

الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique et Populaire

وزارة التعليم العالي والبحث العلمي
Ministère de l'enseignement supérieur et de la recherche scientifique



Université des Frères Mentouri Constantine
Faculté des Sciences de la Nature et de la Vie
Département de Biologie Animale

جامعة الاخوة منتوري قسنطينة
كلية علوم الطبيعة والحياة
قسم بيولوجيا الحيوان

Polycopié du cours

Génétique des eucaryotes

Destiné aux étudiants de Licence

Spécialité : Génétique

Docteur SEDRATI ep ZAAF Khadidja

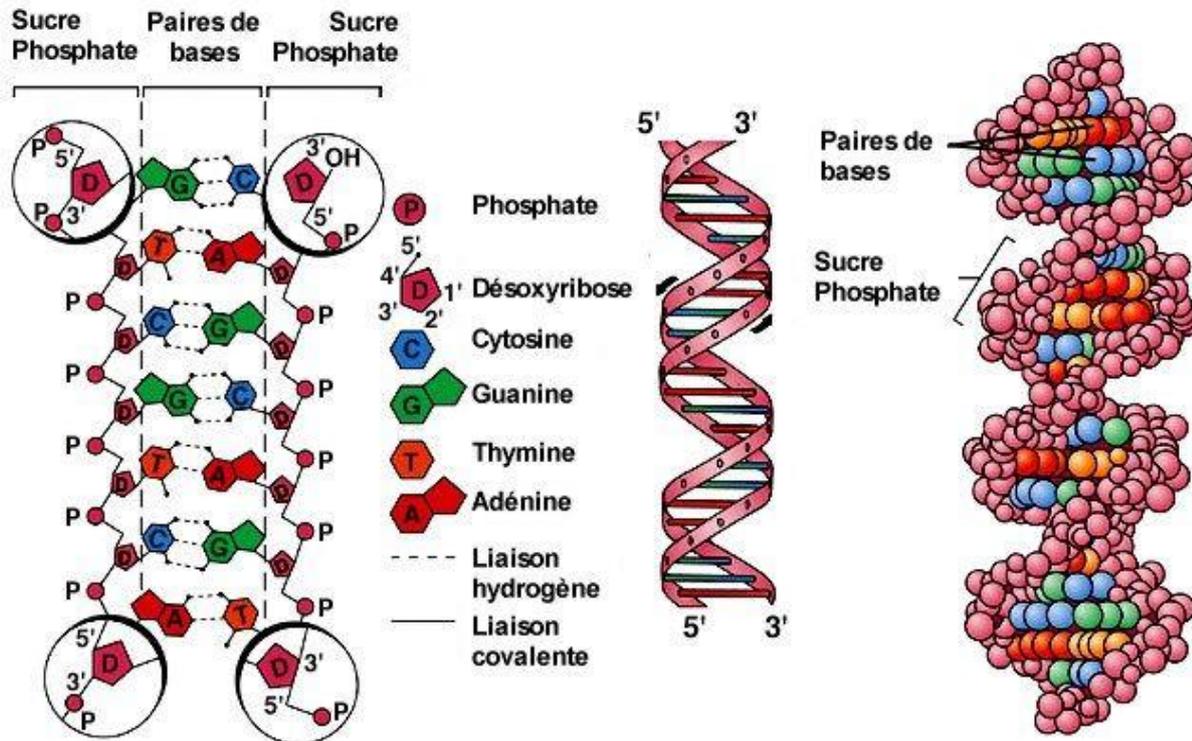
Maitre de conférences classe B

Année universitaire

2020 - 2021

CHAPITRE I : Constitution et dynamique du matériel génétique eucaryote

1. Rappels sur la structure de l'ADN



2. Organisation du Génome eucaryote

Génome : Ensemble de l'information génétique d'un organisme contenu dans chacune de ses cellules sous la forme de chromosomes. Il correspond à l'ADN d'un lot haploïde de chromosome. Il contient les instructions nécessaires au développement, au fonctionnement, au maintien de l'intégrité et à la reproduction des cellules et de l'organisme.

2.1. Taille et constitution des génomes eucaryotes

La taille du génome se mesure en nombre de nucléotides, ou bases. La plupart du temps, on parle de pb (pour paire de bases, puisque la majorité des génomes est constituée de doubles brins d'ADN ou bien d'ARN). On emploie souvent les multiples kb (pour kilobase) ou Mb (mégabase), qui valent respectivement 1 000 et 1 000 000 bases. La taille du génome peut aussi être exprimée en pg (picogrammes), ce qui correspond à la masse d'ADN (haploïde) par cellule. 1 pg représente environ 1 000 Mb.

De manière plus générale, la taille du génome des Eucaryotes s'étend de 10 Mb (Saccharomyces : 12 Mb) à 100 000 Mb (Fritillaire). Certains végétaux ont un génome important (Triticum : 17 000 Mb).

- **La quantité d'ADN n'est pas proportionnelle à la complexité apparente d'un organisme.** Les urodèles, les dipneustes, certaines fougères ou encore certains conifères comme les pins ont des génomes plus de 10 fois plus grands que le génome humain. Ce constat est fréquemment appelé **le paradoxe de la valeur C.**

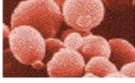
	Taille du génome (nucléotides)	Nbre de gènes (<i>protein-coding</i>)		
	<i>Amoeba dubia</i>	~ 670 000 000 000	?	
	<i>Psilotum nudum</i>	~ 250 000 000 000	?	
	<i>Fritillaria assyriaca</i>	~ 100 000 000 000	?	
	<i>Necturus lewisi</i>	~100 000 000 000	?	
	<i>Homo sapiens</i>	2 900 000 000	23 000	
	<i>Vitis vinifera</i>	487 000 000	30 400	
	<i>Drosophila melanogaster</i>	160 000 000	14 000	
	<i>Arabidopsis thaliana</i>	115 000 000	28 000	
	<i>Caenorhabditis elegans</i>	98 000 000	19 400	
	<i>Saccharomyces cerevisiae</i>	12 500 000	5 800	
	<i>Escherichia coli</i>	4 600 000	4 300	

Figure : représentation de la constitution des génomes de différentes espèces

Définition de la valeur C :

Les premiers biologistes qui mesurèrent la masse d'ADN dans les noyaux des cellules des eucaryotes trouvèrent donc très remarquable que cette valeur soit constante :

1. entre les cellules des différents tissus d'un même organisme.
2. entre les cellules des différents individus d'une même espèce.

On a pris depuis l'habitude de désigner par valeur C, pour valeur constante, la masse de ce que l'on appelle aujourd'hui le génome. L'ADN étant un hétéropolymère linéaire sans ramifications, il y a une relation de proportionnalité entre sa masse et sa longueur, aussi la valeur C est également utilisée pour désigner la taille du génome haploïde.

La taille d'un génome haploïde, exprimée en nombre de paires de bases nucléotidiques (Mb ou million de bases), est constante pour une espèce déterminée. C'est ce qu'on appelle la valeur C

En particulier, la question est de savoir si l'augmentation de la taille du génome est associée à une augmentation de la complexité des organismes. Cette complexité, que l'on peut estimer par le nombre de types cellulaires différents présents dans l'organisme, n'a pu être corrélée de manière significative à la taille du génome. C'est ce qu'on appelle **le paradoxe de la valeur C**.

- **La taille du génome n'est pas directement en relation avec le nombre de gènes** : il y a plus de gènes chez *Caenorhabditis* que chez la Drosophile, alors que son génome est plus petit. Le génome de la levure est 250 fois plus petit que celui de l'homme et pourtant son nombre de gènes n'est pas 100 mais de près de 6000. Le génome humain est 30 fois plus petit que celui de certains végétaux et amphibiens et 200 fois plus petit que celui d'une espèce d'amibe. La teneur en ADN de génomes d'organismes similaires (ex : poissons osseux, amphibiens) peut varier de plusieurs centaines de fois alors que le génome contient grossièrement le même nombre de gènes.
- Des espèces très apparentées et ayant des tailles de génome identique peuvent avoir un nombre très différent de chromosomes de tailles différentes : il n'y a pas de relations simples entre le nombre de chromosomes, la complexité des espèces et la taille totale du génome.

Tableau 01 : comparaison des génomes eucaryote

Organisme	Nombre de paires de bases	Nombre de gènes estimés	Densité génique (Gènes/Mb)
<i>Mycoplasma</i>	0,6.10 ⁶	450	860
<i>E. coli</i>	4.10 ⁶	4 000	950
<i>Saccharomyces cerevisiae</i>	12.10 ⁶	6 000	480
<i>Arabidopsis thaliana</i>	142.10 ⁶	26 000	220
<i>Caenorhabditis elegans</i>	97.10 ⁶	19 000	190
<i>Drosophila melanogaster</i>	137.10 ⁶	14 000	82
<i>Homo sapiens</i>	3.10 ⁶	20 à 25 000	6.25

2.2. Différents classes cinétiques de l'ADN

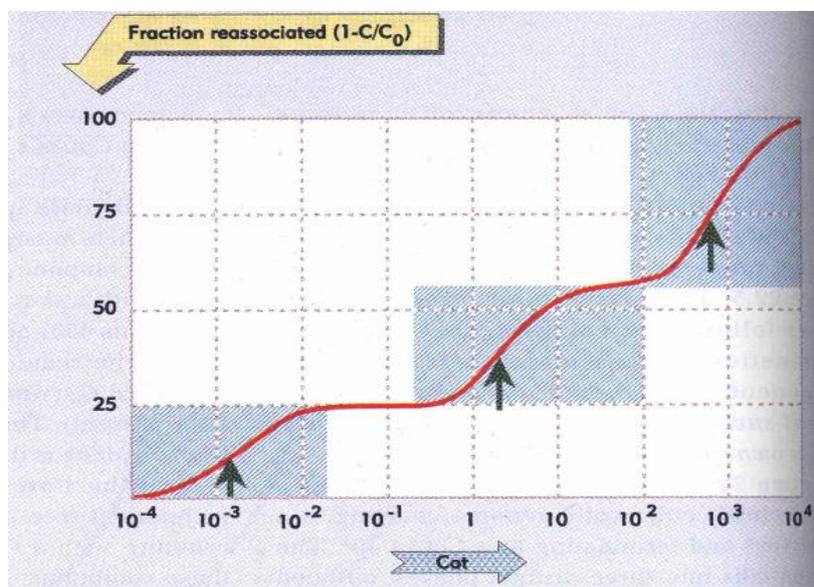
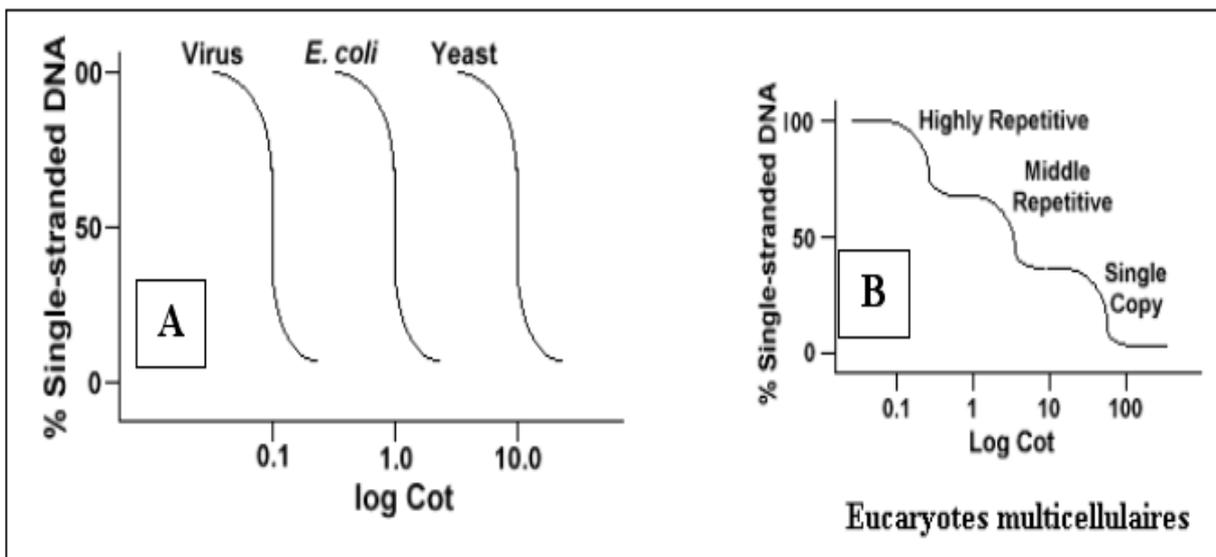
Nous avons vu précédemment qu'il n'existe pas de relation évidente entre la taille des génomes et la complexité des organismes ; c'est ce qu'on appelle « le paradoxe de la valeur C ». Ce paradoxe a été résolu avec la découverte de **l'ADN non-codant**. Les génomes eucaryotes sont, généralement, composés en grande proportion d'ADN qui ne code pas pour des protéines ou pour d'autres éléments fonctionnels. Il est donc possible qu'un génome de taille très importante ne contienne

qu'une très faible proportion de matériel informatif ce qui explique l'absence de lien avec la complexité métabolique, développementale ou comportementale de l'organisme.

2.2.1. Définition de la cinétique de réassociation :

Dans l'expérience suivante, après avoir dénaturé l'ADN génomique d'un organisme, on le place dans des conditions ménagées de refroidissement afin de favoriser la renaturation. On mesure la fraction d'ADN non encore renaturé en fonction du temps.

Sur l'axe des Y on porte le pourcentage d'ADN qui reste simple brin par rapport à la concentration totale d'ADN. L'axe des X porte une échelle logarithmique du produit de la concentration initiale de l'ADN (en moles/litre) par le temps écoulé (en secondes).



Ces courbes sont appelées **courbes de Cot**. Une courbe lisse est l'indice d'une renaturation progressive et régulière de l'ADN. C'est ce qui est observé pour des virus, des bactéries et pour des cellules de levure. Par contre lorsque l'on prend de l'ADN d'organismes multicellulaires (Homo sapiens), l'aspect est totalement différent. La courbe observée semble le résultat de la mise bout à bout de trois courbes. Ce qui indique que dans le premier cas il y avait un seul type d'ADN alors que dans le second trois types d'ADN différents doivent coexister : le premier type se renature très rapidement, le second se renature nettement plus lentement, quant au troisième il ne se renature que très lentement. L'interprétation de cette observation, est que l'ADN qui se renature rapidement n'a aucun mal à retrouver un brin qui lui soit complémentaire parce qu'il contient une (ou des) séquence très fréquente, alors que le troisième type a beaucoup de mal à retrouver son complément parce que cette séquence est " unique ".

2.2.2. Les classes d'ADN chez les eucaryotes :

L'ADN des eucaryotes comporte au moins trois classes : ADN hautement répétitif : réappariement quasiment instantané, ADN moyennement répétitif : réappariement plus lent et ADN à séquence unique : renaturation très lente

- **ADN hautement répétitif (séquence en tandem):** appeler aussi les séquences en tandem. Elles représentent 10% du génome. Ce sont des séquences hautement répétées de quelques bases à quelques centaines de bases qui se suivent en tandem (l'extrémité 3' d'une séquence est suivie par l'extrémité 5' d'une même autre séquence). Elles sont dispersés dans le génome humain (en particulier l'hétérochromatine), mais localisées (toujours positionnées aux mêmes endroits du génome).

a) ADN « satellite » (environ 15% de l'ADNg)

Représenter essentiellement par :

-Séquence CEN (centromérique) appelé aussi « α satellite) Motif de 171 paires de base répété n fois, riche en G-C, dans les centromères. C'est l'ADN correspondant aux bandes C. Cette région servirait d'ancrage aux protéines du kinetochore pour fixer le chromosome aux microtubules du faisceau mitotique.

-Séquence TEL (télomérique) : le motif TTAGGG est répété des milliers de fois. Le motif est rajouté à l'extrémité des chromosomes par la télomérase pour protéger les télomères vis-à-vis des nucléases.

b) Mini-satellite ou VNTR (*Variable Number of Tandem Repeat*)

Ce sont des séquences de 11 à 16 pb répétées en tandem avec un nombre de copie variable d'un individu à l'autre. Il peut y avoir jusqu'à 1000 répétitions de suite. Les séquences sont dispersées dans tout le génome (surtout dans les télomères) excepté sur le chromosome X ou Y. Ils sont très polymorphes car polyallélique (jusqu'à 1000 allèles différents par VNTR)

c) Microsatellite ou STR (*Short Tandem Repetition*)

Ce sont des motifs très court (de 1 à 4 pb) répété en tandem jusqu'à 40 fois. Ils sont très abondants (environ 130.000). Ils sont répartis de façon uniforme sur tous les chromosomes, et sont très polymorphes comme les VNTR. Ils proviennent sans doute d'un dérapage de l'enzyme de réplication.

• ADN moyennement répétitif (Séquences répétées dispersées)

Ce sont des séquences apparentées, réparties sur l'ensemble du génome : Il représente 25 à 40% du génome humain. On distingue :

a) ADN non codant :

Séquences qui sont des retrotransposants (élément génétique mobile et retro : utilisation de la transcription inverse)

-SINE: *Short Interspersed Repetitive Element*. Ce sont des séquences non-autonomes, faisant appel à des protéines synthétisées par d'autres gènes.

Exemple : Famille Alu : séquences de 300pb dispersées dans tout le génome avec 900.000 copies/génome humain.

-LINE : *Long Interspersed Repetitive Element*: Ce sont des séquences autonomes : elles sont aptes à « sauter » grâce à des protéines qu'elles codent elles-mêmes : une rétrotranscriptase (=transcriptase inverse) et une endonucléase, ce qui permet la formation de copies intermédiaires d'ADNc (complémentaire). Ces copies d'ADNc sont capables d'être transposées et intégrées à d'autres régions du génome riches en séquences AT.

b) ADN codant :

Les gènes ribosomiques (ADNr vers ARNr). Ces gènes sont regroupés en Clusters de 200 copies d'une unité de transcription, et ceci sur 5 chromosomes différents. L'ARN polymérase 1 donne un ARN unique de 45S qui donnera ensuite 3 ARNr de 28S, 18S et 5,8S. Ces Clusters d'ADNr

correspondent aux constriction secondaires des chromosomes et aux nucléoles dans la chromatine interphasique.

Les gènes pour ARNt. Ils sont aussi regroupés en Clusters avec des répétitions en tandem. Il y a environ 1.200 copies d'ARNt par génome humain.

- **Les séquences uniques :**

Consiste en séquences qui sont présentées en un seul exemplaire dans le génome. Cet ADN comprend les séquences qui codent les protéines (les gènes) mais aussi une certaine quantité de séquences très conservées, non codantes, réparties sur tout le génome et qui sont souvent situées près des régions régulatrices de gènes des fonctions essentielles de la cellule (gènes du développement, ARNr...). Elles ont un rôle sans la régulation de l'expression des gènes et dans l'appariement correct des chromatides lors de la division cellulaire.

