellule.



- **Les topoisomères** : deux molécules d'ADN qui diffèrent uniquement par le nombre d'enroulements (enlacements).

3

## LES TOPOISOMÉRASES

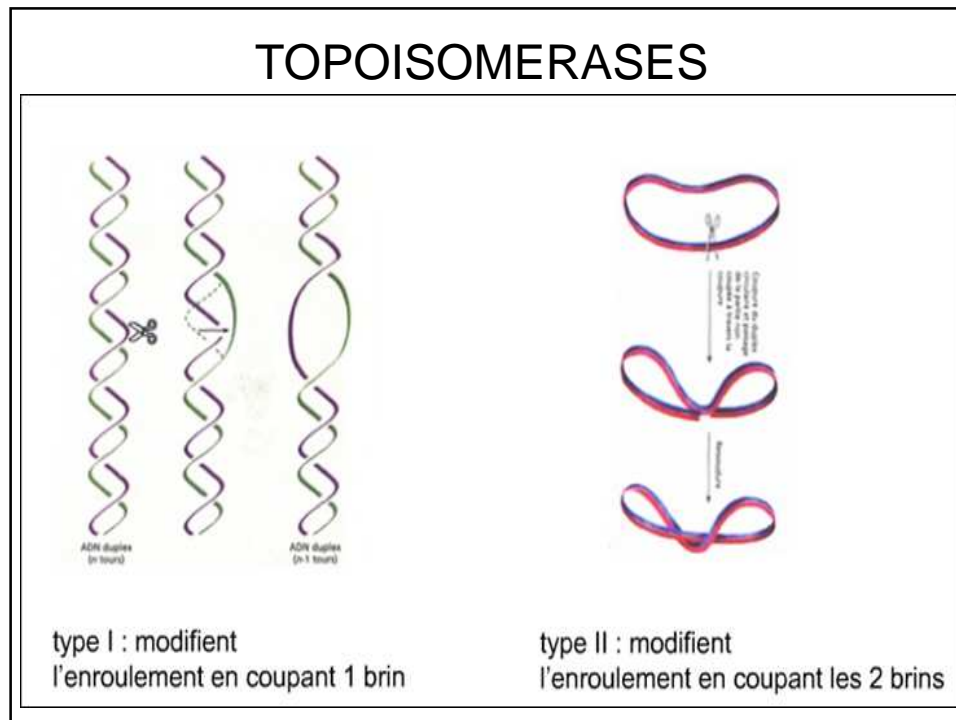
**Topo-isomérases** : enzymes qui modifient les enlacements des brins. Nécessitent la coupure d'un ou deux brins de l'ADN, elles permettent de tordre et détordre les deux brins (et donc de défaire les supertours)

- Elles ont été mises en évidence chez les procaryotes et les eucaryotes.
- Leur importance est capitale dans la réplication, la transcription de l'ADN.

Topoisomérase 1	Topoisomérase 2
Coupe 1 brin (par rupture liaison phosphodiester)	Coupe les 2 brins (par le même procédé)
Défait les supertours négatifs uniquement	Défait les supertours positifs ET négatifs
Ne nécessite pas d'énergie (ATP)	A une fonction ATPasique (besoin d'ATP)
Conduit à l'état relâché	C'est la gyrase d'E. Coli

- Par exemple : l'acide nalidixique inhibe la gyrase des bactéries. Ce principe peut être utilisé dans les traitements anticancéreux
  - UP16 : inhibiteur topoisomérase 2
  - Taxol : inhibiteur topoisomérase 1 (extrait de l'if)





## ORGANISATION DE L'ADN CHEZ LES EUCARYOTES

### Caractéristiques des génomes eucaryotes

- Dans le noyau
- Taille >> génome procaryote
- Plusieurs chromosomes (homme 23, cheval 32, levure 16, drosophile 4...)
- Plusieurs origines de réplication par chromosome
- Gènes « disloqués » (exons, introns)
- Grandes régions intergéniques de fonction inconnue
- Chromatine!

## ORGANISATION DE L'ADN CHEZ LES EUCARYOTES

### Organisation de l'ADN : chromatine

#### □ Problème : stockage de l'ADN

- 1 cellule : 7 pg ADN soit 2,5 m dans un noyau de 5  $\mu\text{m}$  de  $\varnothing$

→ nécessité de compaction

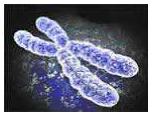
#### □ Comment : 2 états d'organisation

- Chromatine pendant l'interphase sous 2 formes

↪ hétérochromatine : ADN condensé

↪ euchromatine : grandes boucles d'ADN peu condensé

passage réversible d'une forme à l'autre (acétylation des histones et rôle des protéines non histones)



- Chromosomes pendant la mitose

### PROTEINES HISTONES

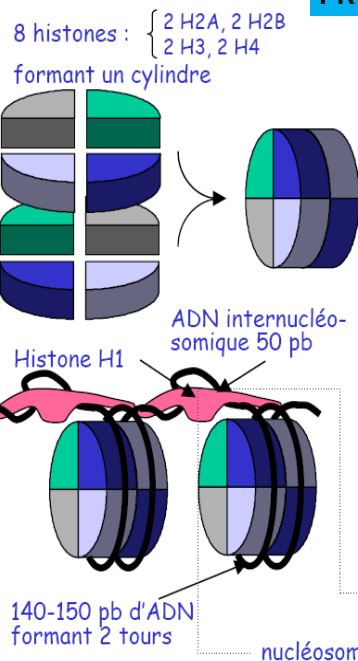
#### □ Eléments de la chromatine

association de protéines et d'ADN

- Protéines ↪ histones  
↪ non histones

- Histones

- petite taille : 100 à 200 aa **60 millions**
- riches en aa basiques (Lys, Arg)
- très conservées au cours de l'évolution
- gènes sans introns
- interactions avec ADN
  - protéines de l'enveloppe
- Classification
  - Histones nucléosomiques : H2A, H2B, H3, H4
  - Histone H1 **n'appartient pas au nucléosome**
- Organisation
  - 146** pb d'ADN formant  $\approx$  2 tours autour de 8 histones formant un cylindre
  - histone H1 permet associat° des nucléosomes



**DNA**

**H1 Histone**

**8-Histone Core**

**30 nm**

**H1**

- ▶ histones : protéines de petite taille (MM 11 à 22 kDa, 100 à 200 aa)
- ▶ protéines riches en acides aminés basiques (charge +): arginine, lysine, histidine
- ▶ séquence primaire en acides aminés très conservée
- ▶ modifications post-traductionnelles: acétylations, phosphorylations ...

