

## **Chapitre 02**

### **Organisation des génomes**

**1 /Chez les eucaryotes**

**2/Chez les procaryotes**

## Organisation des génomes

### Introduction :

Le génome représente l'ensemble du matériel génétique d'un organisme. C'est aussi le contenu en ADN ou en ARN d'un organisme. Chez la majorité des organismes, le génome est composé d'ADN. Cependant, chez certains virus tel que les rétrovirus (SARS-CoV2 et VIH) le matériel génétique est fait l'ARN.

### Génome

#### VIRUS :

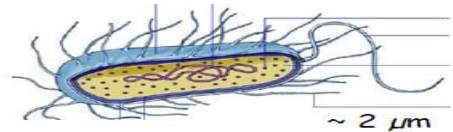
1 molécule d'ADN ou 1 molécule d'ARN\_



~ 0,2  $\mu\text{m}$

#### PROCARYOTES (bactéries) :

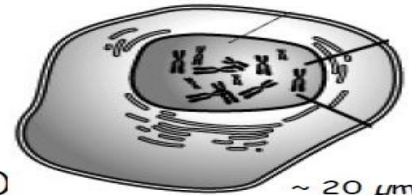
1 unique chromosome  
(4,6 millions de paires de bases ou pb)  
constitué d'une molécule  
d'ADN circulaire (+ plasmides)



~ 2  $\mu\text{m}$

#### EUCARYOTES :

2n chromosomes  
situés dans le noyau  
+ génome mitochondrial (ADN circulaire)



~ 20  $\mu\text{m}$

La taille des génomes varie en fonction des espèces, elle est définie comme la quantité d'ADN présente dans les gamètes haploïdes d'une espèce donnée = Valeur C,

Elle est mesurée +++ sur le spermatozoïde chez l'homme, exprimée en nombre de paires de base +++ (1 pg correspond à 978 Mb).

Cette taille varie de quelques milliers de paires de bases à plusieurs milliards de paires de bases à plusieurs milliards : Homme : 3,2 milliards de pb, Blé : 16 milliards de pb, Bactéries : quelques millions de pb. Il n'y a pas de corrélation simple entre la taille du génome et la complexité d'un organisme.

### I. Génome procaryote :

Le génome des procaryotes est localisé au niveau d'une région cytoplasmique particulière qu'on appelle le nucléoïde.

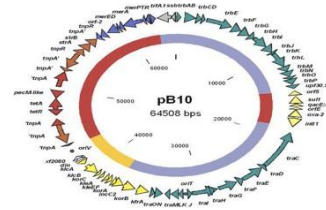
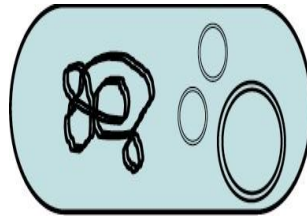
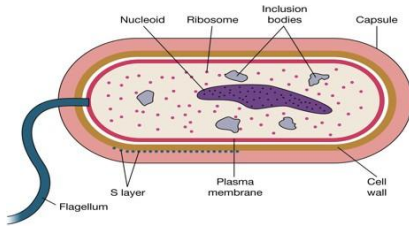
#### I.1. Nucléoïde :

Le nucléoïde est localisé dans une région cytoplasmique de forme irrégulière, habituellement composé d'une seule molécule d'ADN, associée à des ARN et des protéines (surtout des facteurs de transcription).

Il contient le chromosome, une molécule circulaire, unique d'ADN double-brin. Habituellement 1 par cellule. Il n'est pas délimité par une membrane.

Certains procaryotes ont plus de 1 chromosome et d'autres ont des chromosomes composés d'ADN linéaire double-brin.

Certaines cellules procaryotes peuvent aussi contenir une ou plusieurs molécules d'ADN surnuméraires, généralement circulaires, appelées plasmides.



## **I.2. Plasmides :**

-Ce sont des molécules d'ADN extra-chromosomiques, bicaténaires, superenroulées, généralement circulaires. Distinctes de l'ADN chromosomique et ne sont pas nécessaires à la survie de la cellule procaryote.

•**Taille** : de 1 à 400 Méga daltons (kilobases);

**Types** : Plusieurs types de plasmides peuvent coexister dans une même cellule.

**Nombre** : Une copie, pour les grands plasmides et 10 à 20 copies pour les petits plasmides.

**Rôle** : Ils codent pour une information génétique non-indispensable. Et peuvent apporter pour la cellule des avantages évolutifs sélectifs comme la résistance à un antibiotique.