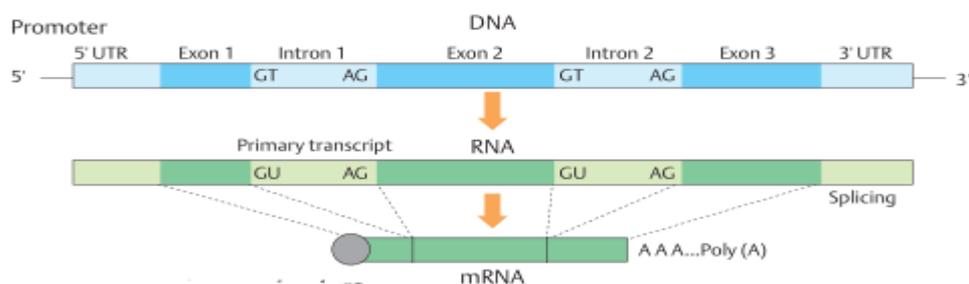
2.3. Structure du gène eucaryote

1.3.1. Définition des gènes :

Ce sont des séquences de nucléotides de l'acide désoxyribonucléique contenant l'information nécessaire pour la synthèse d'un acide ribonucléique (transcription) ou d'une protéine (transcription-traduction). Les gènes sont situés sur les chromosomes, dans le noyau de la cellule. Les gènes sont dispersés et séparés par de l'ADN intergénique

1.3.2. Structure des gènes :

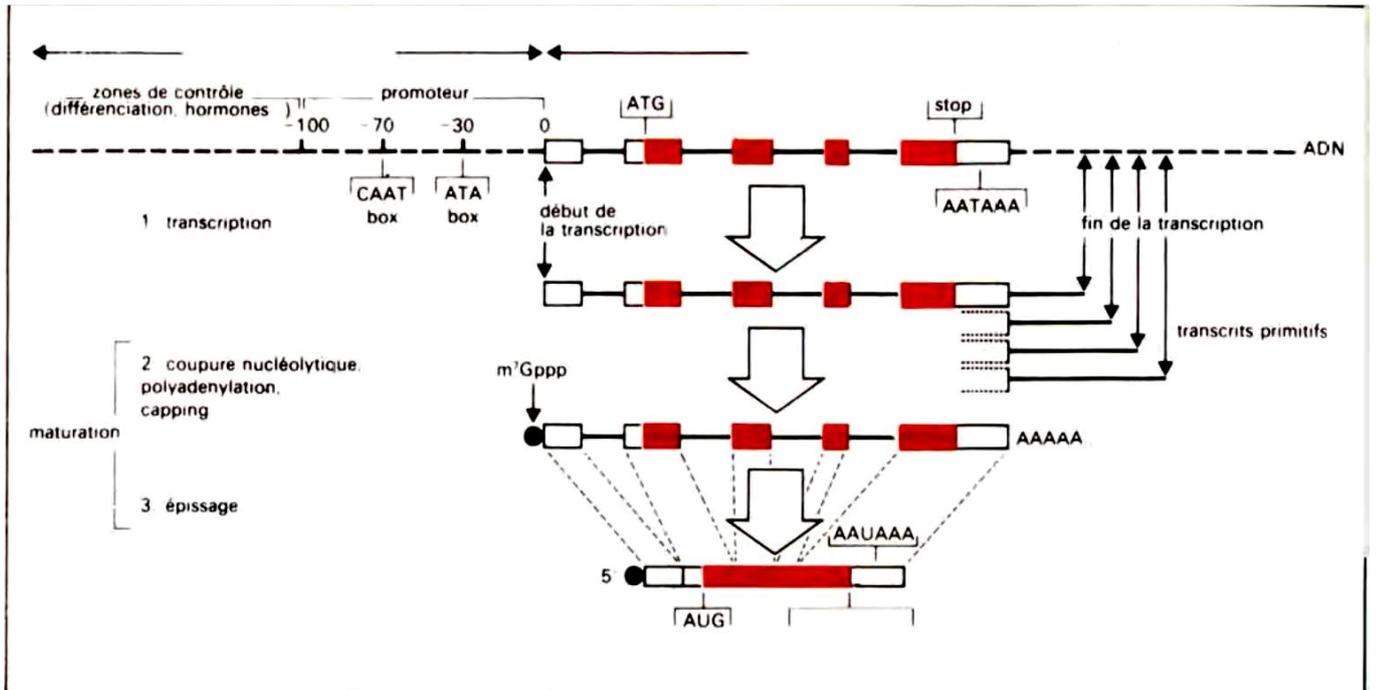
Un gène codant pour une protéine commence au site de départ de la transcription et termine par des séquences qui seront transcrites mais pas traduites : Ceux-ci sont appelés la région 5 'non traduite (5' UTR) et la 3 'UTR à l'extrémité 3'. A quelques dizaines ou quelques centaines de paires de bases en aval, le codon d'initiation ATG est le site de départ de la traduction. Puis vient une succession de séquences tantôt codantes, les exons qui seront transcrits et traduits, tantôt non codantes, les introns (qui séparent deux exons) qui seront transcrits mais pas traduits.



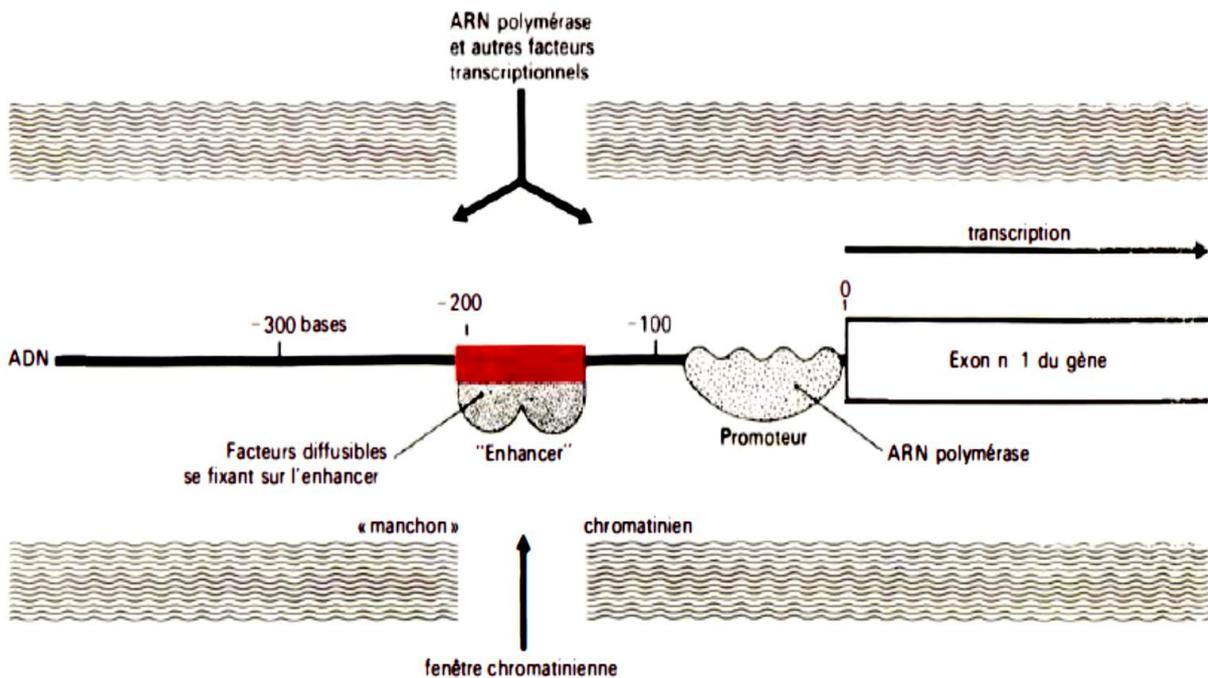
Les exons et les introns sont numérotés dans la direction 5 'à 3' du brin codant. Les deux exons et les introns sont transcrits en un ARN précurseur (transcription primaire). Les segments non codants (introns) sont retirés du transcrit primaire et les exons de chaque côté sont connectés par un processus appelé épissage. L'épissage doit être très précis pour éviter une modification indésirable du cadre de lecture correct. Les introns commencent presque toujours par les nucléotides GT dans le brin 5 'à 3' (GU dans l'ARN) et se terminent par AG. Les séquences à l'extrémité 5 'de l'intron commençant par GT sont appelées site donneur d'épissage et à l'extrémité 3', se terminant par AG, sont appelées site accepteur d'épissage. A la fin du dernier exon ,un codon stop (TAA,TAG ou TGA) est le site de fin de la traduction.

En amont ,coté 5'du gène ,se trouve **le promoteur** : cette séquence ,d'une centaine de paires de bases ,définit le site de départ de la transcription et sa direction, plus exactement le promoteur est un segment d'ADN en amont de la séquence codante comportant le site de fixation de l'ARN polymérase ainsi que les sites de fixation des protéines régulatrice de la transcription. Il comporte des séquences très conservées:

- **La boîte TATA** : Elle est située à environ -25 paires de bases de l'origine de la transcription. C'est une séquence de six nucléotides riches en A et T. La séquence dite consensus (statistiquement la plus rencontrée) est TATAAA.
- **La boîte GC** : (située le plus souvent dans la région entre -110 et -40). Elle peut se présenter sous forme d'hexanucléotides: 5'-GGGCGG-3'. Le motif riche en bases G et C peut être répété plusieurs fois.
- **La boîte CCAAT** : (souvent située dans la région entre -120 et -80). Cette boîte peut être située avant ou après une boîte GC ou même entre deux boîtes GC.



Par ailleurs, des séquences régulatrices sont présentes à des distances très variables du promoteur. ces séquences constituent des éléments fondamentaux de la régulation de l'expression du génome des eucaryotes, leur importance fonctionnelle venant de leur propriété d'être régulées par des facteurs spécifiques de la différenciation tissulaire, voire par des hormones. il s'agit des séquences extinctrices (ou inhibitrices) appelées « silencers » ou des séquences stimulatrices les dits « enhancers ».



1.3.3. Classification des gènes : Les gènes sont classés selon

- **la complexité :**

- a) Gène unique
- b) Famille de gènes
- c) Super famille de gènes

- **le tissu d'expression :**

- a) Gène spécifique
- b) Gène domestique qualifie un gène qui s'exprime dans tous les types cellulaires et dont les produits assurent les fonctions indispensables à la survie des cellules. Ils ne subissent donc pas de régulation.

- **la fonction :**

- a) Gène actif
- b) Pseudogènes Sont des gènes non fonctionnels (par absence de région promotrice, de codon d'initiation, d'un cadre de lecture suffisant ...). Il existe 2 sortes de pseudogènes :
 - Les gènes dupliqués (mal dupliqués ou mutés au cours de l'évolution : ils sont pourvus d'introns)
 - les gènes rétrotranscrits à partir d'ARNm (les rétropseudogènes n'ont ni promoteur, ni introns, mais une séquence poly (A)).

Comme ils ne sont pas fonctionnels, les pseudogènes ne subissent aucune pression de sélection et les mutations s'y accumulent en grand nombre.

- **le type de protéine :**

- a) Gène de structure : est un gène qui détermine la séquence d'acides aminés d'une protéine
- b) Gène de régulation : ou gène de contrôle est un gène qui possède la capacité de contrôler le fonctionnement des gènes de structure. C'est ainsi qu'il élabore selon les cas, un répresseur ou un inducteur de façon à freiner au contraire déclencher voir accélérer la transcription.

- **la transcription :**

- a) Gènes de classe I : ARNr
- b) Gènes de classa II : ARNm
- c) Gènes de classa III : ARNt

1.3.4. Famille de gènes :

Ensemble de gènes, homologues par la structure, la séquence et la fonction de leurs produit, qui ont évolué à partir d'un gène ancestral commun.

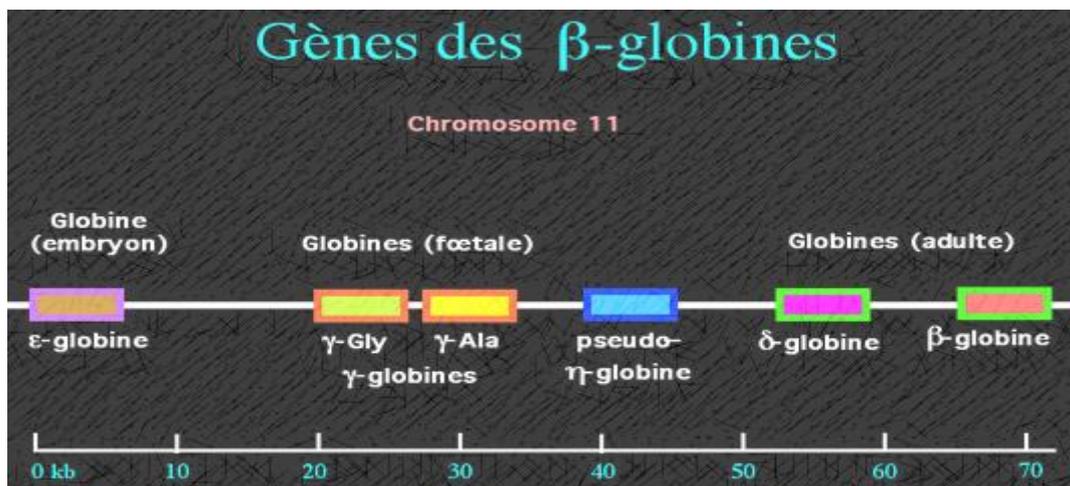
Une forme particulière d'insertion a joué un rôle majeur dans l'évolution des espèces d'eucaryotes : la duplication des gènes. Lorsque la séquence d'un gène a été répétée deux fois dans le même ADN et que les deux gènes continuent à s'exprimer, ils évoluent différemment pour aboutir au bout de

nombreux millions d'années à exprimer des protéines différentes, ce qui peut apporter un avantage sélectif à l'espèce.

Dans les gènes ou dans les protéines qui sont issus d'un gène ancestral unique, on retrouve toujours une homologie de structure (exons, introns) de séquences primaires (DNA, protéines) et de fonctions (activités des protéines). Ces gènes constituent ensemble une famille de gènes dont on peut reconstruire l'arbre phylogénétique en suivant l'évolution du gène ancestral et de ceux qui en sont issus à travers le temps et l'évolution des espèces.

Si des molécules produites se ressemblent tout en ayant des fonctions identiques, c'est qu'elles sont homologues dans la même espèce et le même individu. Ces molécules sont alors apparentées et les gènes qui codent pour elles sont homologues dans une même espèce. Ces gènes ne s'expriment pas au même moment de la vie de l'individu mais existent simultanément dans le génome : il s'agit d'une famille multigénique. On parlera de famille multigénique pour décrire une famille de gènes présentant plusieurs membres dans un même génome.

Gènes des β -globines :

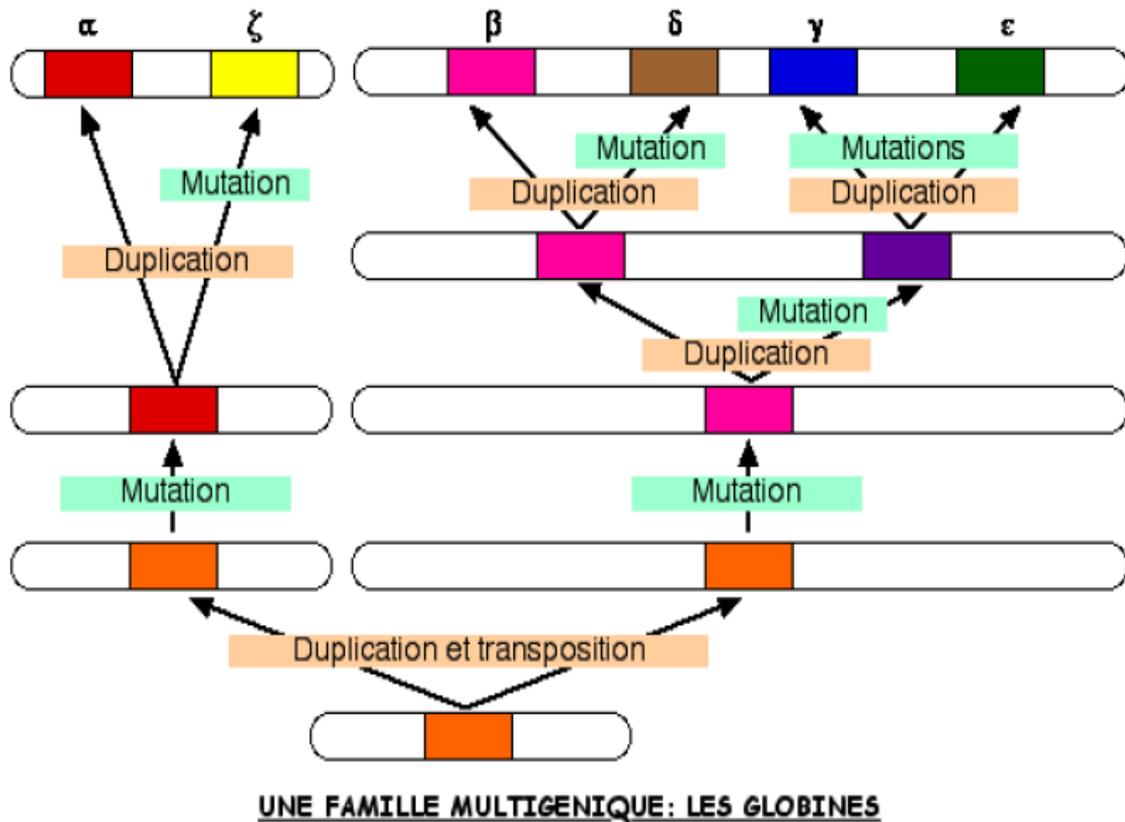


Plusieurs gènes existent pour exprimer les chaînes d'acides aminés qui constituent l'hémoglobine : α -globines sur le chromosome 16 et β -globines sur le chromosome 11. Sur ce dernier chromosome il existe cinq gènes exprimés successivement au cours du développement de l'individu dans les hémoglobines embryonnaires, fœtales et adultes.

L'ensemble des cinq gènes constitue un groupe de gènes (cluster). Les cinq gènes sont dérivés dans l'évolution à partir d'un gène ancestral unique (chez les Invertébrés) par duplications successives.

Le gène ϵ -globine est à l'extrémité 5' du groupe de gènes. Après un long intergène, on rencontre les deux gènes des γ -globines (γ -Gly et γ -Ala). Dans l'intergène suivant se trouve un pseudogène η -

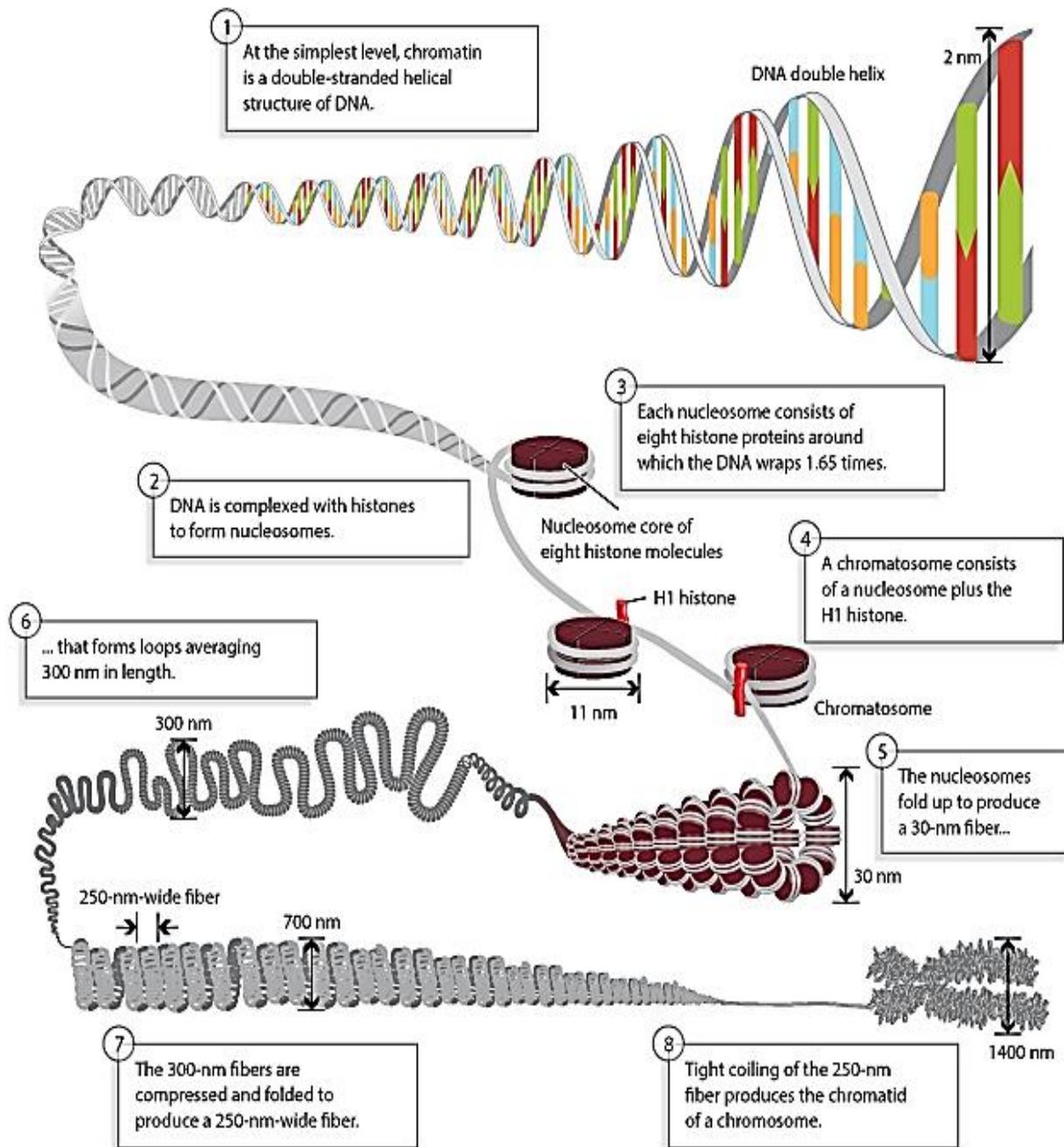
globine, qui n'est plus exprimé. Enfin vers l'extrémité 3' se rencontrent suc-cessivement les deux gènes exprimés chez les adultes δ -globine et surtout β -globine



1.4. Architecture du génome eucaryote dans le noyau

Dans les cellules eucaryotes plus précisément dans les noyaux, le matériel génétique est organisé en une structure complexe constituée d'ADN et de protéines. Cette structure a été baptisée **chromatine**. La masse de la chromatine contient environ deux fois plus de protéine que d'ADN. Les protéines sont deux types : **les histones** (protéines basiques présents en quantité à peu près égale à celle de l'ADN) et **les protéines non histones** (acides représentant entre 10 à 30% de l'ensemble). L'ARN représente environ 10% de la masse d'ADN, la majorité est représenté par des chaînes d'ARN messager naissantes qui sont en transcrit vers le cytoplasme où ils seront traduits.