↳ « code histones »

## Modifications des histones

### 2 - Modification des histones

**Domaine N terminal**

- acétylation
- méthylation
- phosphorylation

**Domaine C terminal**

- ubiquitinylation

**Acétylation liée à activité transcriptionnelle**

- acétylation → transcription
- désacétylation → pas de transcription

**H2A**

N S 1 K 5 K 119 C

**H2B**

N K 5 K 12 K 15 K 20 K 120 C

**H3**

N K 9 S 10 K 14 K 18 K 23 K 27 S 28 C

**H4**

N S 1 K 5 K 8 K 12 K 16 K 20 C

histone-fold domain

*D'après Molecular Biology of the Cell, 4<sup>th</sup> ed*

**L'épigénétique: modification de l'expression des gènes non liés a des changements de séquences (mutations)**

**L'ADN**  
 succession des bases A,T,G,C = définit le code génétique

**La chromatine**  
 gène inactif = définit le code épigénétique (CODE HISTONE)  
 gène actif

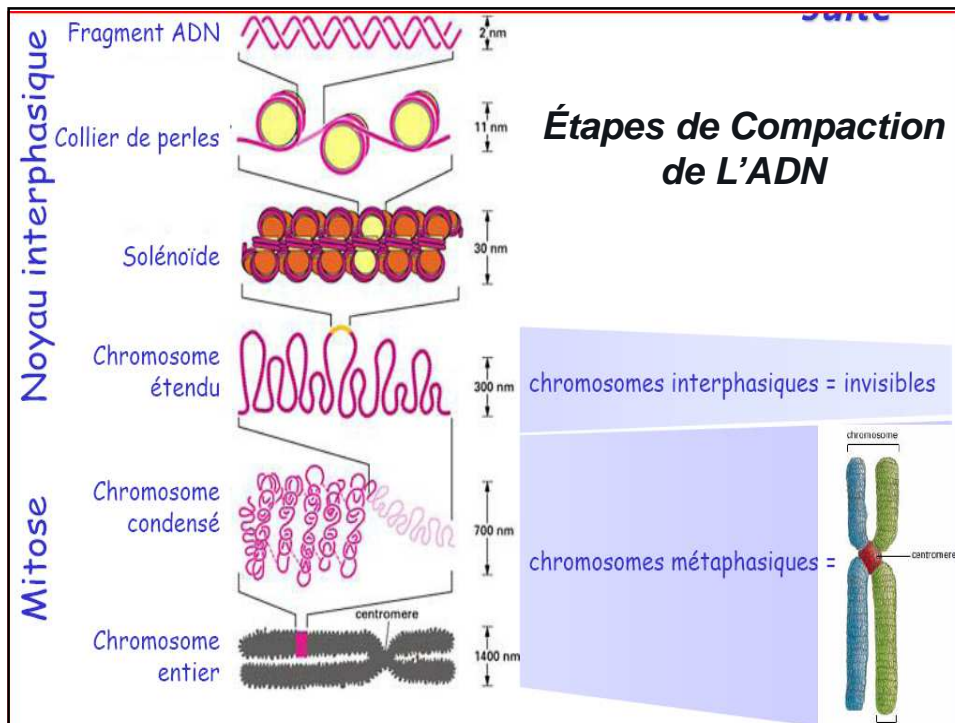
**Repression**  
 Co-repressor  
 Transcriptional Repressor  
 HDAC  
 Hypoacetylated, condensed chromatin

**Activation**  
 Transcriptional Activator  
 HAT  
 Hyperacetylated, accessible chromatin

**EUCHROMATIN** : Regions with high transcriptional activity are loosely packed  
**HETEROCHROMATIN** : Regions with low or no transcriptional activity are densely packed

Legend:  $\text{H}$ : Histone,  $\text{D}$ : DNA,  $\text{A}$ : Acetyl group,  $\text{P}$ : Phosphoryl group,  $\text{M}$ : Methyl group

More methyl - Genes OFF



### □ Chromosomes

- 23 paires chez homme
- 22 paires d'autosomes (1 à 22)
- 1 paire de chromosome sexuel (XX ou XY)
- Structure

Centromère : étranglement

Bras : bras court (p)  
bras long (q)

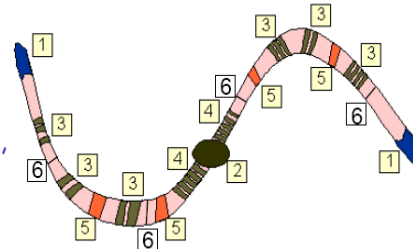
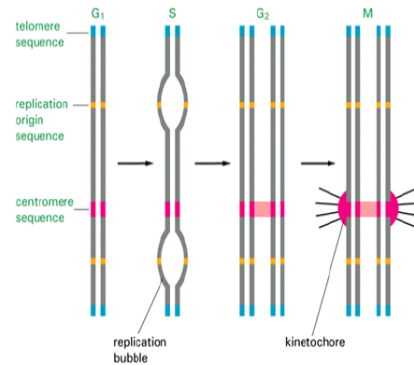
- Classification

Classement par paire en taille décroissante pour les 22 paires d'autosomes → caryotype