**Propriétés** :

- Douées de répllication autonome, Transmissibles de façon stable à la descendance,
- Certains plasmides sont capables de s'intégrer aux **chromosomes** ; on appelle ces plasmides des **Episomes**.
- Ils peuvent infecter des bactéries ou être échangés entre elles.

## **II. Le génome eucaryote :**

### **II.1. Génome nucléaire :**

- Formé de chromosomes linéaires en nombre variable.
- Le nombre de chromosomes n'est pas relié à la complexité de l'organisme et à la taille du génome.  
Exemple : 46 chromosomes chez l'homme ,16 chez la levure et 6 chez la drosophile.

### **II.1.1. Classification de l'ADN eucaryotique :**

Le génome humain est constitué d'environ 3,5 milliards de paires de bases. Les gènes, dont on estime le nombre entre 50 000 et 100 000, ne représentent que 3 à 5% de l'ADN nucléaire.

Le reste de l'ADN est constitué de séquences de régulation de l'expression des gènes, de séquences répétées et de séquences de fonction inconnue.

#### **II. 1.1.1.ADN répétitif :**

ADN hautement répétitif

ADN moyennement répétitif

#### **II .1.1.1.1. ADN hautement répétitif :**

##### **-Microsatellites**

-Répétition en tandem (4 pb ou moins).

-Au niveau de centaines de milliers de sites le long du chromosome. Se trouvent dans les régions centromériques des chromosomes.

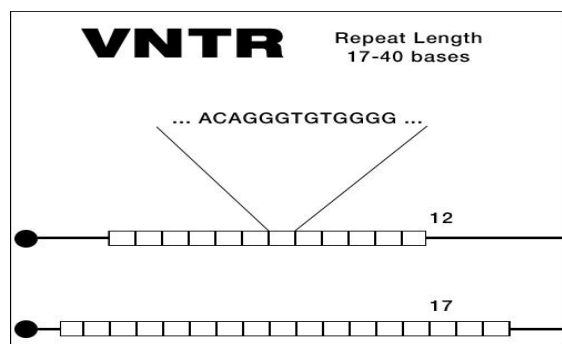


### SÉQUENCES MICROSATELLITES

Répétition en tandem (4pb ou moins) dans cet exemple 2pb

-**Minisatellites** : VNTR : Variable Number of Tandem Repeat

- Il s'agit souvent de séquences répétées en tandem au même site, Le nombre de répétitions varie d'un individu à l'autre et constitue un allèle. Ils sont très polymorphes, tant du point de vue du nombre de répétitions que de la taille des unités répétées (11 à 16 nucléotides).



### -**Microsatellites** :

- Répétition en tandem de centaines de Kb.
- Se trouvent dans les régions centromériques des chromosomes.
- Cet ADN satellite correspond à l'hétérochromatine constitutive.

### II 1.1.1.2. ADN moyennement répétitif :

Les séquences moyennement répétitives sont des séquences nucléotidiques constituées de motifs dont la taille varie de 100 à 1 000 bp, répétées des centaines voire des milliers de fois dans le génome. On distingue deux groupes de séquences moyennement répétées :

- L'ADN répété en tandem : ADN codant
- L'ADN répété dispersé : ADN non codant

#### L'ADN répété en tandem codant : Comporte :

- les gènes de classe I ou les gènes des ARNr 28S, 18S et 5,8S, ces gènes sont regroupés en Clusters de 200 copies, localisés sur les chromosomes :13, 14, 15, 21 et 22.
- les gènes de classe III : codant l'ARNt et l'ARN 5 S. Il y a environ 1200 copies d'ADNt par génome humain.

#### L'ADN répété dispersé non codant :

Comprend les éléments génétiques mobiles capables de transposition à de nouveaux sites. Ce sont des séquences répétées de 100 à 6 000 pb, dispersées dans tout le génome.

Les éléments mobiles les plus abondants peuvent être classés en 2 groupes :

**Petits éléments disséminés ou SINE** : Short Interspersed Repetitive Element.

Ex : séquence Alu : séquences de 300 pb, avec 900.000 copies/génome humain. 5% du génome humain

**Longs éléments disséminés ou LINE** : Long Interspersed Repetitive Element.

Ex : séquence line de 5 à 7 kb, environ 5.000 à 10.000 copies /génome humain

**III. 1.1.2.ADN à copie unique:** constitué de gènes de classe II codant pour des protéines

### **II .1.1.2.1. Définition d'un gène :**

-Unité de l'information génétique. Il est constitué d'une séquence d'acides désoxyribonucléique codant pour la synthèse d'une protéine pour les gènes de classe II et pour un ARN pour les gènes de classe I et III .