- **Structure de la chromatine**



Au sein du noyau interphasique, la chromatine est organisée en deux territoires fonctionnels.

En fonction de caractères morphologiques (en microscopie électronique notamment) et de propriétés fonctionnelles la chromatine a été divisée en :

- Euchromatine
- Hétérochromatine

L'hétérochromatine a été définie comme une structure qui ne change pas d'état de condensation au cours du cycle cellulaire tandis que **l'euchromatine** apparaît décondensée pendant l'interphase.

L'hétérochromatine est localisée principalement en périphérie du noyau et du nucléole tandis que l'euchromatine est répartie à l'intérieur du nucléoplasme.

On distingue :

- **L'hétérochromatine constitutive** qui contient peu de gènes, formée principalement de séquences d'ADN satellite qui peuvent se replier sur elles-mêmes et pourraient jouer un rôle important dans la structure très compacte que présente l'HC constitutive et dont les plus grandes régions sont situées à proximité des centromères et des télomères des chromosomes. L'HC constitutive est stable et garde ses propriétés d'hétérochromatine à toutes les étapes du développement et dans tous les tissus. L'HC constitutive est très polymorphe probablement du fait de l'instabilité de l'ADN satellite. Ce polymorphisme peut concerner aussi bien la taille, que la localisation de l'hétérochromatine et n'entraîne apparemment aucun effet phénotypique.
- **L'hétérochromatine facultative** qui contient des régions codantes pouvant adopter les caractéristiques structurale et fonctionnelle de l'hétérochromatine, comme le chromosome X inactif chez la femelle des mammifères. L'HC facultative est caractérisée par la présence de séquences répétées de type LINES qui pourraient favoriser la propagation d'une structure chromatinienne condensée. L'HC facultative est réversible, son « état hétérochromatique» dépendant du stade de développement ou du type cellulaire étudié. L'X inactif (Corps de Barr) dans les cellules somatiques femelles et la vésicule sexuelle (VS) inactive au stade pachytène de la méiose masculine sont deux exemples d'HC facultative. L'HC facultative, n'étant pas particulièrement enrichie en ADN satellite, ne présente pas de polymorphisme.

Tableau : Propriétés permettant de différencier l'hétérochromatine constitutive de l'hétérochromatine facultative.

HC constitutive	HC facultative
stable	réversible
contient de l'ADN satellite	enrichie en séquences LINES
polymorphisme +	polymorphisme -
bandes C +	bandes C -

INSERM U491

- **Structure des chromosomes**

Les chromosomes sont des structures permanentes du noyau des cellules, bien que n'étant pas toujours visibles. Ils existent sous deux états :

- **l'état filamenteux**, lorsque la cellule n'est pas en division (interphase) ; le matériel chromosomique apparaît alors formé de nombreux filaments enchevêtrés (ou nucléofilaments) constituant la chromatine ;
- **l'état compact**, pendant la division cellulaire ou mitose ; les filaments chromatiniens y atteignent leur condensation maximale, ce qui rend les chromosomes visibles : ce sont les chromosomes métaphasiques.

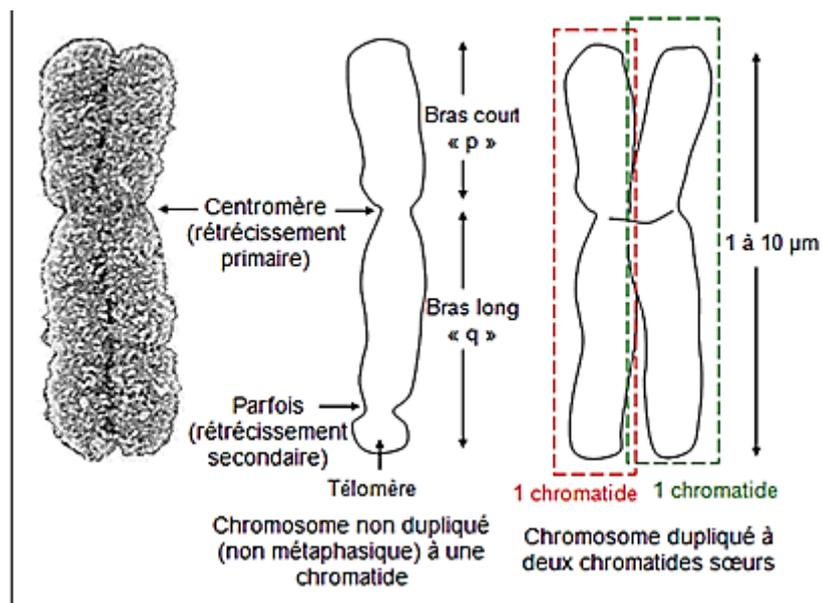


Figure : structure des chromosomes métaphasiques

Chapitre 2 : la réplication chez les eucaryotes

Introduction

La molécule d'ADN doit sa pérennité et la transmission de l'information qu'elle contient au fil des divisions cellulaires à un mécanisme hautement complexe et très reproductible qu'est la réplication (ou duplication). Ce phénomène universel est très spécifique à la fois chez les organismes eucaryotes et procaryotes. Il assure une fidèle duplication du support de l'information génétique qu'est la molécule d'ADN.

Ce mécanisme hautement complexe implique une machinerie enzymatique très spécifique indispensable au processus de la vie des cellules. Cependant, au cours de la vie de la cellule, l'ADN est soumis à de nombreuses modifications comme des lésions et des mutations (deux phénomènes distincts) qui modifient la fidélité de réplication.

Pour pallier à ces perturbations, les cellules possèdent en complément des systèmes de réparation qui permettent la plupart du temps de rétablir l'information génétique originelle lors de la réplication.

1- Caractéristiques :

1-1- La réplication est semi-conservatrice :

À partir d'une molécule d'ADN et de nucléotides séparés, on obtient deux molécules d'ADN parfaitement identiques. Chacune des deux nouvelles molécules obtenues est formée d'un brin de la molécule d'origine et d'un nouveau brin assemblé à partir des nucléotides ajoutés. C'est ce qu'on a appelé un mode de reproduction **semi-conservatif**.

Depuis *Watson et Crick* (1953), on sait que l'ADN est une molécule formée de deux brins antiparallèles, formant une double hélice. Dès leur publication originale sur la structure de l'ADN, *Watson et Crick* ont proposé que cette double hélice puisse s'ouvrir, permettant ainsi la synthèse de nouveaux brins, complémentaires des brins originaux. Ce sont les expériences de *Matthew Meselson* et *Franklin Stahl* qui confirmeront le modèle semi-conservatif imaginé par *Crick et Watson*. Cette expérience date de 1958. Elle permet de démontrer le caractère semi-conservatif de la multiplication de la molécule d'ADN chez les bactéries. Cette expérience a pu être réalisée grâce à plusieurs mises au point techniques :

1 - *Meselson et Stahl* mettent au point une technique d'obtention de gradient de densité par centrifugation. En utilisant du chlorure de Césium de densité moyenne 1,72, ils obtiennent après

24h de centrifugation à grande vitesse un gradient de densité (environ de 1,70 à 1,75), gamme qui englobe la densité de l'ADN (1,710).

2 - Ils cultivent les bactéries dans un milieu dans lequel les substances organiques utilisées comme source d'azote contiennent de l'azote lourd (N^{15}). Au cours de la culture, toutes les molécules azotées et en particulier l'ADN contiennent une forte proportion d'azote N^{15} . L'ADN "lourd" a une densité de 1,724 et peut être distingué de l'ADN "léger" (1,710).

3 - Ils mettent au point une méthode qui permet de synchroniser pendant quelques générations la division des bactéries.

Meselson et Stahl faisaient se reproduire des bactéries dans un milieu contenant des ions ammonium (NH_4^+) constitués d'azote 14 (isotope léger) ou d'azote 15 (isotope plus lourd). Les bactéries fabriquent leurs nucléotides à partir de ces ions ammonium et les utilisent pour synthétiser leur ADN.

- **Hypothèses :**

Pour expliquer la duplication d'un ADN bicaténaire, trois modèles ont été proposés. Ces modèles se basent tous sur l'utilisation de la molécule d'ADN "mère" comme matrice pour sa réplication, mais selon des modalités différentes :

Légendes : en rouge : molécules d'ADN "mère" et son devenir en bleu : ADN néo-formé		
Hypothèse 1: modèle conservatif A partir d'une molécule d'ADN bicaténaire "mère", on forme une nouvelle molécule d'ADN bicaténaire. On garde donc ici une molécule "mère", non modifiée (elle est donc conservée), tout en "créant" une nouvelle molécule ("fille").	Hypothèse 2: modèle semi-conservatif On dissocie les deux brins de la molécule d'ADN bicaténaire "mère". Chaque brin sert donc de matrice à la synthèse d'un brin complémentaire, l'ensemble reformant une molécule d'ADN bicaténaire. Chaque nouvelle molécule "fille" ne conserve donc que la moitié de la molécule "mère".	Hypothèse 3: modèle dispersif On ne conserve aucun brin intact. La copie se réalise par fragments dispersés dans l'ensemble de l'ADN, permettant de former les deux molécules d'ADN bicaténaires "filles".

Des bactéries cultivées depuis longtemps en présence de molécules azotées ^{15}N sont repiquées sur un milieu contenant des molécules azotées ^{14}N et permettant la synchronisation des divisions. Des fractions sont prélevées après différents temps correspondant à 1, 2, 3, ... divisions. L'ADN est extrait, placé dans la solution de chlorure de Césium et centrifugé 24h à 100.000 g. La position des ADN est repérée par une mesure de la densité optique.

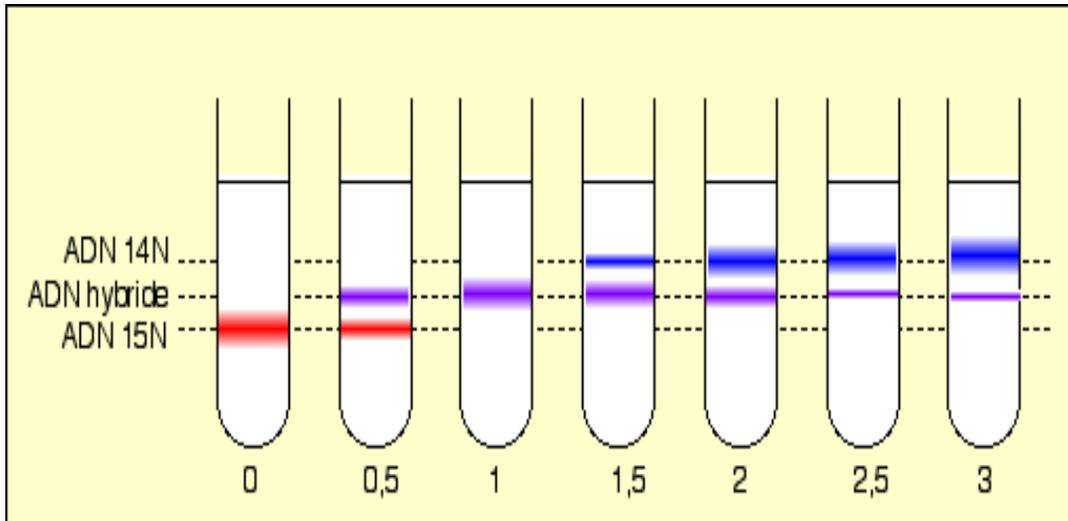
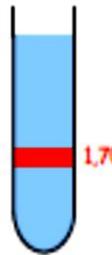
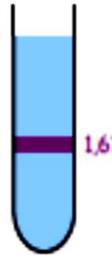
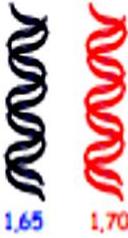
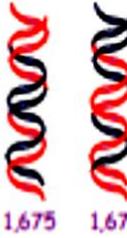
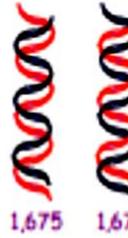
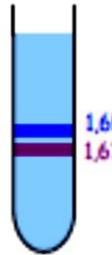
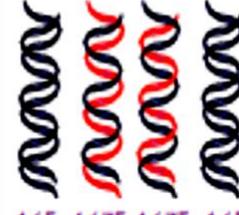
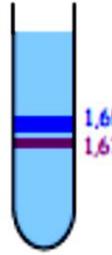
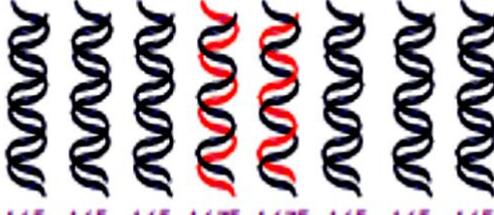


Figure : Position des différentes bandes d'ADN au cours du temps. Les chiffres donnent le nombre de divisions.

Après 1 génération, tout l'ADN est hybride (du point de vue de sa densité). Il n'y a plus d'ADN N^{15} . Ensuite, l'ADN hybride disparaît progressivement au profit d'ADN "léger" (N^{14}).

Interprétation

T	Résultats	Mode conservatif	Mode dispersif	Mode semi-conservatif
t0				
t30		 1,65 1,70 PAS POSSIBLE!	 1,675 1,675 POSSIBLE	 1,675 1,675 POSSIBLE
t60			 1,665 1,665 1,665 1,665 PAS POSSIBLE!	 1,65 1,675 1,675 1,65 POSSIBLE
t90			 1,65 1,65 1,65 1,675 1,675 1,65 1,65 1,65 TOUJOURS POSSIBLE!	

En conclusion, seul le modèle semi-conservatif permet d'aboutir à des résultats attendus correspondant aux résultats observés.

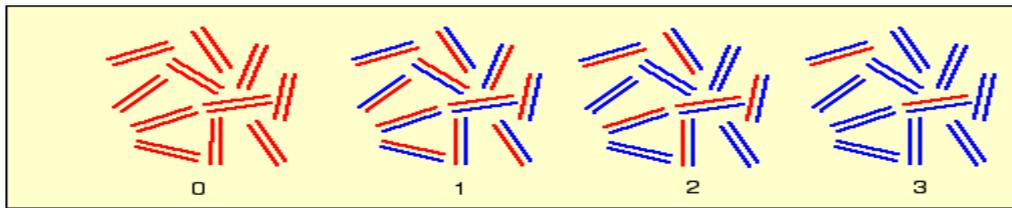
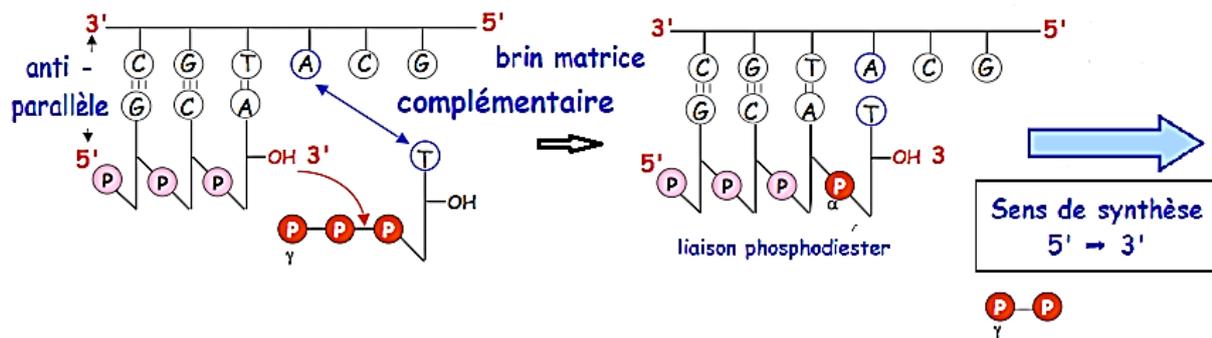


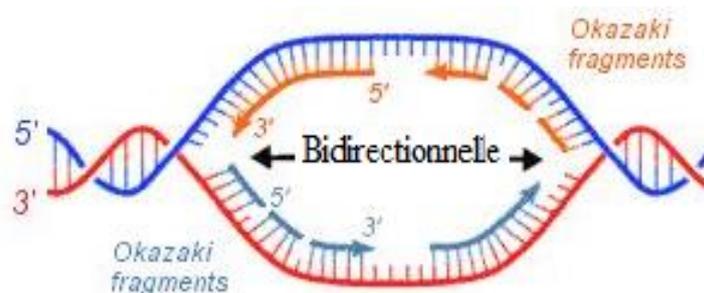
Figure : Représentation schématique de la population de fragments d'ADN au cours des générations. Après 1 génération tout l'ADN est "hybride" et constitué d'un brin "lourd" (N^{15}) et d'un brin "léger" (N^{14}).

Synthèse de l'ADN dans un seul sens ($5' \rightarrow 3'$) :



1-2- La réplication bidirectionnelle et semi-continue :

La réplication d'un ADN bicaténaire est assurée par un complexe enzymatique se déplaçant sur la molécule dans une direction donnée. L'antiparallélisme des deux brins impose un mode de réplication semi-discontinue : l'un des brins est synthétisé d'un seul tenant de $5'$ vers $3'$ en partant d'une unique amorce (c'est le brin à synthèse continue); l'autre brin est synthétisé sous forme de fragments (fragments d'Okazaki), chacun nécessitant un amorçage spécifique (c'est le brin à synthèse discontinue). Au cours de la réplication, les fragments d'Okazaki adjacents sont progressivement jointés et ligaturés, et forment finalement un brin d'ADN continu.



2- Le mécanisme de réplication :

2-1- Changement de forme de l'ADN

La molécule d'ADN doit être accessible aux enzymes chargées de sa réplication. Or l'ADN en son état naturel est surenroulé. Par conséquent, cet état doit être modifié pour laisser l'accès aux enzymes responsables de la réplication. Les enzymes qui assurent ces modifications de l'état d'enroulement de l'ADN sont les topoisomérases. Elles déstructurent et restructurent la forme en isomérisant l'ADN. On distingue les topoisomérases I et II.

La topoisomérase de type I ne coupe que l'un des deux brins de l'ADN en se liant par une des tyrosines au phosphate en 5' libre de l'ADN coupé. La molécule d'ADN peut ensuite librement se dérouler et l'énergie de la liaison phosphodiester libérée est transférée lors de la ligation de l'ADN qui suit le déroulement, ce qui explique que l'énergie sous forme d'ATP n'est pas toujours nécessaire.

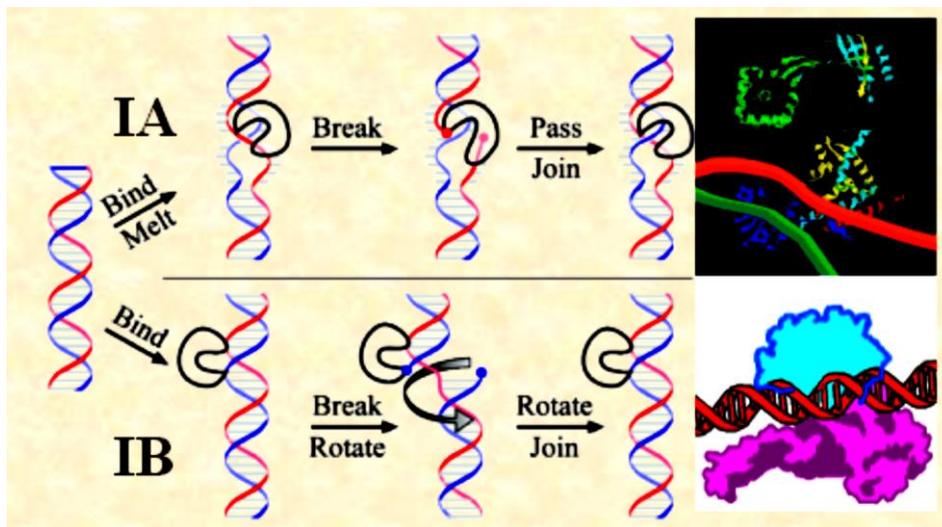


Figure : rôle des topoisomérase I

La topoisomérase de type II agit spécifiquement en coupant les deux brins de l'ADN. Contrairement à la précédente, elle est capable de créer ou de retirer des supertours en consommant de l'énergie.