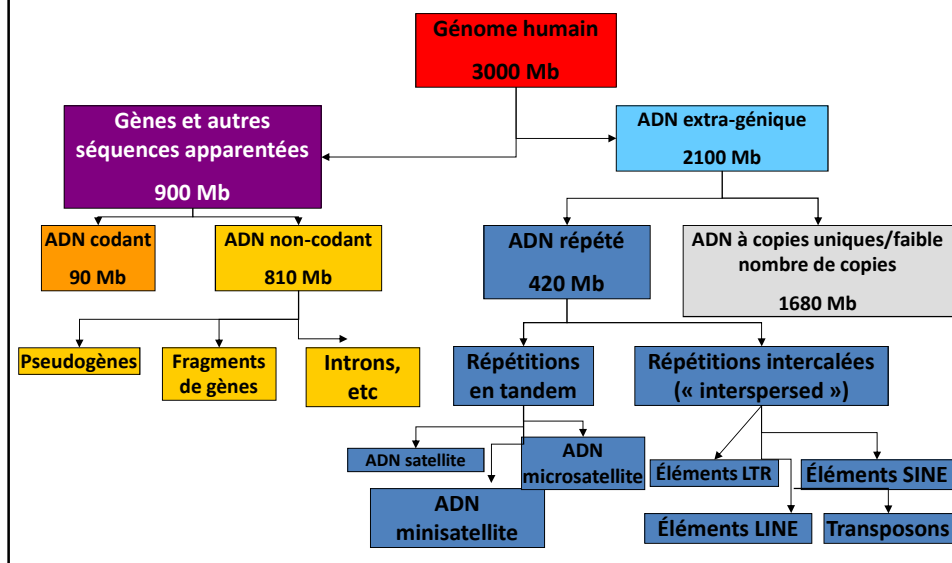
*suite*

□ Zones d'un chromosome

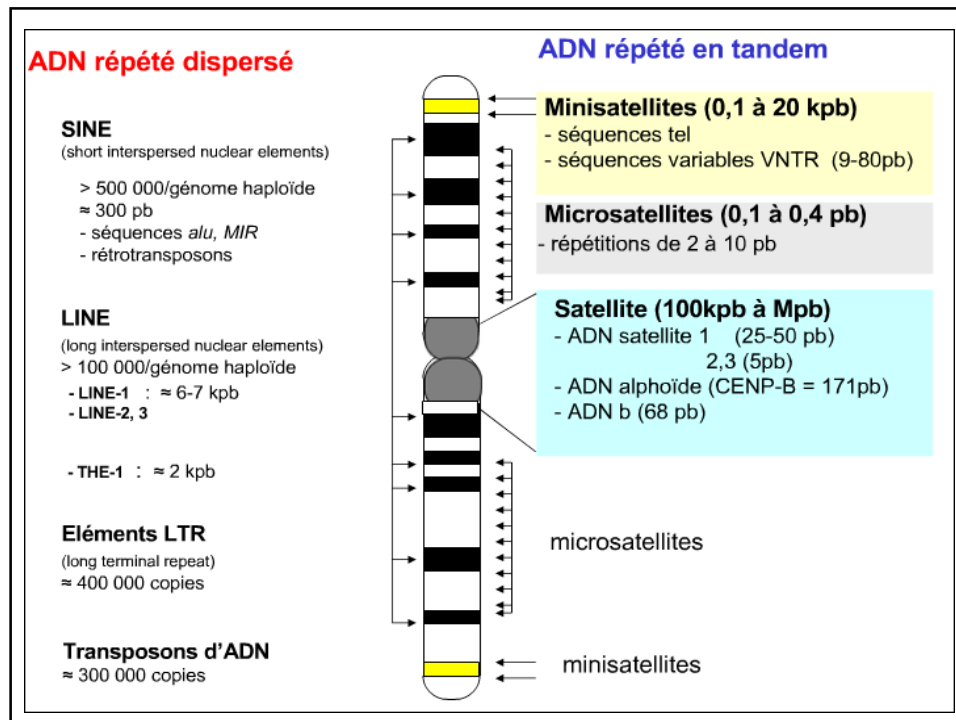
- [1] deux télomères
- [2] un centromère
- [3] des séquences non codantes modérément répétitives
- [4] des zones non codantes très répétitives (zones Alu de 300 bp), ne correspondant pas à un gène
- [5] des gènes
- [6] de multiples origines de réplication



## Le génome humain







## A- Séquences répétitives en tandem

### 1. DNA satellite:

- DNA principalement centromérique
- DNA répété en tandem
- bloc de 100kb à plusieurs Mb de longueur
- Le DNA **satellite** a une **densité différente** du reste du DNA.
- 3-5% du DNA de chaque chromosome

## Séquences répétitives en tandem

### 2 - ADN Minisatellite des régions télomériques :

- bloc de 0.1 à 20kb correspondant à des répétitions de 6 nucléotides (TTAGGG) dans les régions télomériques

• Les télomères sont indispensables pour la réplication et pour la stabilité des chrs : elles sont situées à leurs extrémités grâce à des télomérases (télomère synthase)

Chez l'homme séquence *héxanucléotidique répétée*: TTAGGG

• réduction de la taille télomérique est liée au problème de «fin de réplication».

• Au bout d'un certain nombre de divisions, les télomères vont atteindre une taille critique qui est corrélée avec la mort physiologique de la cellule (horloge biologique)

• implication des télomères dans le phénomène de sénescence.

• la détection de l'activité de la télomérase pourrait constituer

un marqueur intéressant dans le diagnostic de certaines tumeurs.

#### Prix Nobel 2009 pour les télomères

Elizabeth H. Blackburn, Carol W. Gleider et Jack W. Szostak

pour le rôle déterminant de ces chercheurs dans la

compréhension des télomères et la découverte

de la télomérase.



### Minisatellites (VNTR, Variable Number of Tandem Repeats):

régions du génome caractérisées par la répétition en tandem d'une même séquence de DNA. Le nombre de répétitions est variable. Le motif de base est d'une longueur > à 10 nucléotides.