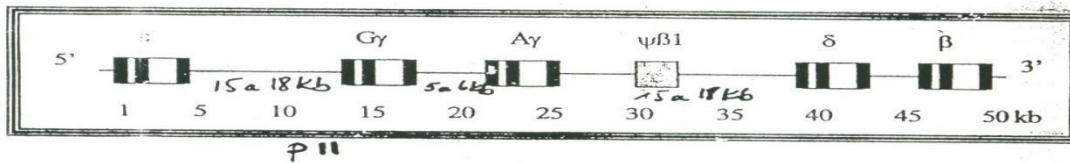
### **II.1.1.2.2. Caractéristiques des gènes :**

-Les gènes sont des entités stables, ils se transmettent des parents à leur descendance sans subir de modification dans leur séquence. Mais occasionnellement, ils peuvent subir des changements dans leur séquences: un tel changement est appelé: **mutation**.

- Un organisme portant le gène normal est dit de type sauvage.

- Un organisme portant le gène altéré est un mutant.

-Pour chaque gène nous héritons de deux allèles, chacun est porté par un des deux chromosomes reçus de nos parents cad par la paire de chromosomes homologues et occupe toujours le même locus.



### **II.1.1.2.3 .Les différents types de gènes de classe II codant des protéines :**

#### **Les gènes uniques ou quasi uniques :**

La très grande majorité des gènes appartient à cette classe.

#### **Les familles de gènes:**

Toute une série de gènes codant pour des protéines grossièrement analogues.

-La famille des gènes globine.

-La famille des gènes actine.

-La famille des gènes myosine

#### **Les gènes domestiques: (House – Keeping genes)**

Ce sont des gènes qui ne s'expriment que dans certains tissus.

Ils codent pour des protéines domestiques comme par exemple les gènes des enzymes de la glycolyse, de la respiration et des métabolismes intermédiaires indispensables à la survie de chaque cellule.

#### **Les pseudogènes**

- Ce sont des séquences nucléotidiques non fonctionnelles, car elles sont ni transcrites ni traduites.
- Leur non fonctionnalité peut résulter :

Soit de l'absence d'un cadre de lecture suffisant (excès de codons stop).

Soit de l'absence de codon Méthionine initiateur ou de région promotrice

#### **II.1.1.2.4. Exemple: Famille des gènes de la $\beta$ globine :**

L'hémoglobine, durant les premiers mois de la gestation est de type embryonnaire  $\zeta_2 \epsilon_2$  (xeta2 epsilon2). Puis dès les premiers mois, des chaînes  $\gamma$  et  $\alpha$  sont synthétisées, ce qui se traduit par l'apparition d'Hb fœtale ( $\alpha_2, \gamma_2$ ). Enfin, immédiatement après la naissance, une diminution brutale de la synthèse des chaînes  $\gamma$  et une augmentation parallèle de la synthèse des chaînes  $\beta$  induisent le remplacement de l'Hb fœtale par l'Hb A ( $\alpha_2, \beta_2$ ). Cette transition Hb F  $\rightarrow$  Hb A porte le nom de switch.

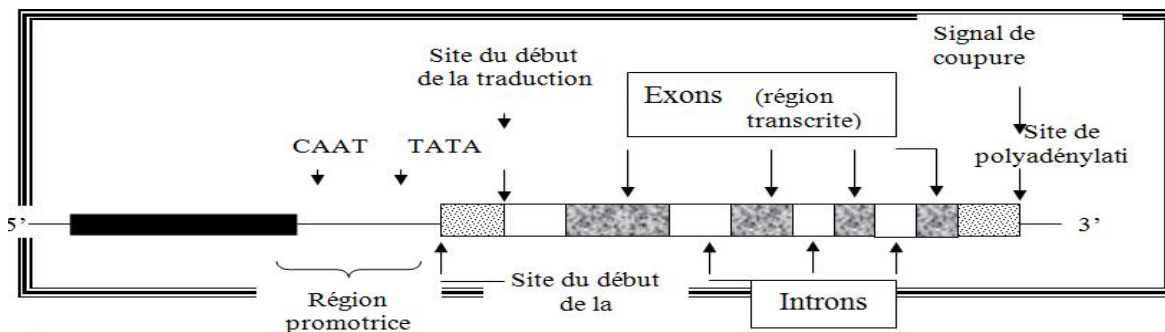
Pour le fœtus, le fait de posséder une hémoglobine différente de celle de sa mère est un avantage; en effet l'Hb F ayant une plus forte affinité pour l'oxygène que l'Hb A, les transferts de l'oxygène entre la mère et le fœtus sont facilités.