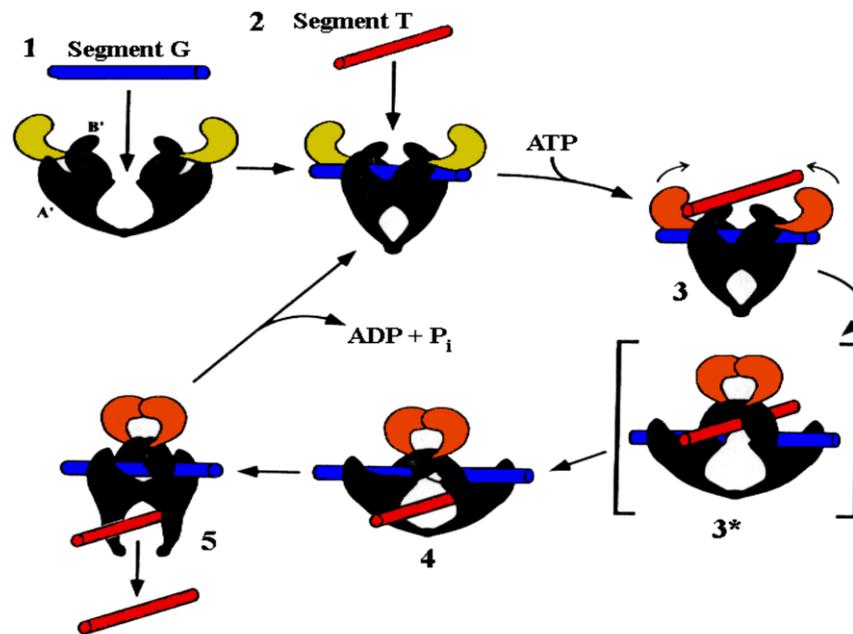


Figure : Rôle des topoisomérase II

Les topoisomérases ne sont pas capables de produire les supertours observés lors de la formation des nucléosomes. En réalité elles jouent un rôle de relâchement dans les boucles pour favoriser la transcription.

Il est donc important de dérouler l'ADN et de le maintenir sous forme simple brin. Les deux brins sont séparés au moyen d'enzymes que sont **les hélicases** (ou déroulases) qui s'associent sur les brins d'ADN, coupent, déroulent et réassocient. Ces activités sont énergie-dépendantes et utilisent de l'ATP. Il existe plusieurs types d'hélicases qui se fixent soit sur le brin 3'-5' soit sur le brin 5'-3' (hélicases II et III).

Les brins d'ADN séparés sont stabilisés sous forme simple brin au moyen de protéines, Les **protéines RPA** (*replication protein A*)

2-2- Initiation de la réplication

Chez les eucaryotes, il y a plusieurs origines de réplication : de 20 000 à 100 000 origines de réplication par cellule. Ces origines de réplication ne sont pas toutes activées en même temps, mais par petits groupes de 20 à 80 OR. Ces groupes constituent **une unité de réplication**.

Ces unités de réplication sont activées de manière **asynchrone**. Selon la localisation, la réplication s'effectue plus ou moins tardivement : les unités de réplication dans l'euchromatine (chromatine

décondensée) sont activées précocement alors que les unités de réplication dans l'hétérochromatine (chromatine condensée) sont activées plus tardivement.

Sur un chromosome, on retrouve donc plusieurs milliers d'OR, à partir desquelles il y a les deux fourches de réplication. A partir de chaque point de réplication, quand il y a activation, il va donc y avoir un œil de réplication qui va se former, avec les deux fourches qui vont progresser dans des sens opposés (bidirectionnel).

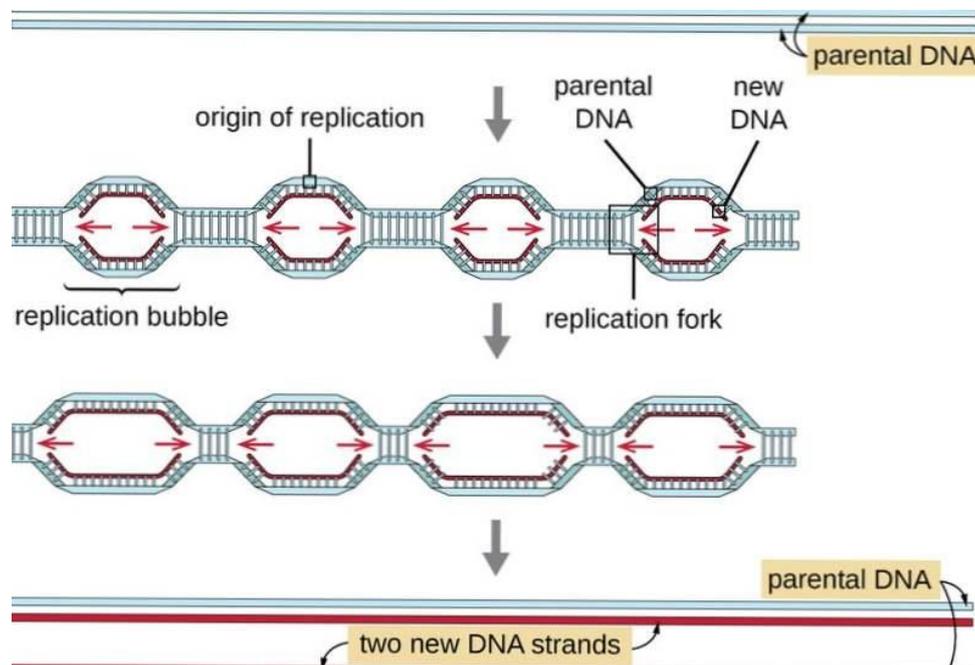


Figure : les origines de réplication

Pour commencer la réplication trois protéines sont nécessaires pour débiter l'assemblage du réplisome : **le complexe de la reconnaissance de l'origine (ORC)**, qui est un composant central de la réplication de l'ADN eucaryote et reste lié à la chromatine aux origines de la réplication tout au long du cycle cellulaire, **les deux protéines Cdc6, Cdt1**. Ce complexe sert à recruter l'hélicase en présence d'un autre complexe appelé **MCM (minichromosome maintenance)** qui formeront ensuite **le réplisome**.

La réplication est corrélée au cycle cellulaire par l'intermédiaire de la disponibilité des protéines Cdc6 et Cdt1, ces protéines sont synthétisées lors de la phase G1 du cycle cellulaire et elles sont détruites après le début de la synthèse. De cette façon, le réplisome peut être assemblé uniquement avant la phase S. une fois la réplication commencée, aucun nouveau réplisome ne peut se former au niveau des origines car les protéines Cdc6 et Cdt1 sont dégradées pendant la phase S et ne sont pas donc disponible.

Les ADN polymérases permettent la synthèse du brin fils par polymérisation en utilisant un brin de l'ADN parental comme matrice. L'activité ADN-polymérasique est la synthèse de l'ADN dans le sens 5'→3' dans la chaîne du brin en cours de synthèse. Cette activité nécessite une initiation : démarrage par une amorce de nucléotides. Pour la réplication, l'amorce est un ARN synthétisée par polymérisation par une primase (ARN polymérase). **Les primases** sont des ARN polymérases ADN-dépendante : elles synthétisent les amorces d'ARN et sont dites ADN-dépendantes car elles se servent de l'ADN comme matrice.

Chez les eucaryotes, un complexe **ADN polymérase α -primase** est formée de quatre sous unités. Les deux plus petites sous-unités du complexe (de 60 et 50 kDa) forment la primase (ou oligoribonucléotide polymérase). L'activité de synthèse d'oligoribonucléotides (2 à 12 nucléotides de long) est portée par la sous-unité de 50 kDa, elle se fait dans le sens 5'→3' du sens de la chaîne en cours de synthèse. Le rôle de la sous-unité de 60 kDa est jusqu'alors inconnu. Ces courts fragments d'ARN sont ensuite allongés par la sous-unité de 180 kDa qui synthétise un fragment d'ADN adjacent de 20 à 40 nucléotides de long. Une amorce ARN-ADN est ainsi créée ; elle va permettre la synthèse du premier fragment d'Okazaki. Le mécanisme par lequel l'ADN polymérase α -primase passe de l'activité primase à l'activité polymérase n'est pas encore établi.

2-3- Elongation :

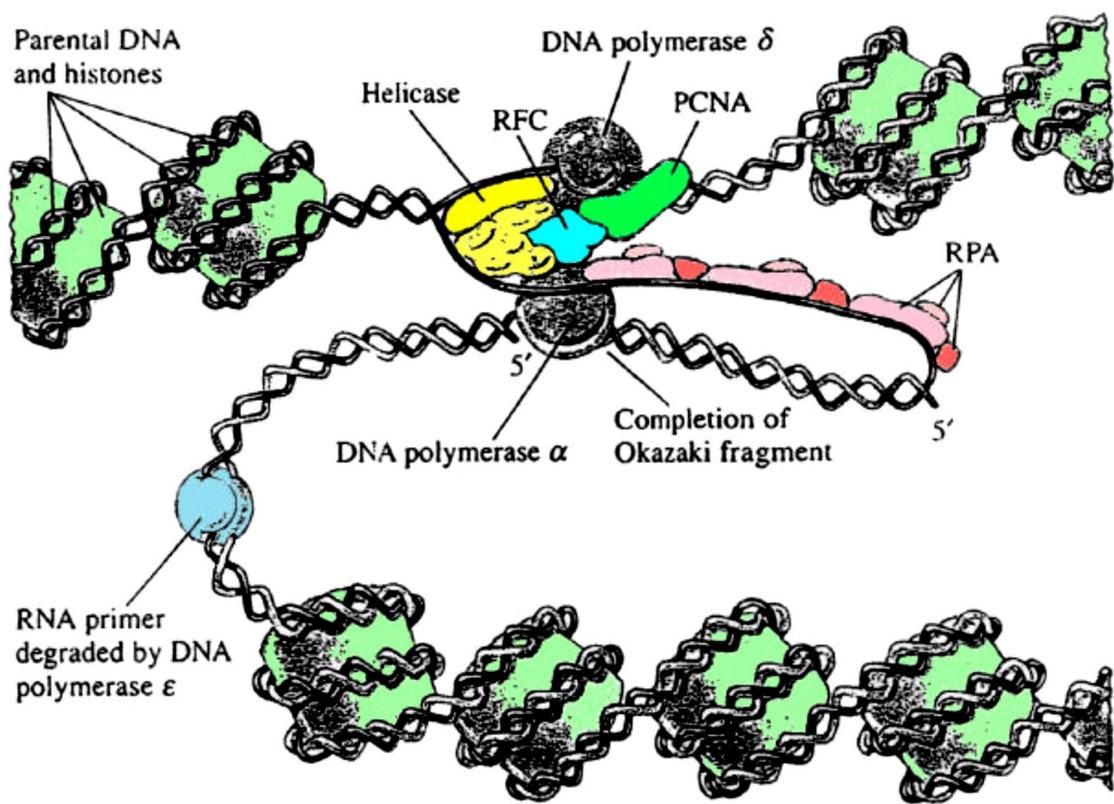
Un complexe **RF-C/PCNA** (facteur de réplication C / antigène nucléaire de prolifération cellulaire) se fixe sur l'extrémité 3'OH de cette amorce ARN/ADN néosynthétisée, dissocie l'ADN polymérase α de la matrice d'ADN, laissant la place à l'ADN polymérase δ ou ϵ qui reconnaît le complexe RF-C/PCNA et sera responsable de la synthèse du brin continu. L'ADN-polymérase δ ou ϵ se positionne sur l'extrémité 3'OH pour polymériser de l'ADN dans le sens 5'→3' après incorporation successive de nucléotides. La synthèse se fait de manière complémentaire et anti parallèle à l'ADN parental. Lors de la formation de la liaison phosphodiester, il y a libération de deux pyrophosphates qui s'hydrolysent en deux phosphates inorganiques.

Sur le brin à synthèse discontinue, l'ADN polymérase α -primase préalablement dissociée remet en route la synthèse d'une nouvelle amorce d'ARN/ADN. L'ADN polymérase α primase est ensuite remplacée par l'ADN polymérase δ qui complète la synthèse pour former **un fragment d'Okazaki** de 300 nucléotides de long. Des fragments d'Okazaki sont ainsi synthétisés de manière discontinue et joints entre eux par de l'ADN ligase I qui engendre la ligature de l'extrémité 5'-P de la séquence de l'amorce du fragment d'Okasaki en amont à l'extrémité 3'OH du fragment d'Okasaki en aval et ainsi de suite.

- **Notion de boucles** : il y a un repliement structurel de l'ADN du brin fils, de façon à ce que l'allongement se fasse dans le sens 5'-3' mais aussi dans le la fourche de réplication.

2-4- Terminaison du brin d'ADN

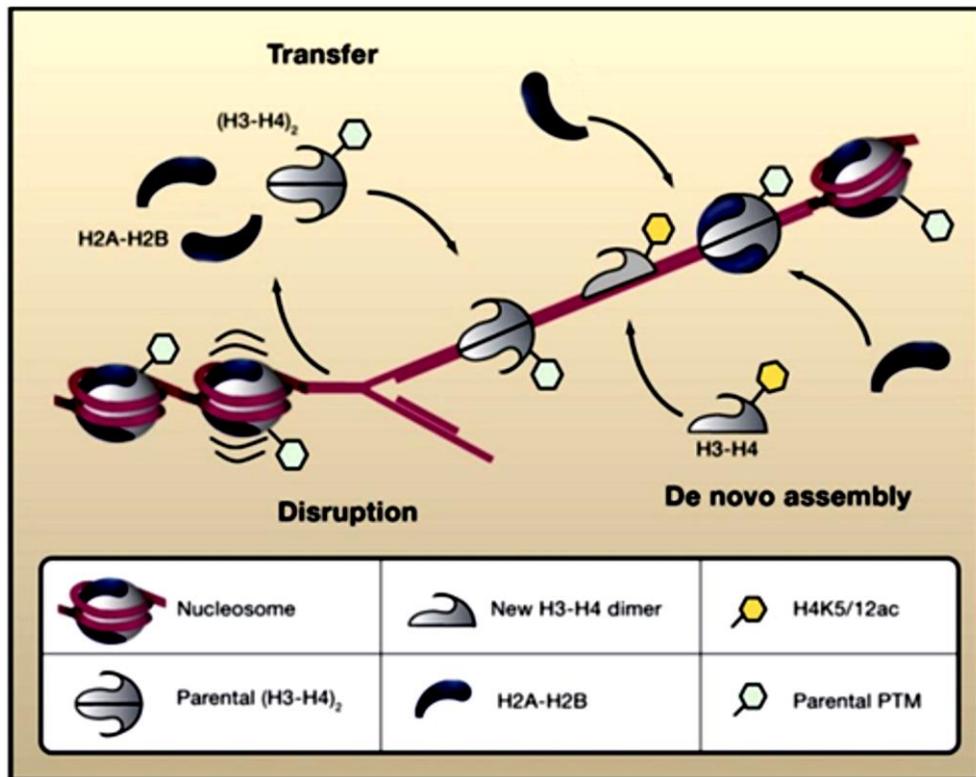
Les amorces d'ARN qui initient la réplication progressivement dégradées par une enzyme spécifique qui hydrolyse l'ARN sur les hybrides ARN-ADN : la **Rnase H** et la nucléase **FEN-1** et la lacune est comblée par l'ADN polymérase ϵ ou δ , puis l'ADN ligase I lie les morceaux entre eux pour former un seul brin. Quand la synthèse des 2 copies est achevée, la topoisomérase II décatène les 2 chromosomes.



Les nucléosomes sont des structures composées de protéines basiques : les histones. L'ADN est enroulé autour de ces histones, et enroulé sur lui-même, dans un état de condensation. Le réplisome doit donc non seulement copier les brins parentaux mais également dissocier les nucléosomes dans brins parentaux et les brins néosynthétisés dans la molécule fille.

Les nucléosomes parentaux restent associés, puis une répartition équivalente se fait entre les deux brins (brin parental et brin fils) de la molécule d'ADN. Pendant la phase S du cycle cellulaire, de nouveaux nucléosomes sont synthétisés pour compléter les nucléosomes parentaux dans les cellules filles. Cette distribution est réalisée grâce à l'acheminement de nouvelles histones en association avec une protéine appelée facteur d'assemblage de la chromatine 1 (CAF-1) jusqu'au réplisome. CAF-1 se fixe aux histones et les dirige vers la fourche de réplication, où elles peuvent être

assemblées avec de l'ADN. Son chargement d'histones arrive au niveau de la fourche de réplication en se fixant au PCNA. Chaque brin possède alors 50% de nucléosomes parentaux et 50% de nucléosomes néosynthétisés. Ce mécanisme de ségrégation des nucléosomes fonctionne avec l'assemblage des nucléosomes de novo pour reproduire une densité correcte de nucléosomes sur les deux brins filles



Chapitre 3 : La transcription et la traduction chez les eucaryotes

I- La transcription chez les eucaryotes

Introduction

La transcription est une biosynthèse d'ARN sur une matrice d'ADN et qui repose, comme celle de l'ADN sur la complémentarité des bases. Ce processus présente des analogies avec celui de la réplication mais également des différences fondamentales:

Contrairement à la réplication qui intéresse la totalité du génome à chaque cycle, le programme de transcription n'est pas fixe : seules, de petites portions du génome sont transcrites à une époque donnée de la vie de la cellule et ces portions varient en fonction du développement, de l'environnement etc...

La transcription commence donc en un point précis de l'ADN pour se terminer en un point également précis, l'espace entre les deux constitue **une unité de transcription**, notion proche du cistron mais pas tout à fait identique.

1. Caractéristiques de la transcription :

- **Un seul brin d'ADN est transcrit**, c'est à dire sert de modèle à la polymérisation des ribonucléotides. En effet, un seul brin de l'ADN en un endroit donné à un sens en termes de protéine c'est pourquoi l'on écrit généralement une séquence d'ADN sous forme d'une succession de bases de 5' à 3'. Dans cette convention on représente effectivement le brin qui possède le code (**brin codant**) mais, c'est l'autre brin (3' → 5' dans notre convention) qui est transcrit. Le résultat est une molécule d'ARN dont l'orientation 5' → 3' correspond à l'orientation NH₂ - COOH de la protéine. La lecture du code (la traduction) se fait dans le même sens que la transcription.
- **La transcription est assurée par des ARNs polymérases** qui utilise l'ADN simple brin (une dénaturation locale de la molécule d'ADN est nécessaire) mais elle polymérise des ribonucléotides en regard des désoxyribonucléotides.

Les cellules eucaryotes expriment 3 ARN polymérases (ARN pol) nucléaires différentes, responsables de la synthèse des différentes classes d'ARN:

Tableau : les principaux ARN polymérases eucaryotes

Type	Present in	Transcribes
RNA polymerase I	All eukaryotes	Large rRNAs
RNA polymerase II	All eukaryotes	Pre-mRNA, some snRNAs, snoRNAs, some miRNAs
RNA polymerase III	All eukaryotes	tRNAs, small rRNAs, some snRNAs, some miRNAs
RNA polymerase IV	Plants	Some siRNAs
RNA polymerase V	Plants	RNA molecules taking part in heterochromatin formation

Une 4^e ARN polymérase, l'ARN polymérase IV, a été découverte chez les plantes. Elle se trouve dans le noyau et elle transcrit les ADN mitochondriaux et des ARNs qui jouent un rôle dans la méthylation de l'ADN et dans la structure de la chromatine.

Les ARN polymérases sont constitués d'au moins 10 sous-unités, certaines d'entre elles étant communes aux 3 enzymes. Les plus grandes sous-unités de chaque polymérase sont homologues entre elles et aux sous-unités α , β et β' de l'ARN polymérases de *E.coli*.

* Remarque : Les ARN polymérases A, B et C avaient été initialement baptisées I, II et III en raison de leur ordre d'élution de certains supports chromatographiques.

- **De nombreux facteurs protéiques** interviennent également pour assurer l'initiation, l'élongation et la terminaison mais ce ne sont pas les mêmes que ceux intervenant dans la réplication.