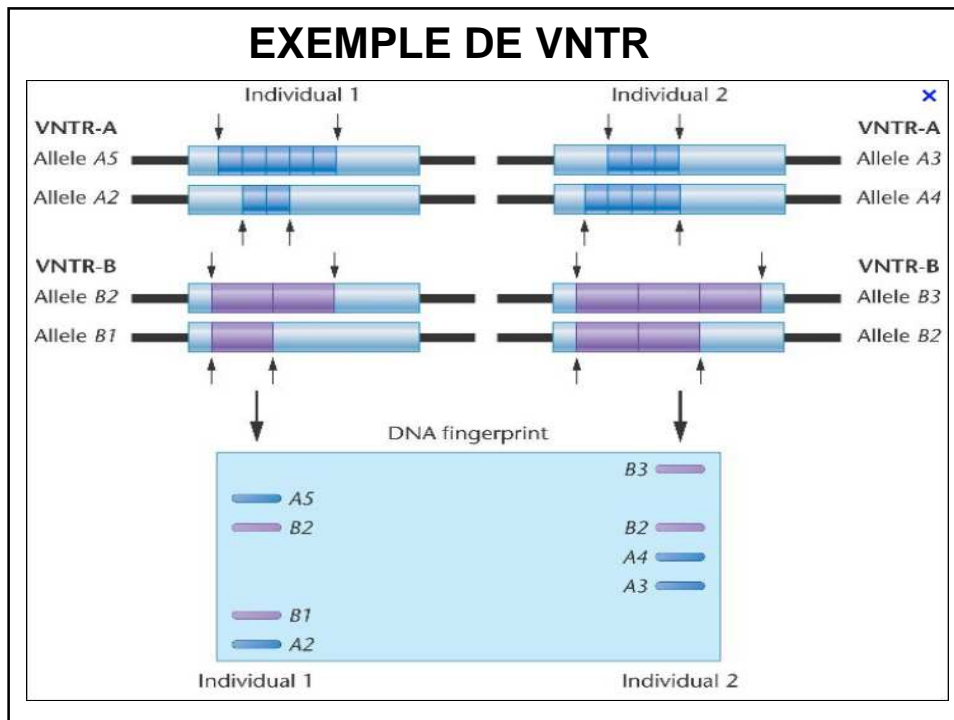
```

5'---- CGCCTCGGCCTCCCAGAGTGGCTGAGATTACAGGCGTGAACCACCATGCCTAGCCGTTAGCTCCCA
CTTATGAGTGAGAACAGGTGATGTTTTGGTTTTCCATTCCCTGAGTTACTTTACCCAGAATTGTTGT
CTCCAATCTCATCGAGGTCTCTGCGAATGCCAGTAATTCATTCCCTTTTATGGCTAAGTAGTATT
CCATCGTATATATACATATACATATATATGTATACACACATATACATATATATGTATACACACAT
ATACATATATATGTATACACACATATACATATATATGTATACACACATATACATATATATGTATA
CACACATATACATATATATGTATACACACATATACATATATATGTATACACACATATACATATAT
ATGTATACACACATATACATATATATGTATACACACATATACATATATATGTATACACACATATA
CATATATATGTATACACACATATACATATATATGTATACACACATATACATATATATGTATACAC
ACATATACATATATATGTATACACACATATACATATATATGTATACACACATATACATATATATG
TATACACACATATACATATATATGTATACACACATATACATATATATGTATACACACATATACATA
TATATGTATACACACATATACATATATATGTATACACACATATACATATATATGTATACACACAT
ATACATATATATGTATACACACATATACATATATATGTATACACACATATACATATATATACATATA
TACACAAACACACATGCACCCGCACTTTTTTTTTTTTTTTTTTTTTTTTGGAGATGGGAGTCTCACTCTAT
CACAGGCTGGAGTGCCAGTGGTGTATCTTGGCTCACTGCAACCCTCTGTCTCTGGGTTCAAAGCT
ATTCTCCTGCTTCAGCCTCCTGAGTAGCTGGGATTACAGGTGCTCACCCACATGCCAGCTAAATT
TTTGTATTTTTACCATGTTGGCCAGGATGATCTCCATCTCCTGACCTCCTGATCCTCCTGCCCTTG
GGGTCCC AAAAGTGCAGGGATTGCAGGCATGAGCCACTGCATGTGGCCACACCACACTTCCTTTAT
CCACTTGTGTGGATGGCATTGTGTTGAAAAGAGTCTATCTCTGTTTTTTCTTTCCCTGAACG----3'

```

← CCTTTTATGGCTAAGTAGTATTCCATCGT → } Amorces  
 ← TTCAGTAGAGCAAAAAGAAAGGGACTTGC →

Le minisatellite est ici constitué par la répétition (25 fois) du motif de base de séquence consensus GTATACACACATATACATATATAT (24 nucléotides) avec un très faible % de variation interne (les substitutions sont en bleu foncé). Les séquences (30 nucléotides chacune) utilisées pour le choix des amorces sont en rouge. Il est facile de prévoir quelle sera la taille du fragment amplifié par PCR qui révélera la présence de cet allèle du marqueur CEB 310. On le retrouvera sur le gel, après électrophorèse au niveau du marqueur de taille 1000 pb.

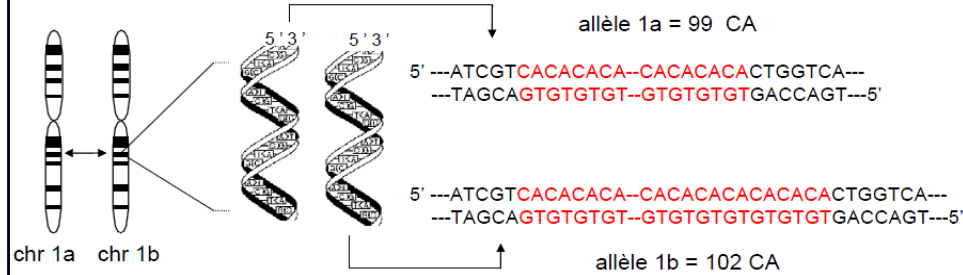


- Les minisatellites peuvent être détectés par RFLP (polymorphisme de longueur de fragment) en utilisant des enzymes de restriction qui coupent un grand nombre de fois le génome mais jamais dans les minisatellites.
- Un RFLP est alors révélé par l'existence d'un nombre de répétitions différent entre individus, ce qui produit des fragments de tailles différentes.
- Les minisatellites sont révélés après PCR ce qui nécessite la mise au point d'amorces spécifiques.