### **II.1.1.2.5. Anatomie d'un gène codant pour une protéine ou gène classe II :**

➤ Le gène commence en 5' par des séquences régulatrices non transcrites, dont la présence est nécessaire pour le déroulement de la transcription.

➤ Vers - 100 par rapport au site d'initiation de la transcription commence la région dite "PROMOTRICE".

-Vers - 70 à - 80 se trouve très souvent une séquence CAAT, où se fixent un ou plusieurs facteurs protéiques de la transcription.



-Vers -25 à -30 pb: on retrouve la séquence TATA appelée TATA box où se fixent la RNA polymérase II et certains facteurs de la transcription.

➤ Vient ensuite le site d'initiation de la transcription, la base correspondant à ce site est le plus souvent: une purine, suit une partie non codante de longueur variable et ce jusqu'à la séquence ATG, codon Méthionine qui signale le lieu d'initiation de la traduction.

➤ Suivent ensuite une alternance de séquences présentes (Exons) ou non (Introns) dans la version finale de l'ARNm cytosolique.

- On appelle "**Exon**" toutes les séquences transcrites retrouvées dans le messager cytosolique final.
- On appelle "**Intron**" toute séquence transcrite mais éliminée après par épissage au cours de la maturation du transcrit primaire, donc non retrouvée dans le messager cytosolique final mûr.

Le signal pour l'arrêt de la traduction est donné par l'un des codons stop: UAA, UGA et UAG.

Enfin 10 à 20 bases avant la fin du dernier exon est retrouvée une séquence AATAAA ou séquence de polyadénylation : site de la fixation de la poly A polymérase.

Le gène se termine par une région 3' parfois par des séquences régulatrices.

Les limites des gènes sont imprécises, leurs tailles sont très variables pouvant atteindre jusqu'à 2 millions de pb ; il n'y a pas de relation directe entre la taille de la protéine et la longueur du gène.

### **IV. 1.1.3. ADN intercalaire non classé: espaceur:**

Il constitue 75% du génome humain de fonction inconnue.