2. Les différentes phases de la transcription :

2-1- Le complexe d'initiation :

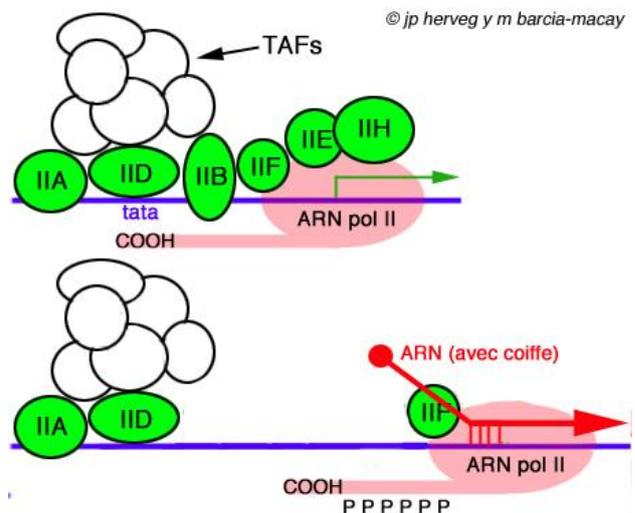
L'ARN polymérase II des eucaryotes ne se fixe pas directement sur le promoteur. Elle se fixe par l'intermédiaire de facteurs généraux de la transcription (Les GTFs) comprenant plusieurs protéines dénommées TF pour « transcription factor » et II pour l'ARN polymérase II: TFIIA, TFIIB.... Ainsi se forme le complexe de préinitiation ou machinerie basale de la transcription. Ces protéines associées à l'ARN polymérase II se lient aux promoteurs en amont des sites d'initiation de la transcription.

La première étape de l'initiation de la transcription est la reconnaissance de la boîte TATA par le facteur TFIID, qui est composé du facteur TBP (TATA binding protein) et d'une série de facteurs associés appelés TAFs (TBP-associated factors).

Les TAFs sont importants pour 2 raisons : ils sont capables de reconnaître des promoteurs dépourvus de boîte TATA et ils permettent à TFIID de répondre à des facteurs de transcription situés parfois très loin.

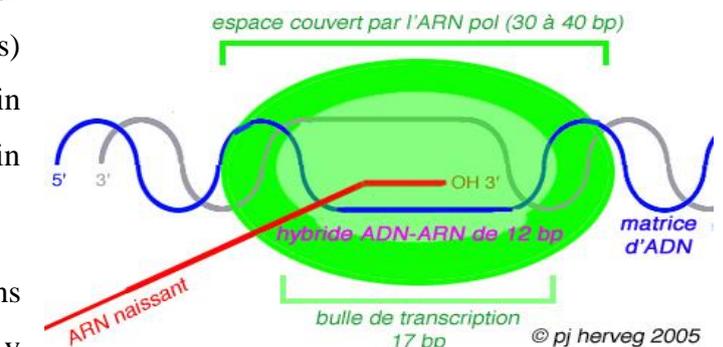
D'autres facteurs multimériques se lient ensuite de manière séquentielle au promoteur. TFIIH amène l'ARN polymérase au complexe. TFIIH est essentiel dans le complexe de preinitiation puisqu'il possède une activité hélicase qui contrôle la dénaturation du promoteur.

Un premier complexe, TFII-D reconnaît le promoteur et permet la fixation de TFII-A puis des interactions protéiques entre cet assemblage, TFII-B et la polymérase II permettent la fixation de celle-ci. Le recrutement de TFII-E suit celui de l'ARN pol II, il semble impliqué dans l'ouverture de l'ADN double brin. D'autres facteurs (TFII-H et J) participent à la modification de la topologie de l'ADN.



2-2- L'élongation de la transcription

Lorsque le site promoteur est libéré par la progression de l'ARN polymérase II sur l'ADN constituant la phase d'élongation. L'ARN polymérase II lit le brin patron (ou brin antisens) qui est complémentaire du brin codant (ou brin sens) depuis son extrémité 3' vers son extrémité 5'. Le brin d'ARN néosynthétisé est donc identique au brin codant d'ADN.



La bulle de transcription : C'est l'endroit dans laquelle les deux brins d'ADN ont été séparés. On y trouve l'hybride ADN-ARN qui mesure 12 bp.

La molécule d'ARN n'est transcrite par l'ARN polymérase qu'à partir d'un seul des deux brins de la molécule d'ADN. Contrairement aux ADN polymérases, les ARN polymérases n'ont pas besoin d'une amorce pour initier cette transcription et elles sont incapables d'identifier et d'éliminer un nucléotide mal apparié dans une chaîne en synthèse.

2-3- La terminaison :

La terminaison est assurée par des signaux spécifiques de terminaison portés par le brin qui annoncent la fin de la transcription sur le brin d'ADN matrice (TTATTT) dont le signal de polyadénylation AAUAAA. La polymérase continue sa transcription un peu après ce motif puis est libérée sous l'action de divers facteurs. Elle arrête bientôt son travail de transcription et libère l'ARNpm qu'elle vient d'assembler.

Les séquences situées en aval du site de polyadénylation sont riches en GU ou en U. Cette séquence AAUAAA sert de signal pour qu'une endonucléase coupe la chaîne d'ARN naissante à un site spécifique 10 à 15 bases en aval.

3. La maturation du transcrit primaire :

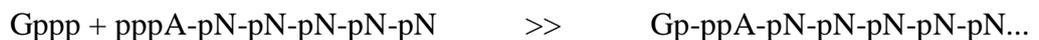
La synthèse d'ARN chez les eucaryotes donne généralement naissance à un produit de transcription "primaire" (ou pré-messager) qui devra subir une maturation pour fournir le messager cytoplasmique fonctionnel. Il existe trois grands types de modifications, catalysées chacune par des enzymes de nature protéique ou ribonucléique.

3-1- L'addition d'une coiffe en 5' :

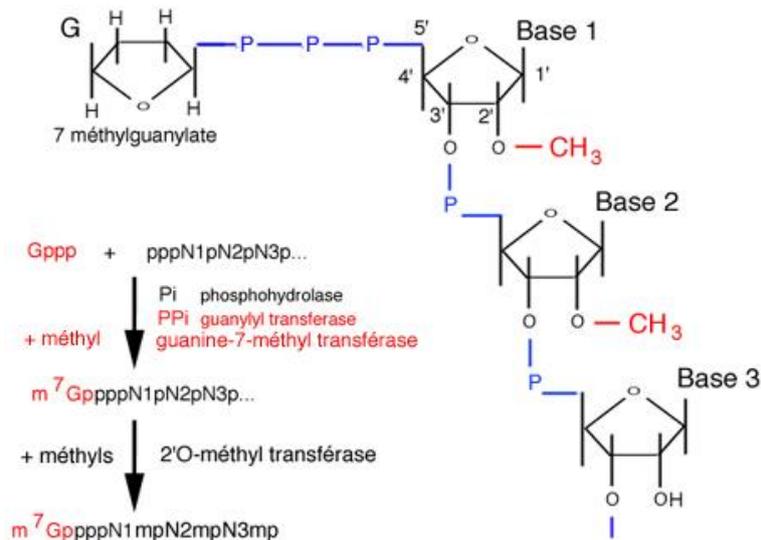
La coiffe est formée par addition d'une guanosine triphosphate : le premier nucléotide du messager est généralement une purine, A ou G et représente théoriquement l'extrémité 5' triphosphorylée de la molécule :



En fait, une guanine est ajoutée par une liaison inhabituelle 5'-5' :



Diverses méthylations peuvent se produire ensuite notamment une en position 7 de la guanine et sur le ribose pour compléter cette structure que l'on retrouve dans tous les messagers eucaryotiques.



La coiffe joue sans doute un rôle dans la protection de l'ARNm vis-à-vis de la dégradation par les nucléases et sert également de site de reconnaissance pour la machinerie de synthèse protéique.

3-2- Poly-adénilylation en 3' par la poly-A polymérase (PAP)

La polyadénylation est un ajout post-transcriptionnel de nucléotides adényliques au niveau d'un site de polyadénylation du transcrit primaire. Le site de clivage déterminé par le dinucléotide CA est entouré par une séquence AAUAAA très conservée située 10 à 30 nucléotides en amont du site de clivage, et par une séquence **DSE (Down Stream Element)** riche en U ou en GU situé une trentaine de nucléotides en aval du site de clivage. Le site est reconnu par un complexe protéique appelé CPSF (Cleavage and Polyadenylation Specific Factor), et la séquence DSE par un complexe CstF (Cleavage Stimulation Factor) en plus d'une poly-A polymérase. Tous ces composants vont interagir en formant le complexe de clivage. Cette structure va former l'extrémité 3' du message, elle peut aller jusqu'à 200 nucléotides.

La présence de poly(A) aurait également une fonction de protection des ARNm sur l'extrémité 3'. Elle stabilise de nombreux ARNm en augmentant leur durée de vie, et donc leur disponibilité pour la traduction. Elle protège l'ARNm de la dégradation par des enzymes présentes dans le cytoplasme. Elle a aussi un rôle facilitateur de l'attachement des ribosomes à l'ARNm.

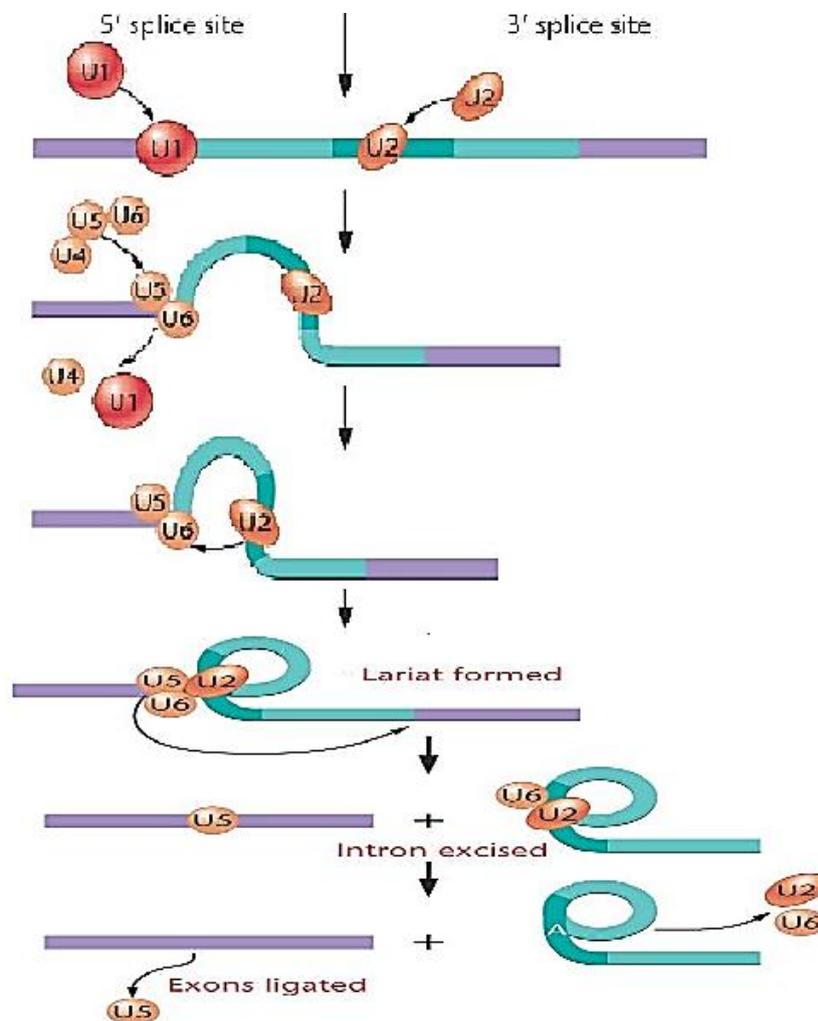
3-3- L'excision-épissage

Après l'addition de la coiffe et la polyadénylation, le transcrit primaire est encore soumis à l'excision des introns et l'épissage des exons ; les introns sont ainsi éliminés

L'analyse de nombreux introns de levure a fait apparaitre trois séquences consensus que l'on a pu généraliser : la première, GU est appelée consensus gauche car elle représente l'extrémité 5' de

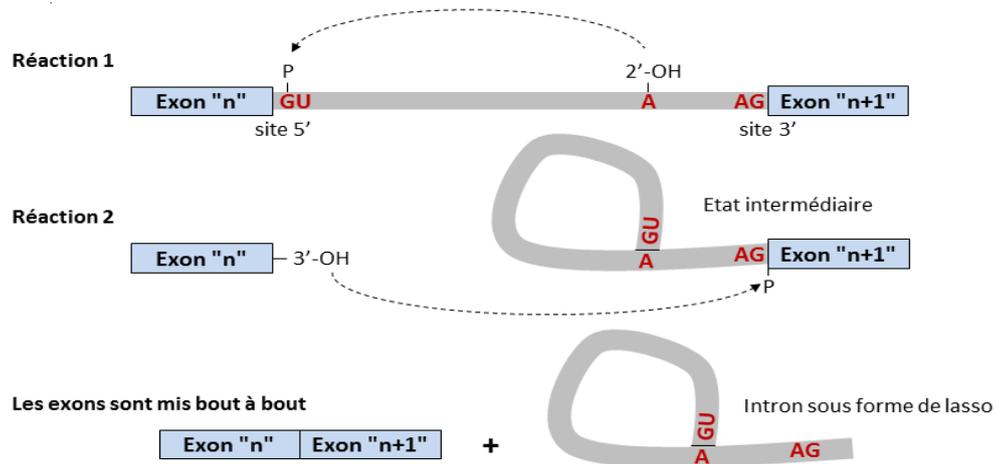
l'intron, de même que la droite (AG) représente la jonction 3' intron - 5' exon suivant, une autre, UACUAAC chez les messagers de levure (plus généralement Py N Py Py Pu A Py chez les eucaryotes supérieurs, Py représente une pyrimidine, N un nucléotide quelconque, Pu une purine, l'adénine étant remarquablement constante) est située peu avant l'extrémité 3' de l'intron et appelée séquence de branchement. Ces séquences sont reconnues par des ribonucléoprotéines qui vont former un complexe nécessaire à l'épissage, appelé **spliceosome**

Ce mécanisme de splicing implique cinq petits ARN nucléaires ou small nuclear RNA (U1, U2, U4, U5 et U6) et une douzaine de protéines. U1 joue un rôle essentiel dans la reconnaissance du site donneur de splicing. U2, U5 et U6 sont les seuls snRNAs directement impliqués dans les réactions de splicing proprement dites. De ces 3 snRNAs, U6 est celui qui est le plus impliqué dans la catalyse de la réaction de splicing. U5 permet de retenir l'exon amont une fois qu'il a été détaché de l'intron suite au clivage du site donneur de splicing. Il permet aussi de conserver les exons alignés de manière à les lier l'un à l'autre au cours de la deuxième étape du splicing.



Par la liaison des différents Snurps, va replier l'intron et permettre une liaison curieuse de l'extrémité 5' de l'intron à l'hydroxyle 2' de l'adénine de la séquence consensus dite "de

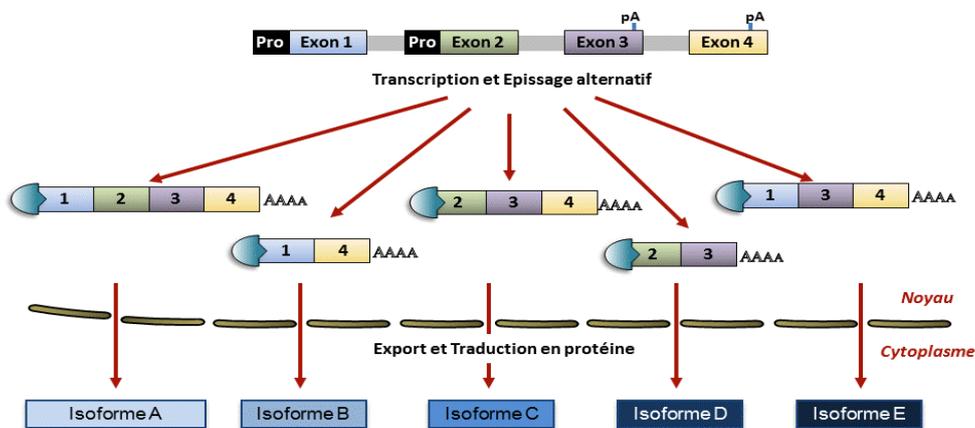
branchement". L'extrémité 3' va être détachée de la 5' de l'exon suivant et une liaison phosphodiester peut se créer entre les exons. Le branchement en 2' de l'adénosine fait prendre une structure dite en "lasso" à l'intron, in vivo, ce lasso est immédiatement dégradé.



4. L'épissage alternatif

Un gène commence par un promoteur permettant l'initiation de sa transcription et se termine par une séquence terminatrice. L'initiation et la terminaison peuvent avoir lieu à différents endroits au sein du gène et ainsi mener à la production de transcrits plus longs ou plus courts. Le pré-ARNm peut être épissé d'une manière alternative, il s'agit du processus **d'épissage alternatif**.

En effet, **il n'est pas obligatoire que tous les exons d'un gène soient inclus dans l'ARN messager mature**, certains exons sont donc considérés comme « **alternatifs** ». Ainsi un seul gène peut produire différents ARNm matures et par conséquent, plusieurs isoformes protéiques ayant des fonctions biologiques différentes, voire opposées.

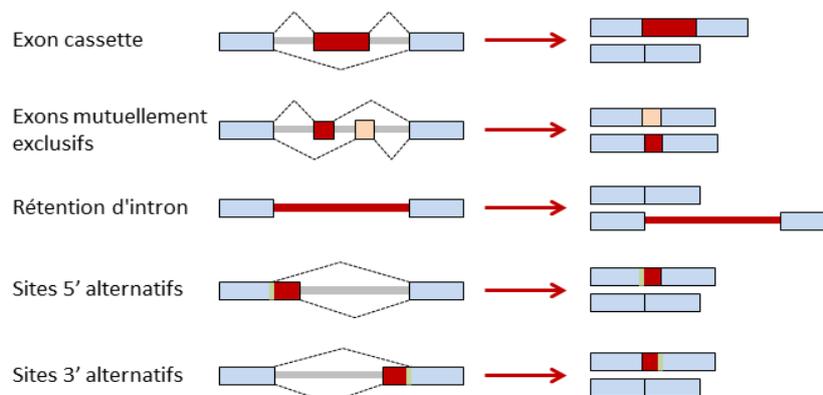


L'épissage alternatif augmente significativement la diversité des transcrits grâce à l'utilisation alternative d'exons ou d'introns, mais aussi de signaux de polyadénylation situés dans des exons alternatifs. La majorité de ces événements se produit dans un maintien du cadre de lecture (la

succession de trinuécléotides qui permet de décoder la séquence d'acides aminés de la protéine correspondante), résultant en l'expression de différentes isoformes protéiques.

Les exemples d'épissage alternatif peuvent être classés en cinq modèles différents :

1. L'exon cassette (ou le saut d'exon) : il s'agit d'un évènement où un exon est entièrement inclus ou retenu dans le transcrit mature. C'est le cas d'épissage alternatif le plus simple, qui prédomine chez les mammifères, donnant lieu à des transcrits plus ou moins longs.
2. Les exons mutuellement exclusifs : l'un des deux exons, qui sont généralement de taille similaire, est retenu dans l'ARNm mature, mais jamais les deux ensemble. Cet évènement change peu la taille du transcrit mature mais modifie les propriétés de la protéine produite.
3. La rétention d'intron : un segment de pré-ARNm peut être épissé comme un intron ou simplement retenu. Si l'intron ne modifie pas le cadre de lecture, il code pour des acides aminés comme les exons qui l'envoient, produisant ainsi une protéine d'un poids moléculaire supérieur. Dans le cas contraire, il entraîne l'apparition d'un codon stop prématuré, ce qui provoque la dégradation du transcrit mature. La rétention d'intron est l'évènement d'épissage le plus rare chez les mammifères.
4. Site d'épissage alternatif en 5' (site donneur d'épissage) : ce variant d'épissage modifie la borne de l'exon dans sa partie aval.
5. Site d'épissage alternatif en 3' (site accepteur d'épissage) : d'une manière similaire au variant précédent, cet évènement modifie la borne de l'exon dans sa partie amont, incluant ainsi un segment de l'extrémité 3' de l'intron épissé.