**Test de paternité par PCR-RFLP**  
**L'enfant 2 est probablement illégitime !**

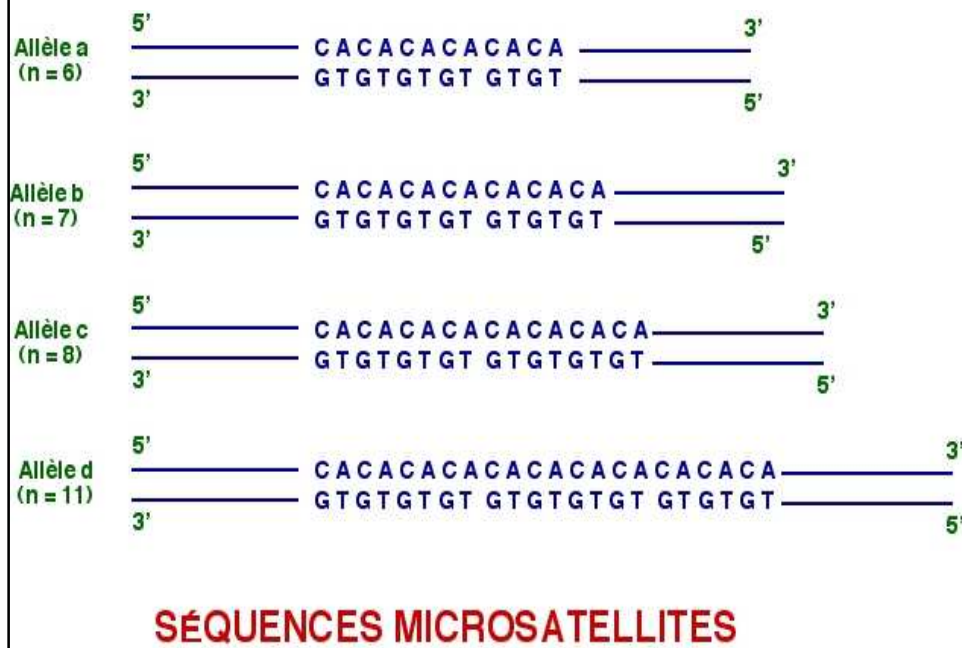
© Yong-Ming YUAN

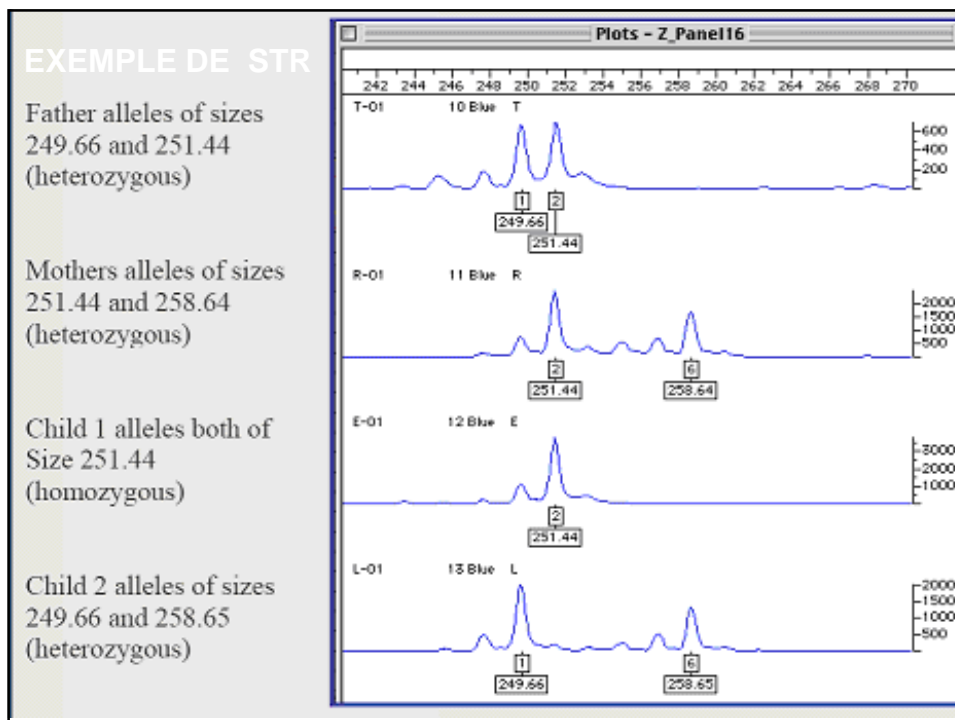
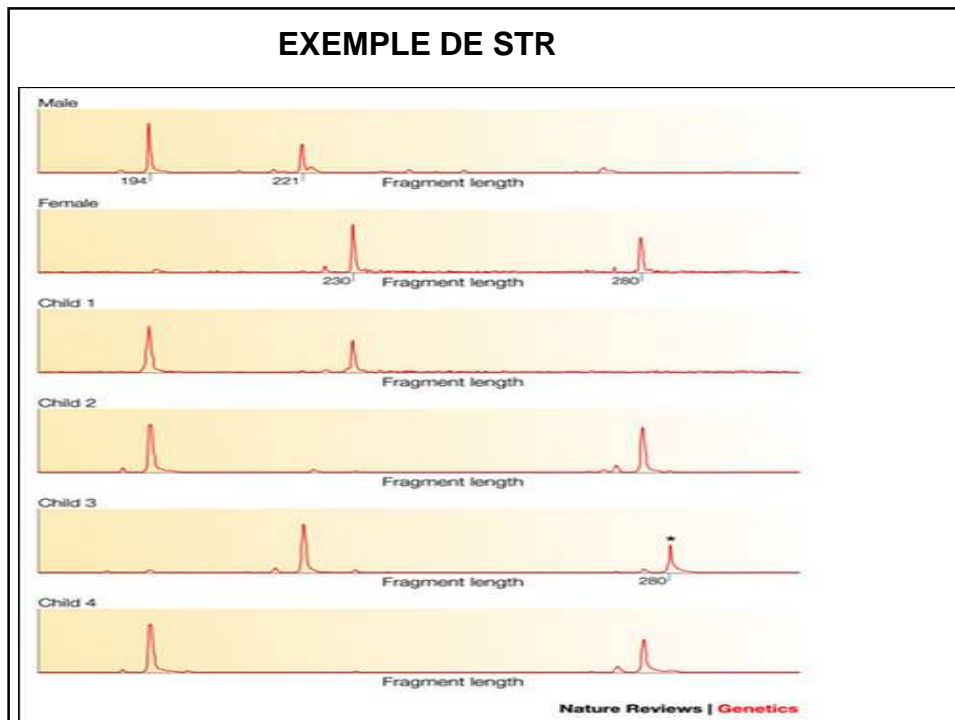
### 3 - Répétitions de type Microsatellites STR (Short Tandem Repeat)