### **IV.1. Le génome des mitochondries :**

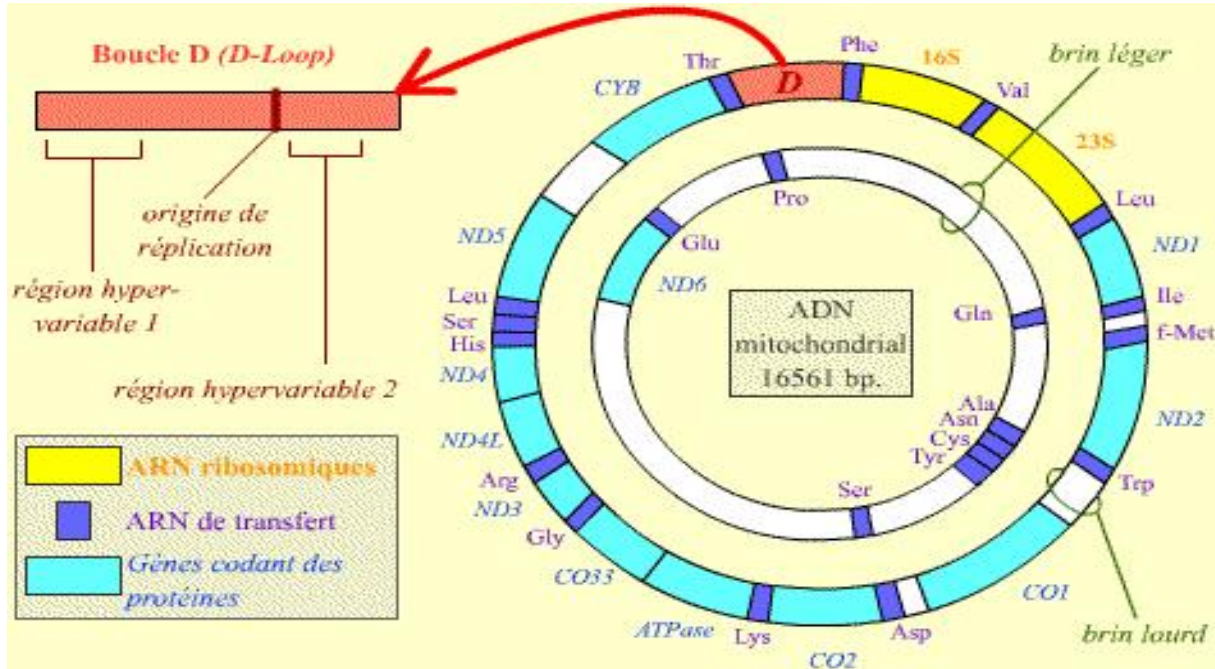
#### **L'ADN mitochondrial**

Bien que chez les eucaryotes, la majeure partie de l'ADN soit présente dans le noyau, une partie de cette ADN est située dans les mitochondries.

- L'ADN mt est situé dans la matrice mitochondriale.
- Il contient 4 à 6 molécules de DNA bicaténaire.
- Organisé en nucléoïde et non associé à des histones
- 1 à 2% de la masse totale d 'ADN de la cellule.

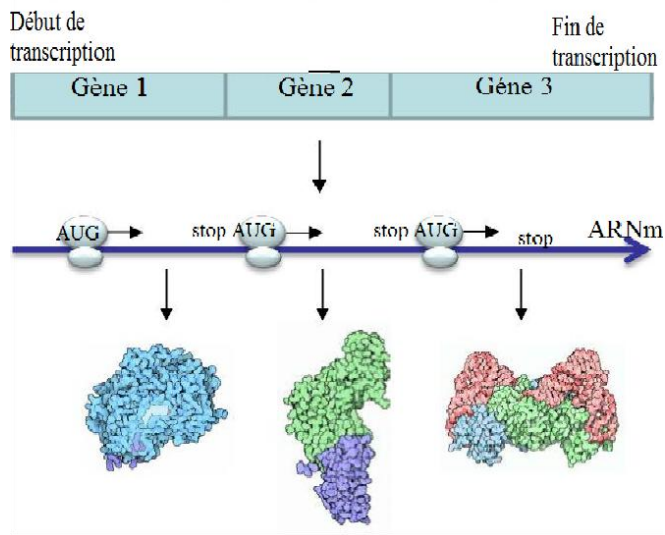


- Il code certaines des protéines constituant le système de phosphorylation oxydative et des ARNt, ARNr et certaines protéines, utilisés dans la synthèse des protéines mitochondriales.
- L'ADN mt humain, une molécule circulaire a été entièrement séquencée, Il contient 16 569 pb
- Contrairement à l'ADN nucléaire, il est dépourvu d'introns et ne contient pas de longues séquences codantes.



### III - ADN procarvotique :

- Les gènes sont **continus (pas d'introns)**.
- Organisés sous forme d'**opérons**.



- Le messenger correspondant est **polycistronique**.

Eucaryotes	Procaryotes
<ul style="list-style-type: none"> <li>• Dans le noyau</li> <li>• Plusieurs chromosomes</li> <li>• G</li> <li>• grande taille &gt;&gt; procaryotes</li> <li>• 98% est non codant</li> <li>• Grande région intergénique de fonction méconnue</li> <li>• Gènes disloqués (exons, introns)</li> <li>• Dans mitochondrie</li> <li>• ADN mitochondriale</li> </ul>	<ul style="list-style-type: none"> <li>• Dans le cytoplasme</li> <li>• Un chromosome circulaire + plasmide</li> <li>• Fraction codante élevé &gt; 90%</li> <li>• Génome compact</li> <li>• Séquence des gènes continue (pas d'introns)</li> <li>• Génome organisé en opérons</li> <li>• Un opéron est une unité d'ADN fonctionnelle regroupant des gènes qui opèrent sous le signal d'un même promoteur.</li> </ul>

### Références bibliographiques

- Précis de Génomique" (2004) G. Gibson & S. Muse - Ed. De Boeck - ISBN : 2-8041-4334-1
- Biologie moléculaire du gène, Watson J. (2012) Pearson éducation
- Biologie moléculaire de Christian Moussard
- Biologie moléculaire de Simon Beaumont.