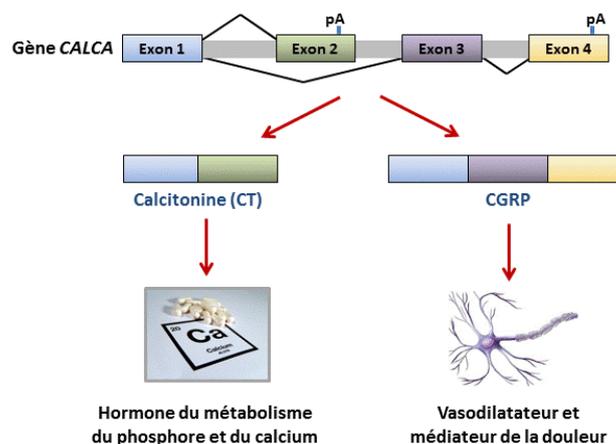


Même si ces formes décrivent les cas les plus répandus de l'épissage, elles ne reflètent pas la complexité du mécanisme. Les exons ne sont pas choisis par hasard. L'épissage est régulé par des protéines de liaison à l'ARN (répresseurs et activateurs), qui se lient sur des séquences cis-

régulatrices ("*silencer*" et "*enhancer*"). La structure secondaire de l'ARN peut elle aussi influencer sur l'épissage.

Exemple :

Les transcrits primaires du gène CALCA produisent, par épissage alternatif dans la glande thyroïdienne, la Calcitonine (CT), une hormone impliquée dans le métabolisme du phosphore et du calcium. En revanche, dans le système nerveux central, les transcrits CALCA produisent un peptide nommé CGRP, un vasodilatateur et médiateur de la douleur



II- Le code génétique

1- Définition

Le code génétique est un code qui permet la conversion d'une séquence de nucléotides (ADN puis ARN) en séquence d'acides aminés (protéines). Le code implique les bases A, C, T et G ainsi que les 20 acides aminés. L'unité de base du code génétique fut appelée **un codon**. Chaque position de nucléotide dans l'ARNm ne peut être occupée que par une des quatre bases A, G, C ou U. Avec trois nucléotides par codon, il y a $4^3 = 64$ codons différents possibles ce qui est plus que suffisant pour coder 20 acides aminés différents.

		2ème position					
1ère position (5')	U	C	A	G	3ème position (3')		
U	UUU } phe	UCU } ser	UAU } tyr	UGU } cys	U		
U	UUC } phe	UCC } ser	UAC } tyr	UGC } cys	C		
U	UUA } leu	UCA } ser	UAA } stop	UGA } stop	A		
U	UUG } leu	UCG } ser	UAG } stop	UGG } trp	G		
C	CUU } leu	CCU } pro	CAU } his	CGU } arg	U		
C	CUC } leu	CCC } pro	CAC } his	CGC } arg	C		
C	CUA } leu	CCA } pro	CAA } gln	CGA } arg	A		
C	CUG } leu	CCG } pro	CAG } gln	CGG } arg	G		
A	AUU } ile	ACU } thr	AAU } asn	AGU } ser	U		
A	AUC } ile	ACC } thr	AAC } asn	AGC } ser	C		
A	AUA } met	ACA } thr	AAA } lys	AGA } arg	A		
A	AUG } met	ACG } thr	AAG } lys	AGG } arg	G		
G	GUU } val	GCU } ala	GAU } asp	GGU } gly	U		
G	GUC } val	GCC } ala	GAC } asp	GGC } gly	C		
G	GUA } val	GCA } ala	GAA } glu	GGA } gly	A		
G	GUG } val	GCG } ala	GAG } glu	GGG } gly	G		

2- Caractéristiques du code génétique

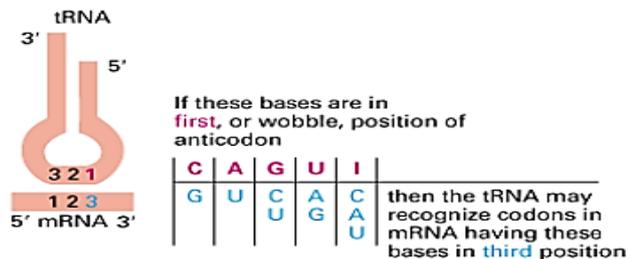
Le code génétique possède différentes caractéristiques :

- **Les codons sont des triplets de nucléotides** et ils codent pour un acide aminé.
- **La séquence du gène et la séquence de la protéine codée sont colinéaires**, c'est-à-dire que la longueur du gène et la longueur de la structure primaire de la protéine finale sont proportionnelles.
- **Le code génétique est universel**. En effet chaque acide aminé dispose d'un ou plusieurs codons et ceci au niveau d'une multitude d'organismes vivants.
- **Le code génétique est redondant (ou dégénéré)**. Plusieurs codons codent pour un même acide-aminé

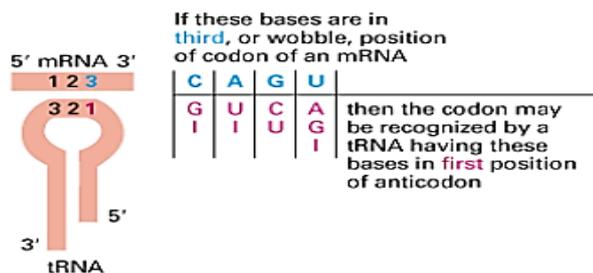
On trouve 64 codons. Parmi les 64 codons, trois sont des **codons stop**, qui signalent la fin de la traduction. Les 61 codons restants, **appelés codons sens**, codent les 20 acides aminés trouvés dans les protéines. Le code contient donc plus d'informations qu'il n'en faut pour spécifier ces 20 acides aminés, et on dit qu'il s'agit d'un **code dégénéré**. Seuls le Trp et la Met ne sont spécifiés que par un seul codon. D'autres acides aminés sont spécifiés par deux codons, et certains, comme l'Arg et la Leu sont spécifiés par six codons différents. Des codons qui spécifient le même acide aminé sont dits **synonymes**.

De nombreux codons synonymes ne diffèrent que par la base en 3^o position. Exemple : Ala est codée par 4 codons qui commencent par GC. Souvent se sont les deux premiers nucléotides du codon qui définissent l'acide aminé, la redondance est donc due au troisième nucléotide du codon.

Il existe une flexibilité dans l'appariement des bases en position 3 du codon et en position 1 de l'anticodon, cette flexibilité s'appelle **le wobble**. Quand le codon dans l'ARNm s'apparie avec l'anticodon de l'ARNt, la 1^{ère} base (5') du codon s'apparie avec la 3^o base (3') de l'anticodon, strictement : A avec U, C avec G. Ensuite, les bases du milieu du codon et de l'anticodon s'apparient aussi strictement selon les règles. Après que ces paires de bases aient formé des liaisons hydrogène, les 3^o bases ne s'alignent pas parfaitement, ce qui provoque le flottement (les contraintes structurales ne permettent que des ajustements conformationnels limités. Ainsi, C ou A en 3^o position (5') de l'anticodon exigent un appariement strict. Par contre, U ou G permettent l'appariement avec deux codons différents).



Codon position 3	Anticodon position 1
U	A, G, I
C	G, I
A	U, I
G	C, U

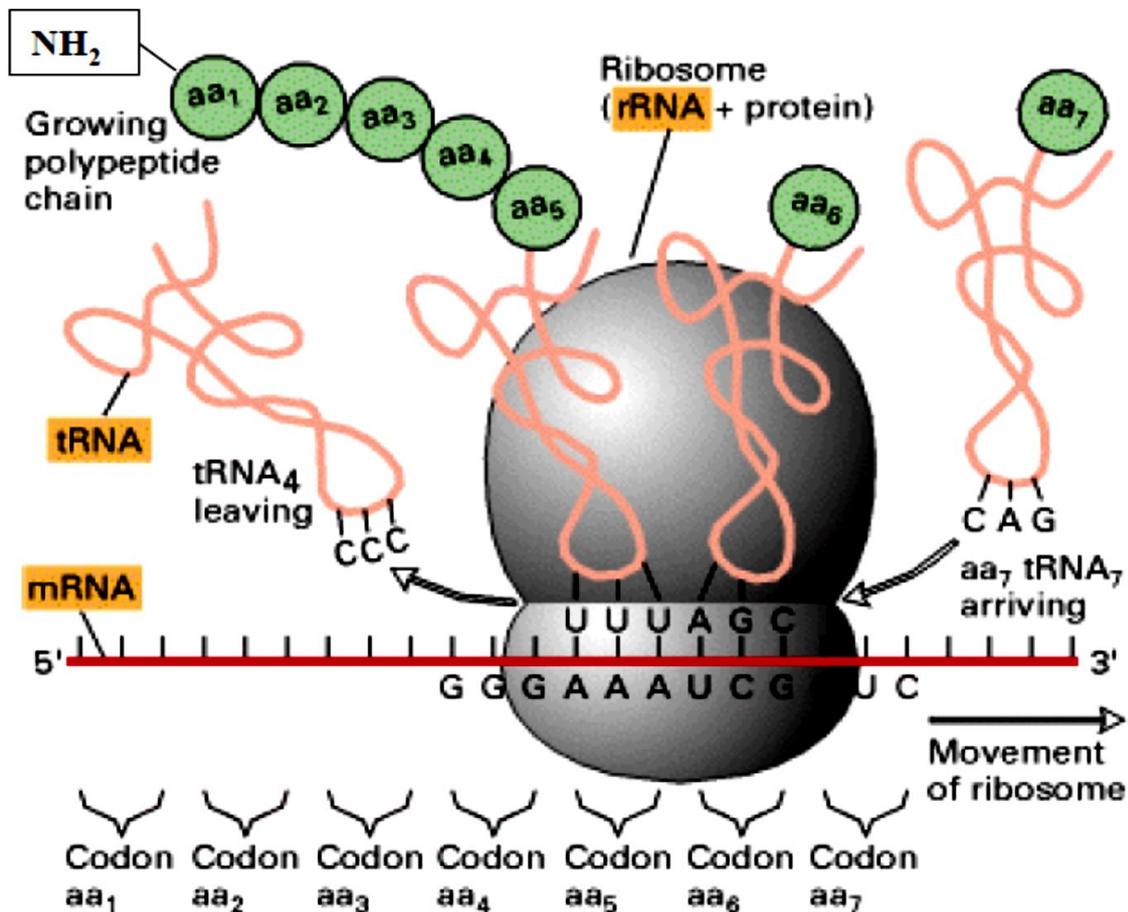


Remarque : I est une base purique qui peut s'apparier avec U, C et A mais pas G.

- **Le code génétique est non-chevauchant.** Les nucléotides d'un codon ne participent qu'au code d'un seul acide aminé, ainsi le prochain acide-aminé sera codé par le prochain codon présent sur l'ARNm. On parle du **cadre de lecture** (ou *reading frame*).
- Un cadre de lecture qui consiste exclusivement en triplets qui correspondent à des acides aminés est appelé cadre ouvert de lecture : **ORF = open reading frame**
- **Le code possède un système de ponctuation.** Le codon d'initiation est le codon AUG (GUG pour ma mitochondrie) et les codons de terminaison sont les codons UAA (ocre),

UAG (ambre) et UGA (opale). Le codon UGA (opale) n'est pas présent au niveau de la mitochondrie.

III- La traduction chez les eucaryotes



La traduction est le passage de séquence de nucléotides à des séquences d'Aa en respectant le code génétique. La traduction s'effectue dans le cytoplasme de la cellule. Une fois l'ARNm a atteint le cytoplasme, il va servir pour synthétiser un polypeptide, avec le concours de l'ARNt et du ribosome. Elle a aussi besoin d'énergie, des ribosomes, des ARNt, de l'activité enzymatique et d'un plan apporté par l'ARNm.

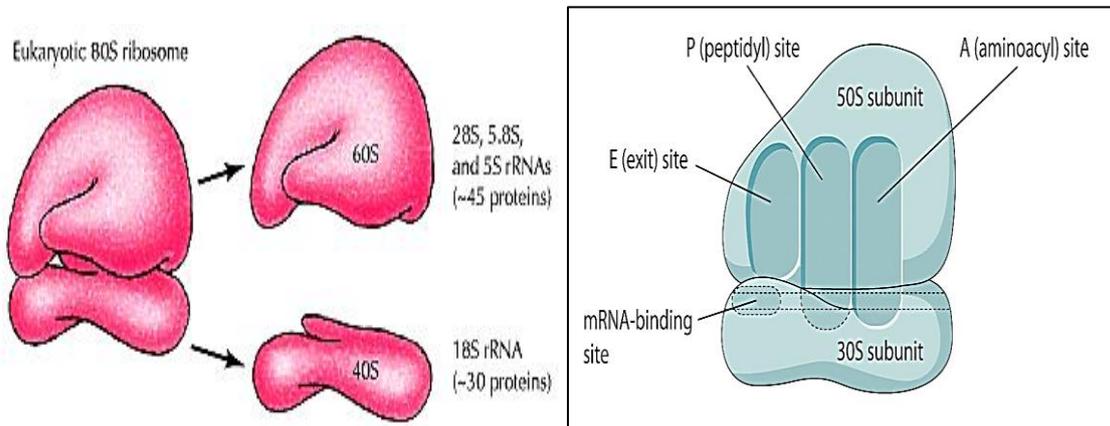
1- Les acteurs de la traduction

Les acteurs de la traduction sont l'ARN messager (ARNm), les ARN de transfert (ARNt), les ribosomes, les acides aminés, les amino-acyl tRNA synthétases, le Mg²⁺, le GTP et l'ATP.

1-1- Les ribosomes

Les ribosomes sont constitués d'ARN ribosomiques (ARNr) et de protéines et sont structurés sous forme de deux sous-unités. Leur taille est définie en unité Svedberg.

Les ribosomes eucaryotes (80S) sont constitués d'une petite sous-unité 40S et d'une grande sous-unité 60S. La sous-unité 40S est constituée d'un ARNr 18S et de 33 protéines. La sous-unité 60S est constituée des ARNr 28S, 5,8S et 5S ainsi que de 49 protéines.



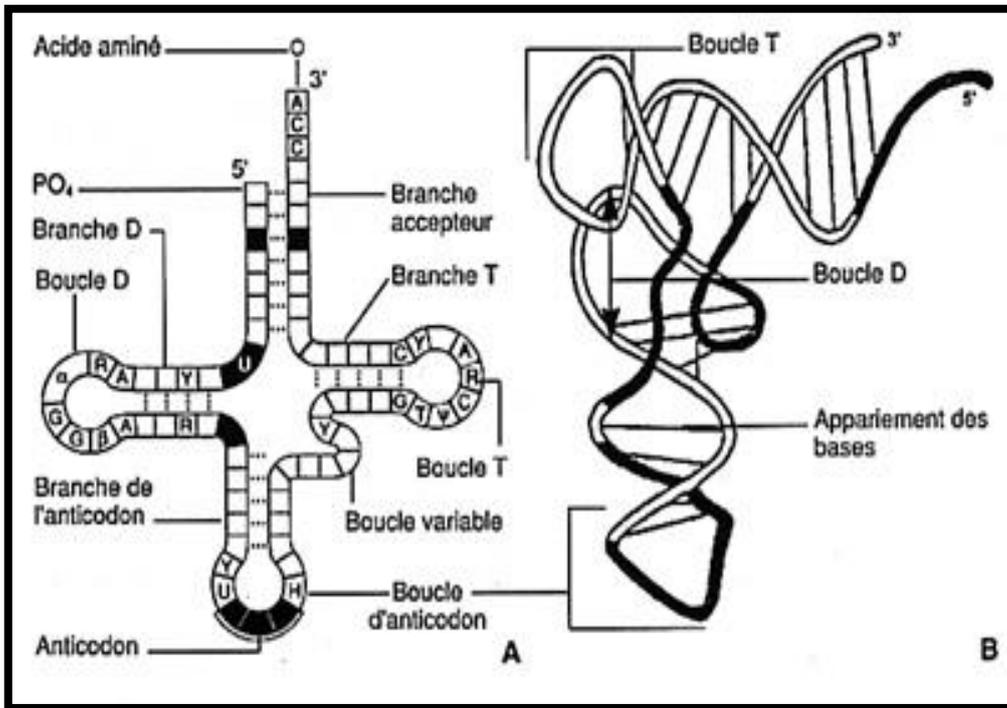
Les ribosomes ont 3 sites pour lier les ARNt :

- ✓ Le site A : (= site Acide-aminé ou Accepteur) où viendra l'ARNt porteur de l'AA
- ✓ Le site P (= site Peptidique ou Donneur) pour l'ARNt porteur de la chaîne peptidique en cours d'élongation.
- ✓ Le Site E : (= site Exit) sortie de l'ARN de transfert.

1-2- Les ARNt :

L'ARNt qui porte la Met initiale a une conformation qui lui permet d'être logé d'emblée exceptionnellement dans le site peptidique.

Certains nucléotides d'un ARNt sont complémentaires et forment des liaisons hydrogène intramoléculaires, créant une structure en feuille de trèfle. Tous les ARNt ont la même séquence terminale en 3' (CCA), où s'attache l'acide aminé.



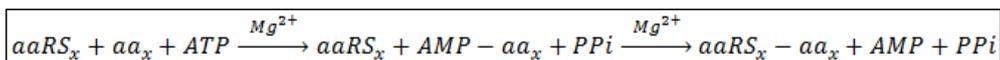
Les ARNt ont une structure secondaire en forme de trèfle à 3 feuilles et une structure tertiaire en forme de L à l'envers. Lors du mécanisme de traduction il y a un appariement antiparallèle entre l'ARNm et l'ARNt pendant la synthèse de la protéine: reconnaissance codon-anticodon au niveau de la boucle de l'anticodon.

Les ARNt possèdent également un bras de l'acide aminé qui le fixe en 3' (CCA) sur le ribose, il s'agit d'une liaison covalente : liaison ester riche en énergie. Les acides aminés ne vont ainsi par arriver libre sur le ribosome mais associés à leurs ARNt respectifs. On trouve 40 à 60 ARNt différents par cellule, il existe donc plusieurs ARNt différents pour un acide aminé, on les appelle **ARNt iso-accepteurs**.

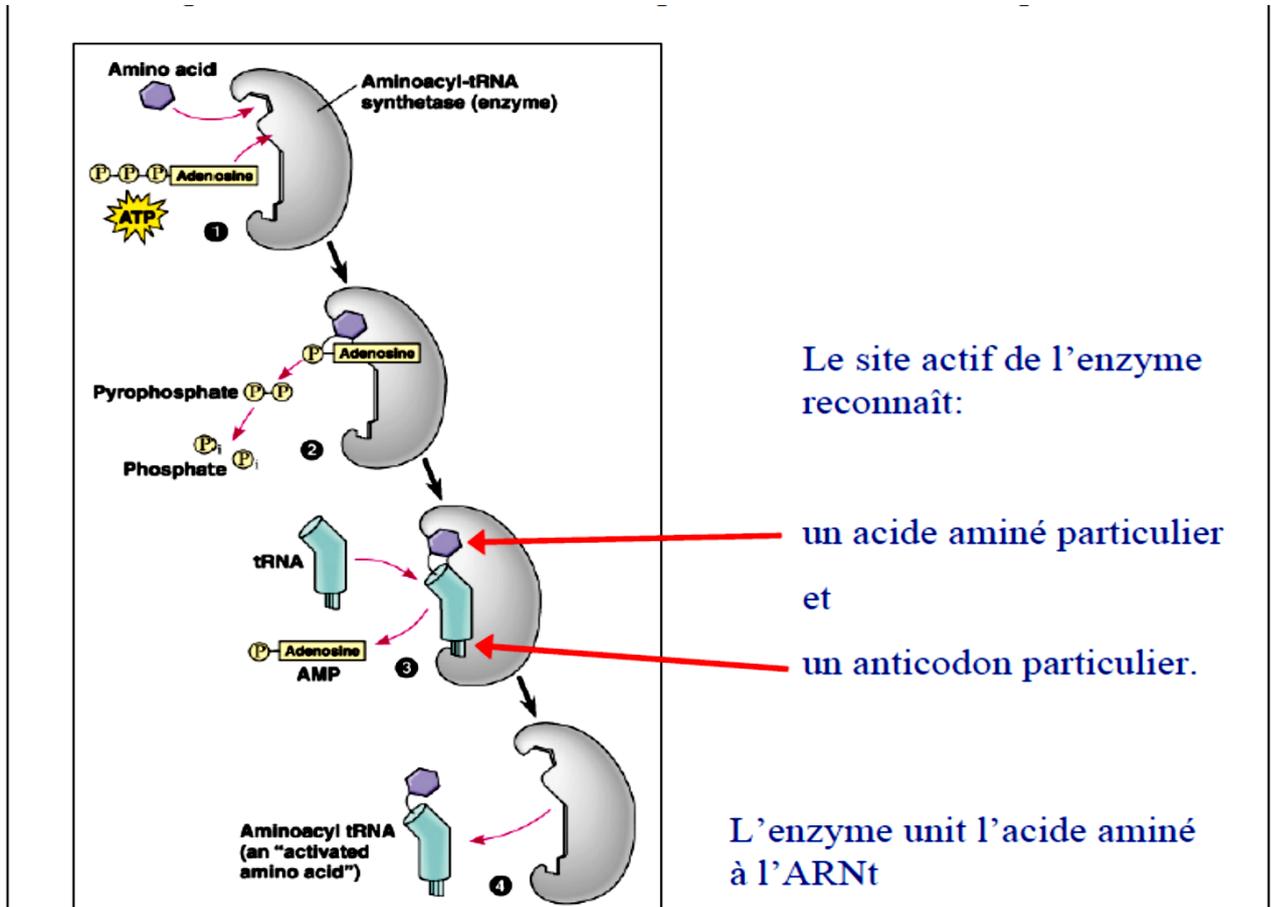
2- Chargement de l'acide aminé sur l'ARNt :

La formation du complexe amino-acyl-tARN (aa-tARN) nécessite une Amino-acyl-tRNA-synthétase spécifique de l'acide aminé, qui doit ainsi reconnaître toutes les formes de codon de cet acide aminé. Le chargement correct de l'ARNt est un élément important dans la fidélité de la traduction.