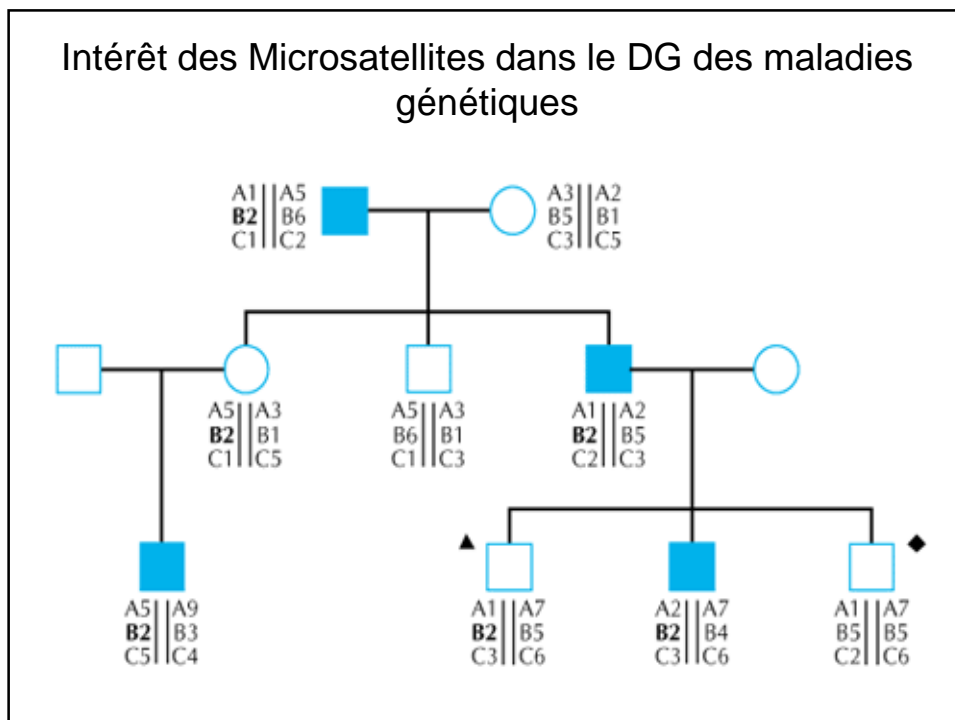
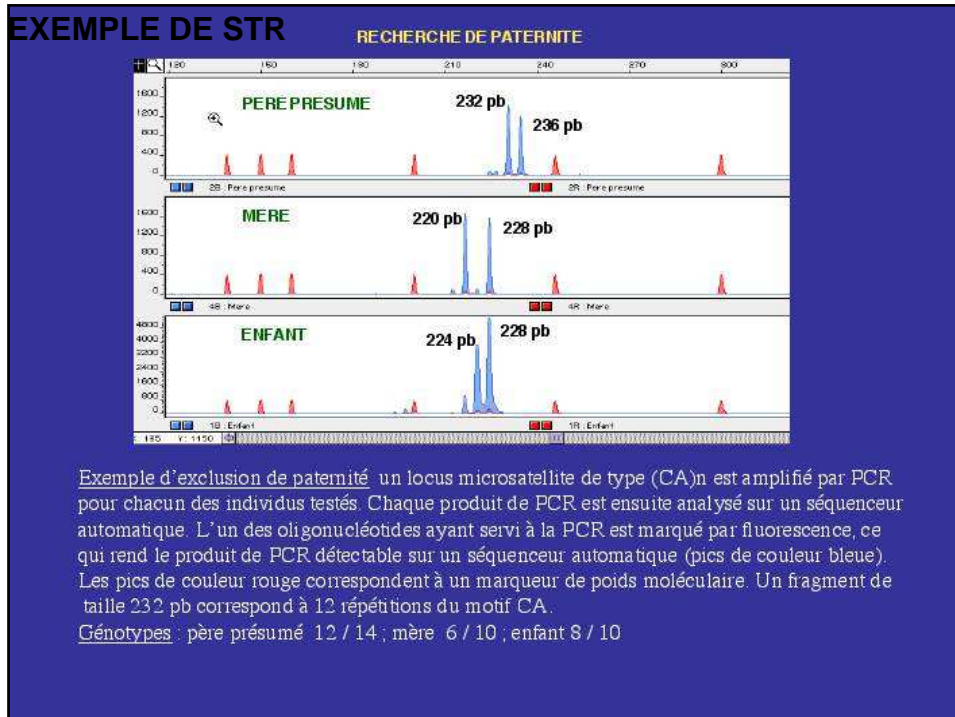


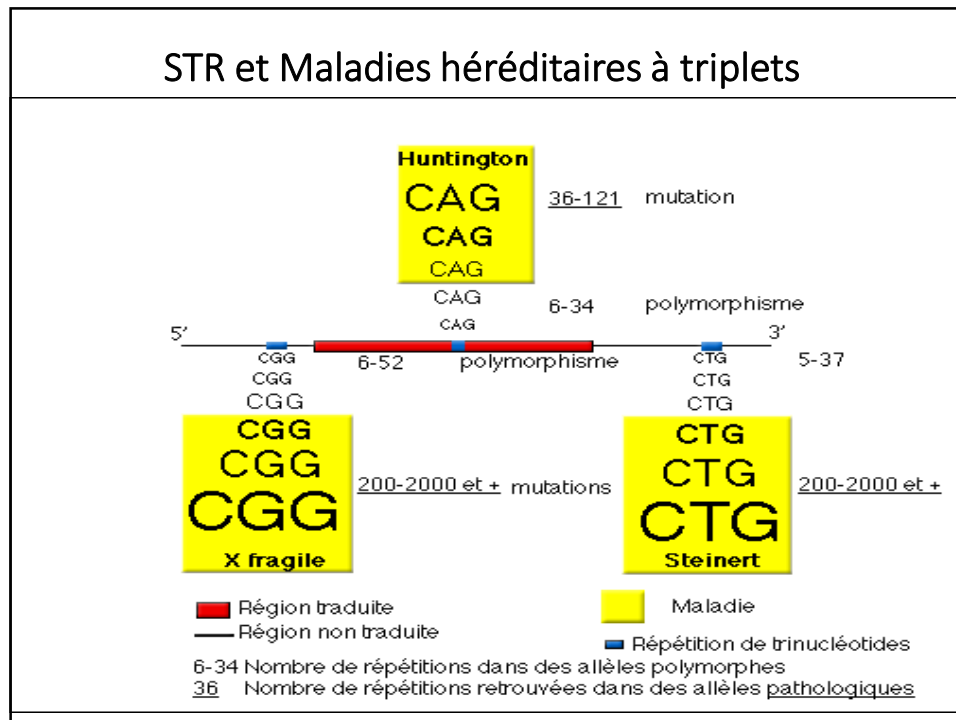
- motifs de 2, 3 ou 4 nucléotides
- nombreux (> 50 000 dans le génome humain)
- le nombre de répétition varie suivant le microsatellite étudié
- le nombre d'allèles est généralement compris entre 2 et 10
- le nombre de répétitions ne varie généralement pas lors de la transmission
- les séquences de part et d'autre du microsatellite sont identiques pour les différents allèles d'un microsatellite

#### EXEMPLE DE STR









## B- DNA dispersés

La majorité des gènes eucaryotes sont constitués de gènes répétés qui se déplacent dans le génome : Eléments mobiles du génome

- 1) les transposons
- 2) les rétrotransposons

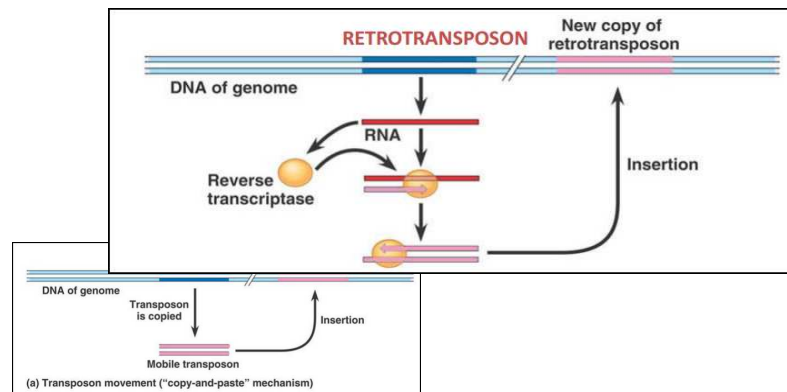
Ces deux groupes se différencient par leur mode de mobilisation.

Les rétrotransposons utilisent un intermédiaire ARN pour leur mobilisation. Ils utilisent donc une transcriptase inverse pour retransformer l'ARN en ADN, comme les rétrovirus.

Les SINEs, les LINEs sont des exemples de rétrotransposons.

Ces éléments mobiles peuvent avoir un rôle mutagénique.

## Mécanismes de transposition : *Transposases et rétrotransposases*



- Une séquence Alu est un court fragment d'ADN de type SINE (Short INterspersed repetitive Element) caractérisé par la présence d'un site de restriction de l'endonucléase de restriction : l'enzyme Alu I, spécifique de l'Homme et de certains primates.
- Elle mesure à peu près 300 pb et on en a à peu près 500 000 copies dans le génome.
- les séquences Alu sont les éléments mobiles les plus abondants du génome humain.
- Des insertions de séquences Alu ont été impliquées dans plusieurs maladies humaines héréditaires, dont plusieurs formes de cancer .
- L'étude de séquences Alu a joué un rôle important dans l'étude génétique des populations humaines : Les séquences Alu sont particulièrement spécifiques de l'Homme.



gène A      gène B

gènes : unités d'information génétique transmise par un individu à sa descendance et correspondant à une séquence d'acides nucléiques codant pour une protéine (ARNm) ou un ARN (ARNt, ARNr, ARNsn...)

densité génique variable

- entre espèces : - ≈ 11 gènes /100 000 pb chez l'homme  
- ≈ 479 gènes /100 000 pb chez la levure
- au sein d'un chromosome : peu de gènes vers les centromères

100 kb

Spectre des couleurs  
Haute densité      Basse densité

### 1- Structure d'un gène procaryote

promoteur    site d'initiation    gène 1    gène 2    gène 3    terminateur

ARN polycistronique

protéine 1    protéine 2    protéine 3

Les gènes ont une structure « en mosaïque »

### 2. Gène Eucaryote

exons = séquence exprimée

introns : séquence non exprimée

bornes exon-intron

5' ----tcagTGTT-----CGTACCTTAAAgtagtc---- 3'  
3' ----agtcACAA-----GCATGGAATTTcaatcag---- 5'

intron    exon    intron

Le nombre d'exon-intron varie suivant les gènes

- gène de l'interféron α :	600 pb	1 exon
- gène de la dystrophine :	2 400 000 pb	79 exons
- gène du récepteur de la ryanodine 1 :	161 000 pb	106 exons

Gènes chevauchants:

Gènes introniques

gène A    gène B

gène A    exon 1    gène B    exon 2    exon 3

Structure d'un gène eucaryote

➔ les gènes ont des tailles variables :

gène de l'ARNt <sup>tyr</sup> :	100 pb
gène de l'insuline :	1 400 pb
gène CFTR :	250 000 pb
gène de la dystrophine :	2 400 000 pb

➔ gènes uniques : 2 copies / cellule  
gènes multicopies : ARNt, ARNr (ADNr), certains ARNm (histones....)