L'acide aminé (aa) est tout d'abord activé et cette activation nécessite de l'énergie sous forme d'ATP pour permettre la formation d'aa-AMP (liaison anhydride mixte).



La liaison formée entre l'ARNt et l'acide aminé est une liaison covalente de type carboxy-ester. Les Amino-acyl-tRNA-synthétase sont au nombre de 20 dans la cellule, autant qu'il y a d'acides aminés qui rentrent en compte dans la traduction. L'acide aminé complexé peut ainsi s'associer à la chaîne.



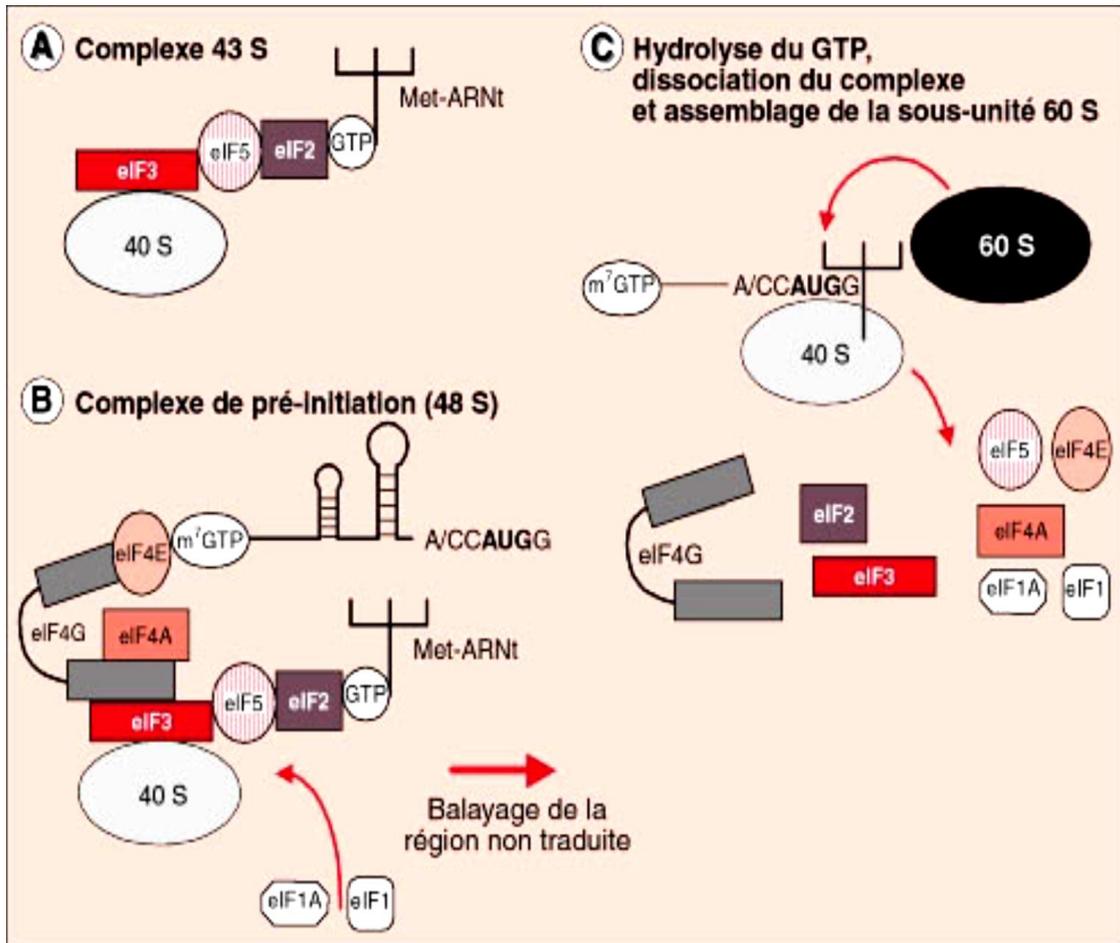
3- Les différentes étapes de la traduction

La traduction chez les eucaryotes s'effectue en trois étapes : initiation, élongation et terminaison. Chaque étape requiert des facteurs de traduction (facteurs d'initiation eIF, facteurs d'élongation eEF, facteurs de terminaison eRF) qui sont associés de manière transitoire aux ribosomes.

Etape	Initiation	Elongation	Terminaison
Facteurs	Facteurs d'initiation IF/eIF	Facteurs d'élongation EF/eEF	Facteurs de terminaison RF/eRF (R = release)
Signaux	Codon d'initiation 5' AUG _{3'}	Codons sens 5' NNN _{3'}	Codons de terminaison ("stop", non-sens) 5' UAA _{3'} , 5' UAG _{3'} , 5' UGA _{3'}

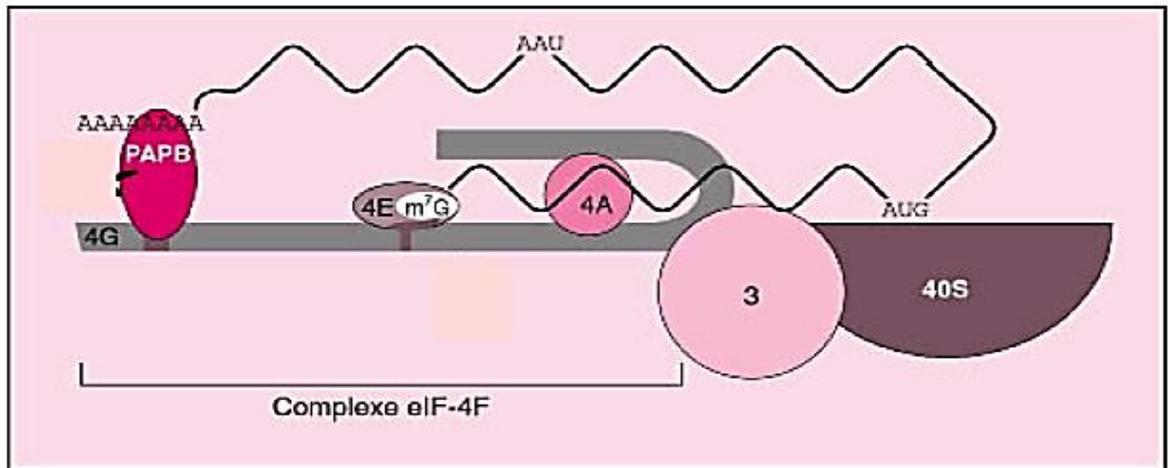
3-1- Initiation

Chez les eucaryotes, des facteurs d'initiation positionnent l'ARNt initiateur portant une méthionine (Met-ARNti) sur la petite sous-unité ribosomale.



- **Le complexe 43S** : Le complexe ternaire, formé par le Met-ARNti, eIF2 et une molécule de GTP, est recruté au niveau de la sous-unité 40S. Cela est sous le contrôle du facteur eIF3 qui est fixé à la sous-unité 40S, et par eIF5 qui peut se lier à la fois à eIF3 et à eIF2.
- **Formation du complexe de pré-initiation 48S** : Le facteur multiprotéique eIF4F, qui se compose de 3 protéines distinctes, eIF4E, eIF4A et eIF4G, pilote l'attachement du complexe 43S sur l'ARNm. eIF4E lie spécifiquement la coiffe méthylée en 5' de l'ARNm et eIF4G assure la cohésion de l'ensemble car elle est liée à eIF4E par son extrémité amino-terminale et à eIF3 par son extrémité carboxy-terminale.

La poly A binding protein (PABP). Reconnaît la queue poly (A) de l'ARNm et possède également un site de liaison avec la région N-terminale de eIF-4G. L'interaction des complexes polyA/PABP et coiffe/eIF-4E, par l'intermédiaire de eIF-4G, permet une circularisation de l'ARNm et favorise le déclenchement de la traduction

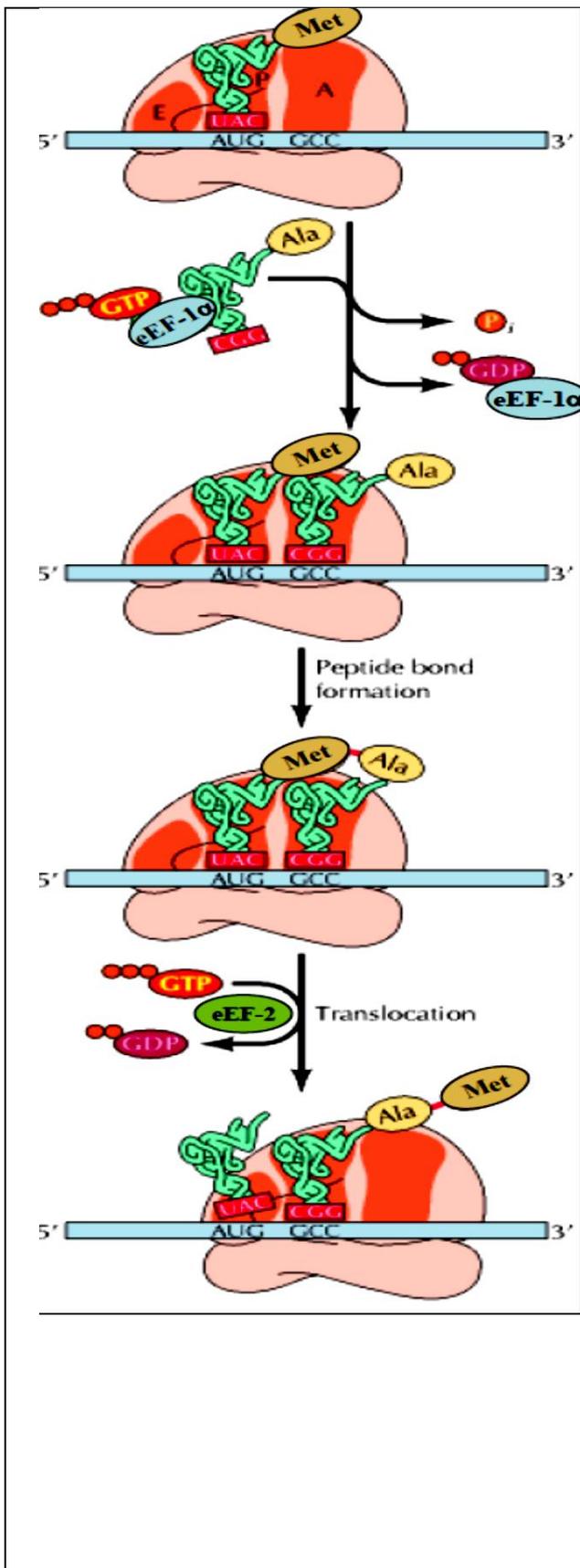


Le ribosome est alors prêt à balayer la région non traduite. Lors de cette progression, l'activité hélicase du facteur eIF4A déroule les structures secondaires de l'ARN.

- **Formation du ribosome 80S** : Lorsque le complexe de pré-initiation reconnaît le codon AUG d'initiation, eIF5 catalyse l'hydrolyse du GTP, ce qui provoque la dissociation du complexe de pré-initiation. La sous-unité 60S peut alors s'assembler pour former le ribosome 80S

3-2- L'élongation

L'élongation correspond à une synthèse protéique par ajout d'acides aminés à l'extrémité C-Terminale de la chaîne peptidique naissante, réaction catalysée par l'activité peptidyl-transférase de la grande SU des ribosomes.



Avant la formation de la liaison peptidique, le site P (P= Peptidyl) contient le peptidyl-ARNt(codon n).

L' aminoacyl-ARNt entre dans le site A (A= Aminoacyl) (codon n+1)

- l'hybridation codon/aniconon est basée sur la complémentarité des bases. Ainsi, elle consomme de l'énergie libérée par l'hydrolyse du GTP fixé sur eEF-1 α

Formation d'une nouvelle liaison peptidique. La sous-unité 60S catalyse l'activité peptidyl-transférase entre le COOH de la Met initiale ou du dernier acide aminé fixé sur le peptide en cours de synthèse porté par l'ARNt sur le site P et le NH2 de l'acide aminé porté par l'ARNt sur le site A.

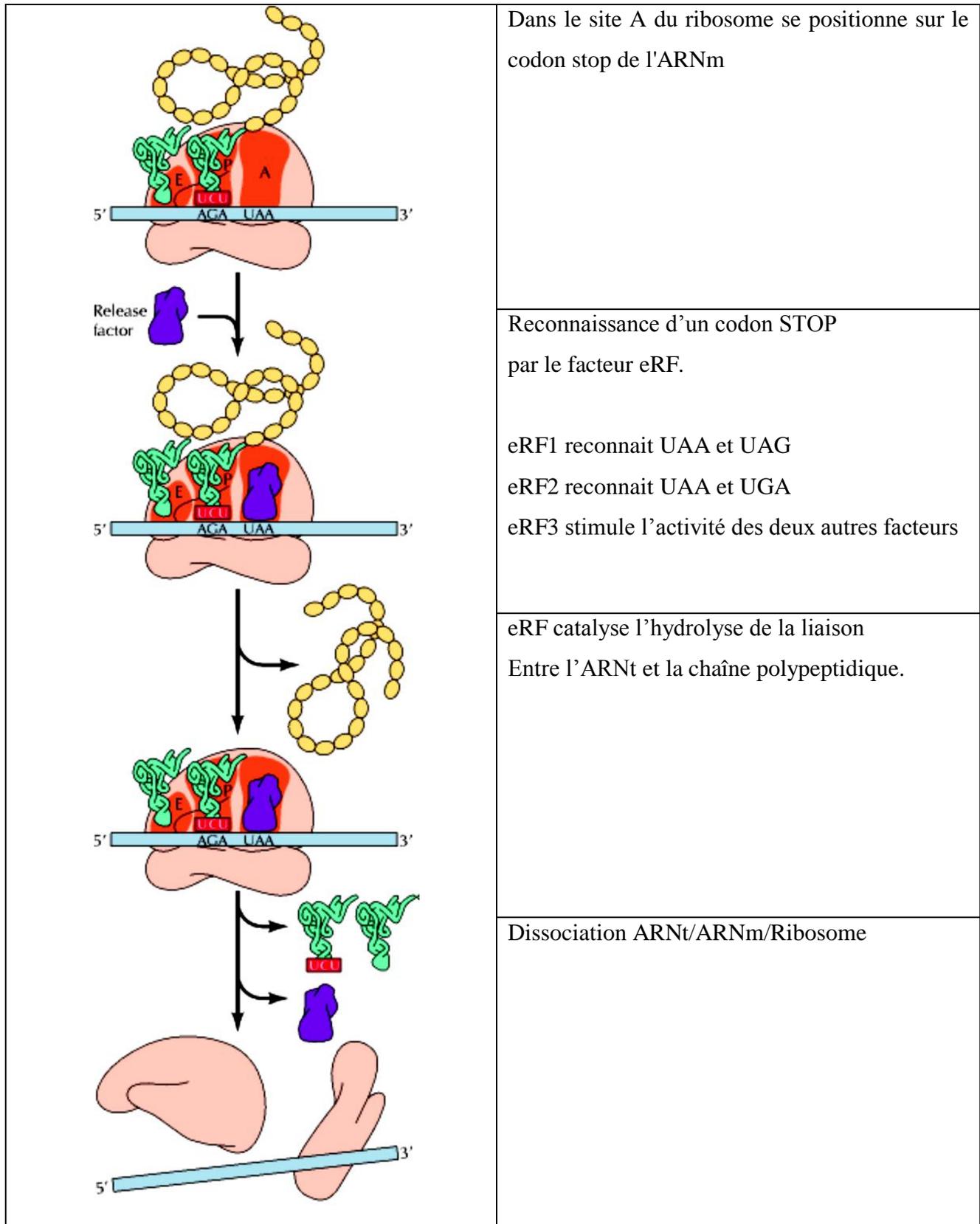
Cela entraîne le transfert du polypeptidyl-ARNt du site P vers l' aminoacyl-ARNt au site A

La translocation déplace le ribosome d'un codon; place le peptidyl-ARNt dans le site P. L'ARNt déacylé quitte le ribosome en passant par le site E (E= Exit). le site A est libéré et peut recevoir un autre acide aminé-ARNt. la translocation consomme l'énergie fournie par l'hydrolyse eEF-2-GTP => eEF-2-GDP,

3-3- La terminaison

Finalement, le site A du ribosome se positionne sur le codon stop de l'ARNm (UAA, UAG et UGA) qui ne codent pour aucun acide aminé. Ces codons sont reconnus par les facteurs de terminaison

eRF1, eRF2 et eRF3 (Eukariotic release factor), ces facteurs eRF activés par du GTP s'associent au ribosome et clivent la liaison entre le peptide et l'ARNt, l'hydrolyse du GTP permet : la libération de eRF; la dissociation des sous-unités du ribosome; la libération de la protéine, de l'ARNt et de l'ARNm.



Chapitre 4 : La cytogénétique et mécaniques chromosomiques

Introduction

Si le terme de chromosome a été proposé dès 1888 par Waldeyer pour désigner les éléments colorés visibles au cours de la division cellulaire, leur dénombrement a longtemps été difficile du fait de l'enchevêtrement des chromosomes visibles à la métaphase, et il aura fallu attendre 1956 date à laquelle Tjio et Levan confirme définitivement que le nombre de chromosomes était de 46 dans les cellules somatiques humaines.

2- Définitions :

2-5- La cytogénétique :

La cytogénétique est l'étude des phénomènes génétiques au niveau de la cellule, plus précisément au niveau des chromosomes sans la nécessité d'extraire l'ADN : C'est la discipline qui permet d'étudier la structure, la fonction et les anomalies structurales et numériques des chromosomes. Ces remaniements chromosomiques peuvent être constitutionnels ou acquis. Les remaniements constitutionnels sont présents dès la conception ou se forment lors des premières divisions du zygote. Les remaniements acquis, sont des remaniements qui vont apparaître au sein d'une cellule au cours de la vie. Dans la majorité des cas ces remaniements acquis sont trouvés dans les cellules tumorales.

On distingue :

- **Cytogénétique Conventiionnelle ou classique** : Elle étudie le caryotype seul (rare car Souvent accompagnée ou remplacée par des analyses moléculaires)
- **Cytogénétique Moléculaire** : Etudie les anomalies quantitatives ou structurales de l'ADN constituant les chromosomes : FISH
- **Cytogénomique** : Etudie quantitativement le génome entier après extraction de l'ADN: CGH « Array » (Hybridation Génomique Comparative Sur micro réseaux) ou puces ADN

2-6- Le caryotype :

Le caryotype est une technique qui permet l'étude d'équipement chromosomique caractéristique d'une espèce donnée, elle correspond à une analyse numérique et structurale de l'ensemble des chromosomes d'une cellule. Cette technique permet d'obtenir une image, en microscopie optique, des chromosomes d'une cellule au cours de la métaphase ou de la prométaphase de la mitose.

La cytogénétique a pour objet l'étude cytologique des chromosomes. Leur analyse est effectuée au cours d'un stade du cycle cellulaire où les chromosomes sont bien individualisés et peuvent être

observés au microscope. C'est le plus souvent la métaphase mitotique qui est observée, mais l'étude de la méiose est également possible et, plus récemment, grâce à des techniques d'hybridation *in situ* utilisant des sondes d'ADN non radioactives, s'est développée celle des noyaux en interphase.