✓ familles multigéniques « simples » : les ARNr

- ≈ 2000 gènes ARNr 5s sur le chr 1
- 280 copies des ARNr 28s, 5.8s, 18s en 5 groupes sur chr 13, 14, 15, 21, 22

✓ familles multigéniques « complexes »

- les gènes de la globine
- les gènes de l'aldolase : 5 gènes sur les chr 3, 9, 10, 16 & 17

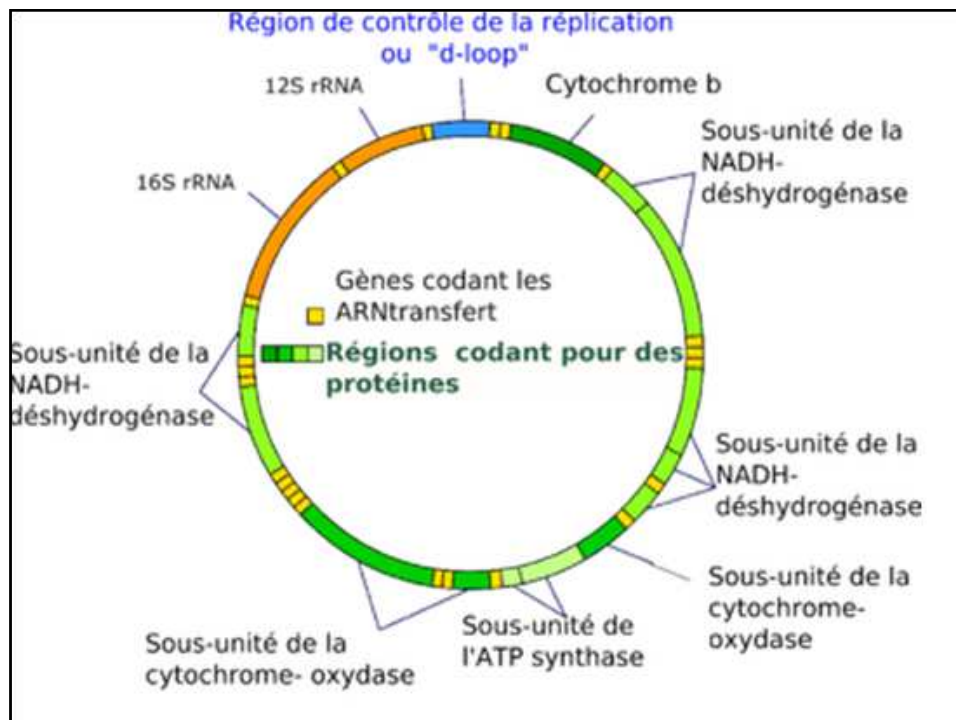
## Comparaison de génomes entre espèces

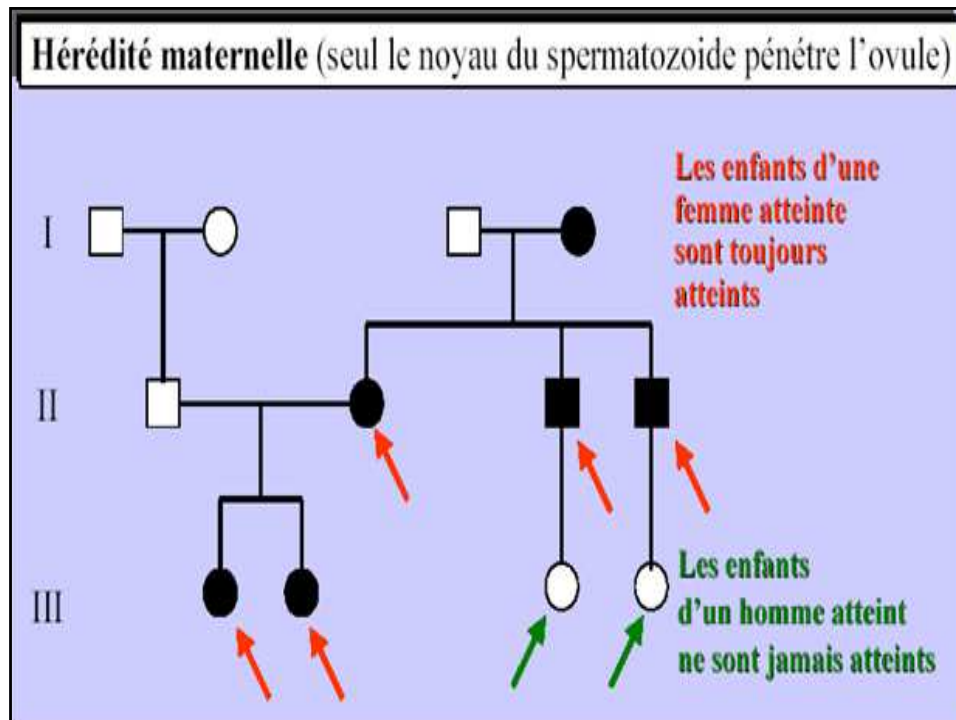
- Homme A / homme B
  - 99,5% identique (0,5% différence)
- Homme/chimpanzé
  - Codant: 98,5% identique
  - Non codant: ~97% identique
  - Quelques duplications/délétions importantes de région de quelques dizaines de kb
- Homme/souris
  - Codant: 90% identique
  - Non codant: la majorité est sans identité apparente, mais on trouve quand même de nombreux segments semblables (« conservés »)
- Homme/poulet
  - Codant: 80% identique
- Homme/poisson
  - Codant: 70% identique

## L' ADN MITOCHONDRIAL

### La mitochondrie a son propre génome: ADN circulaire, double brins:

- **16561pb** : 37 gènes sans introns et sans protéines histones), sans séquences répétées
- code 22 ARNt, 2 ARNr et 13 protéines membranaires impliquées dans la production d'énergie par la mitochondrie
- 5 à 10 copies par mitochondrie
- Grande mutabilité : 10 X > par rapport au génome nucléaire
- Dans le DNA mitochondrial, le code génétique est inhabituel:
  - Ex: AGA - stop codon et non pas codon pour l'Arg.
- **Importance du DNA mitochondrial dans l'hérédité:**
  - le DNA mitochondrial est d'origine maternelle
  - correspond à 0.5% du DNA dans une cellule somatique
  - Mode de transmission caractéristique
- **Importance du DNA mitochondrial dans la filiation (études phylogéniques)**





## HEREDITE MITOCHONDRIALE

### Pathologie: caractères généraux

**Hétéroplasmie:** toutes les mitochondries n'ont pas le même ADN (mutabilité + milliers de mito / cellule et / individu et nb copies / mito)  
hétérogénéité génétique mitochondriale = variabilité clinique +++

**Homoplasmie :** toutes les mitochondries ont le même ADN  
Uniformisation des cellules pendant l'embryogénèse.  
Mutations homoplasmiques rares et graves (voir létales)

**Pathologie de la chaîne respiratoire ++**  
Phénotype dépend du besoin énergétique du tissu:  
SNC puis muscle squelettique puis cœur, rein, foie