3- Techniques conventionnelles du caryotype constitutionnel

3-1- Le prélèvement :

La cytogénétique constitutionnelle consiste à étudier les chromosomes d'un individu (foetus, enfant ou adulte) à partir d'un prélèvement effectué soit en post-natal (sang veineux, biopsie cutanée), soit en prénatal (liquide amniotique, sang fœtal, villosités choriales), selon l'indication du caryotype.

Germinale ou Constitutionnelle	Somatique ou Acquise
<p><u>Pré natale</u> : caryotype fœtal à partir de :</p> <ul style="list-style-type: none"> • Amniocentèse (cellules du liquide amniotique) • Choriocentèse (cellules trophoblastiques, villosités choriales) • Cordocentèse (cellules issues du sang du cordon ; difficilement réalisable) 	<p><u>Onco hématologique</u> pour l'étude des leucémie et des lymphomes : caryotypes médullaires, ganglionnaires et des tissus hématopoïétiques.</p>
<p><u>Post natale</u> : 2 types de caryotypes possibles :</p> <ul style="list-style-type: none"> • Lymphocytaire (le plus fréquent car le prélèvement sanguin est facilement réalisable) • Tissulaire (si le lymphocytaire n'est pas réalisable ; on utilise alors des fibroblastes cutanés dans le cadre d'un mort né par exemple (car il prolifèrent encore) ou dans le cadre du syndrome de Pallister-Killian (tétrasomie du bras court du chromosome 12) où les anomalies chromosomiques sont absentes dans la lignée sanguine) 	<p><u>Tumeurs solides</u> bénignes ou malignes Ce sont les études les plus fréquentes <i>Rmq</i> : les prélèvements pour l'étude des tumeurs solides peuvent venir de liquides pathologiques (ex : liquide pleural, ascite)</p>

3-2- La mise en culture :

Afin d'obtenir des cellules en métaphase, les prélèvements sont mis en culture (le temps de culture est variable selon le type de cellule analysé) dans un milieu de culture stérile additionné d'antibiotiques et qui peut contenir des facteurs de croissance ou des substances chimiques favorisant la prolifération des cellules.

Un prélèvement de sang veineux périphérique recueilli stérilement sur tube héparinate de lithium est le plus souvent utilisé pour réaliser un caryotype constitutionnel en période postnatale. Le sang total est incubé 48 à 72 heures dans un milieu de culture contenant une lectine à fort pouvoir mitogène (phytohémagglutinine ou PHA) permettant une stimulation de la croissance des lymphocytes T. Les fibroblastes de la peau sont également couramment utilisés en période postnatale mais demandent une culture cellulaire de une à trois semaines.

En période anténatale, les amniocytes du liquide amniotique, les cellules trophoblastiques des villosités choriales ou les lymphocytes du sang fœtal sont utilisés. L'établissement d'un caryotype à partir de liquide amniotique nécessite, en moyenne, un temps de culture de 6 à 10 jours. Le temps de culture peut être plus long, parfois de plusieurs semaines, en raison de la faible quantité de cellules initiales en cas d'amniocentèse très précoce ou en raison des nombreuses cellules en apoptose et des débris cellulaires en cas d'amniocentèse tardive.

3-3- Obtention des cellules en métaphase

Après la phase de multiplication des cellules, celles-ci sont bloquées au stade de métaphase afin de pouvoir observer les chromosomes. Pour se faire, on utilise un poison du fuseau de division (classiquement c'est la Colchicine qui est utilisée ou son équivalent synthétique la Colcémide). Cette substance empêche la polymérisation de la tubuline et donc la formation du fuseau mitotique. La mitose est bloquée au stade de métaphase.

3-4- Le choc hypotonique

Cette étape, indispensable pour avoir un étalement correct, entraîne le gonflement des cellules par différence de pression osmotique. Les membranes cytoplasmique et nucléaire sont fragilisées.

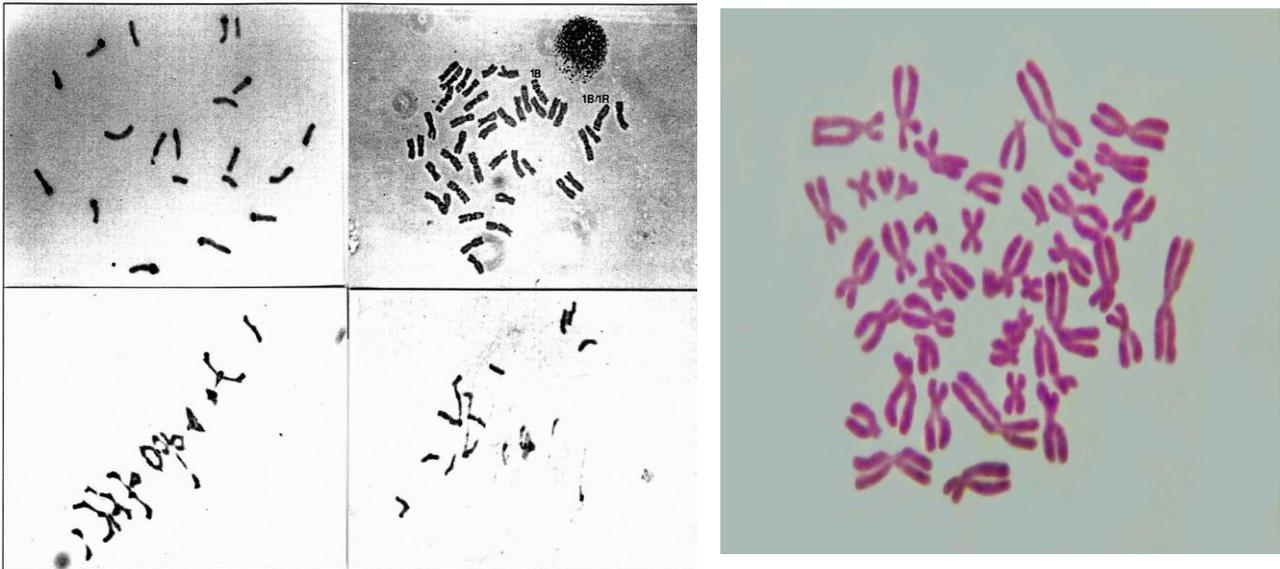
3-5- La Fixations/ l'étalement

Les constituants cellulaires sont ensuite fixés grâce à un fixateur, par exemple un mélange acide acétique et du méthanol. Cette acidification du milieu permet l'arrêt du choc cellulaire. Les cellules sont alors prêtes à être étalées. On laisse tomber quelques gouttes de cette préparation sur une lame de verre. Le fait de faire tomber la suspension cellulaire fait éclater les membranes fragilisées, libérant ainsi les chromosomes qui restent toutefois groupés. Les lames sont ensuite observées en microscopie optique.

3-6- Dénaturation /Coloration

Lorsque l'on colore des préparations chromosomiques avec du **Giemsa**, les chromosomes prennent un aspect rose violacé à peu près homogène sur toute leur longueur. On ne peut donc les distinguer les uns des autres que par leur taille et leur forme. Cependant, ces critères sont insuffisants pour assurer la reconnaissance et l'interprétation correcte des anomalies chromosomiques. Pour reconnaître spécifiquement chaque paire chromosomique, on utilise donc des techniques de marquage particulières qui permettent d'obtenir une coloration inhomogène des chromosomes par le Giemsa et l'apparition de bandes. C'est la succession de bandes sombres et claires le long d'un

chromosome, identique chez tous les individus pour un chromosome donné, qui en permet l'identification précise, selon le même principe qu'un code à barres.



Il existe deux principales techniques de marquage en bandes des chromosomes (banding), utilisées en routine :

- **Les bandes G** : obtenues par digestion trypsinique modérée des chromosomes suivie d'une coloration au Giemsa. Elles sont les plus utilisées en raison de leur facilité d'obtention.
- **Les bandes R** : obtenues par dénaturation thermique ménagée.
 Dans les deux cas, les bandes ne deviennent visibles qu'après une coloration avec le Giemsa. Ces deux techniques donnent un marquage réciproque, c'est à dire que là où l'on obtient une bande sombre avec l'une des deux techniques, on observe une bande claire avec l'autre. D'autres techniques de marquage complémentaires existent qui permettent d'analyser certaines régions particulières du génome :
- **Bandes C** : cette coloration par le Sulfate de Baryum permet de mettre en évidence l'hétérochromatine constitutive, qui correspond à des régions non codantes du génome comme les régions centromériques.

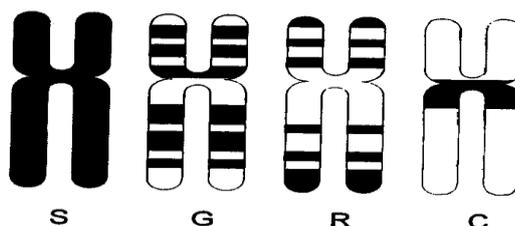


Figure : Comparaison des différentes colorations : S : avec Geimsa, G : bandes G, R : bandes R, C : bandes C

- **Bandes Q** : obtenues après coloration par la moutarde de Quinacrine. Sous UV les chromosomes présentent des bandes fluorescentes, de même que les bandes G.
- **NOR** : cette technique consiste en un dépôt de nitrate d'argent qui met en évidence les organisateurs nucléolaires. Ces structures correspondent aux régions du génome contenant les gènes qui codent pour les ribosomes.
- **Bandes T** : une dénaturation thermique poussée ne laisse persister le marquage qu'au niveau des télomères.

Tableau : récapitulatif de différentes techniques de coloration

Technique de marquage	Traitement utilisé	Marquage obtenu
Bandes Q	Moutarde de quinacrine	Régions riches en liaison AT Régions à réplication tardive Régions pauvres en gènes actifs
Bandes G	Trypsine Solution saline à 60°C	Régions riches en liaison AT Régions à réplication tardive Régions pauvres en gènes actifs
Bande R	Dénaturation à 87°C (solution de pH 5,8)	Régions riches en liaisons C-G Régions à réplication précoce Régions riches en gènes actifs
Bandes T	Dénaturation poussée à 87°C (solution de pH 5,8)	Régions télomériques
Bandes C	Solution chauffée d'hydroxyde de barium	Régions centromériques Constrictions secondaires Bras long du chromosome Y
Bandes NOR	Nitrate d'argent	Satellites de chromosomes acrocentriques (13, 14, 15, 21, 22)

4- Classement des chromosomes métaphasiques

4-1- Structure du chromosome métaphasique

Classiquement, le chromosome est défini comme une structure subcellulaire colorable (d'où son nom), portant les gènes. Un chromosome est caractérisé par un centromère, point d'attache au fuseau mitotique, relié à un bras court (p) et à un bras long (q) qui comportent une ou deux chromatides selon le stade de la mitose, et dont les extrémités constituent les télomères.

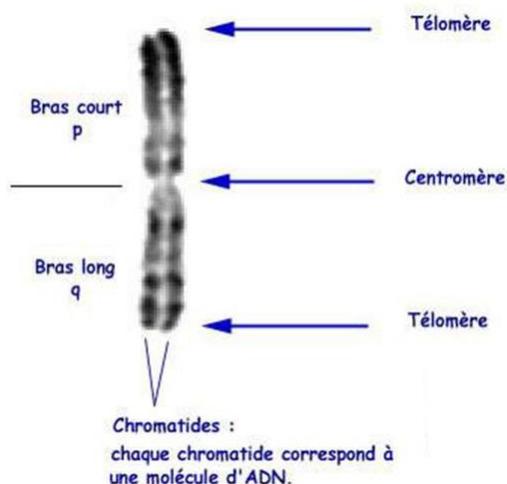
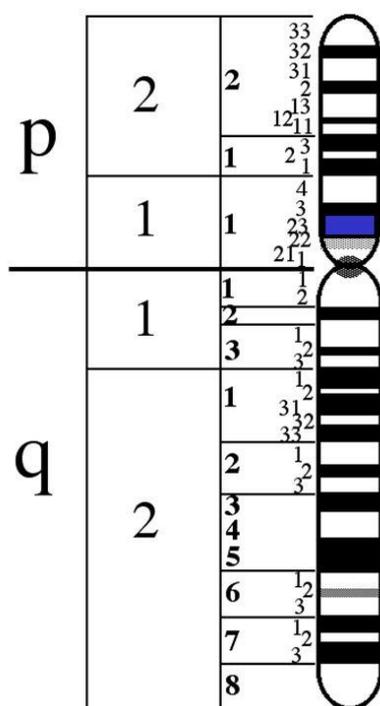


Figure : Aspect morphologique d'un chromosome métaphasique

Bandes chromosomiques :

Après coloration les chromosomes peuvent être subdivisés en régions de 1 à 3, chaque région est divisée en bandes numérotées, et certaines bandes en sous bandes.

Exemple : Xq2, 7, 3 : bras long du chromosome X, région 2, bande 7, sous bande 3.



Caryotype : un chromosome

- Bras (p ou q)
- Région (eg : p1)
- Bande (eg : p11)
- Sous-bande (eg : p11.2)
- Sous-sous-bande (eg : p11.23)

89

4-2- Critères de classement des chromosomes

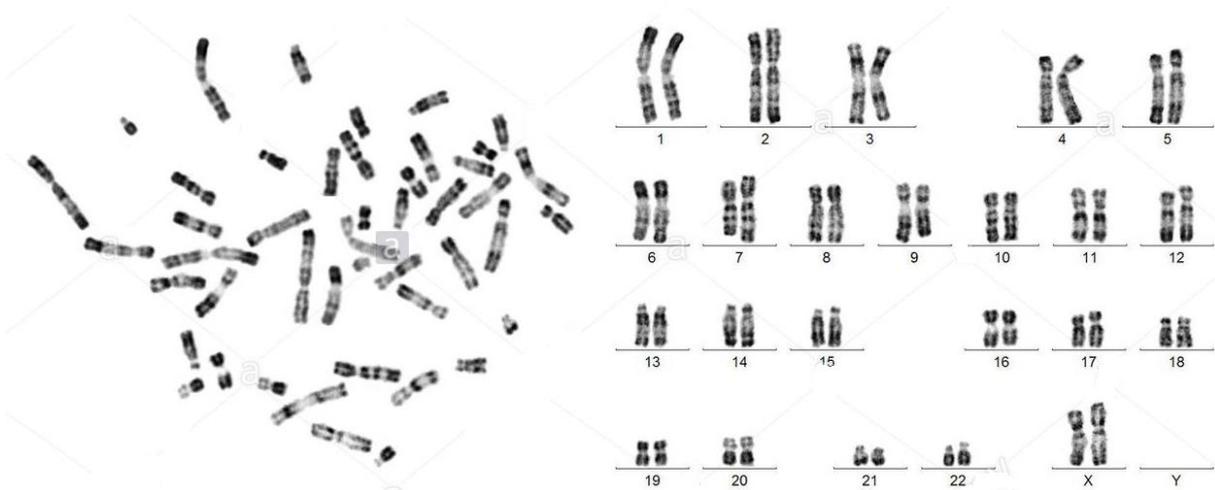
Plusieurs critères vont permettre de reconnaître et de classer les chromosomes :

- **La taille** : Par convention, les chromosomes sont classés du plus grand au plus petit.

- **L'index centromérique (IC)** : c'est-à-dire le rapport entre la taille du bras court et la taille totale du chromosome ($p/p+q$)

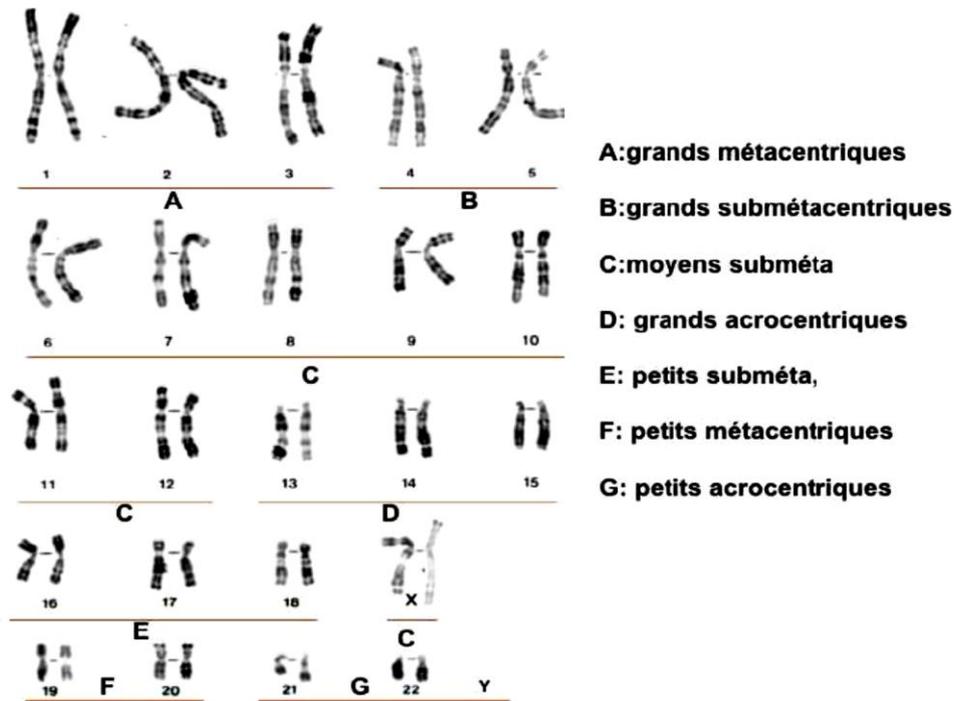
Cet index permet de reconnaître trois familles de chromosomes :

- Les chromosomes métacentriques : dont les deux bras ont une taille à peu près équivalente ($IC = 1/2$).
- Les chromosomes submetacentriques : qui ont un bras franchement plus petit que le bras long ($0 < IC < 1/2$).
- Les chromosomes acrocentriques : dont le bras court est quasi inexistant ($IC \approx 0$).



En fonction de la taille et de la position du centromère, les chromosomes sont classés en 7 groupes :

- Le groupe A : Les grands métacentriques et submétacentriques 1, 2, 3.
- Le groupe B : Les grands submétacentrique 4, 5.
- Le groupe C : Les métacentriques et submétacentriques moyens 6, 7, 8, 9, 10, 11, 12 et X.
- Le groupe D : Les grands acrocentriques 13, 14 et 15.
- Le groupe E : Les petits submétacentriques 16, 17 et 18.
- Le groupe F : Les petits métacentriques 19, 20.
- Le groupe G : Les petits acrocentriques 21, 22 et Y.



5- La formule chromosomique

Il est indiqué successivement, le nombre total de chromosomes; suivi d'une virgule, les chromosomes sexuels; une autre virgule et l'anomalie chromosomique quand elle existe. Exemple:

- Caryotype masculin normal 46,XY → 46 chromosomes/cellule dont un chromosome X et un chromosome Y.
- Trisomie 21 :47,XY,+21 → 47 chromosomes /cellule dont un chromosome X et un chromosome Y, plus un chromosome 21 surnuméraire.
- Translocation:46,XX,t(1;18) → 46 chromosomes /cellule dont deux chromosomes X et une translocation entre le chromosome 1 et le chromosome 18.

Remarque

- Les anomalies sont dites **homogènes** lorsque toutes les cellules nucléées de l'individu ont la même formule chromosomique. - Les anomalies sont dites en **mosaïques** lorsque chez le même individu coexistent au moins deux populations cellulaires de formules chromosomiques différentes. Du point de vue nomenclature ; les différentes populations cellulaires sont indiquées l'une après l'autre et séparées par une barre diagonale. Exemple:46,XX/47,XX,+21.

6- Intérêt du caryotype

L'analyse du caryotype a pour objet

- de repérer certaines anomalies chromosomiques de structure et de nombre.
- d'établir la cartographie physique du génome (localisation des gènes).
- permet de confirmer un diagnostic cliniquement évident.
- déterminer le sexe
- de déterminer aussi l'espèce
- surtout essentiel pour prodiguer un conseil génétique.

Chapitre 5 : La génétique extra-chromosomique

Introduction

La majeure partie du génome eucaryote est contenue dans les chromosomes du noyau. Cependant, en plus de l'ADN nucléaire, certains organites cellulaires : Mitochondries et Chloroplastes contiennent également un type de chromosome qui leur est «spécifique».

Les chromosomes mitochondriaux et chloroplastiques sont des molécules d'ADN double-brin. Les mitochondries et les chloroplastes contiennent de multiples copies de leurs chromosomes, et, dans chaque cellule, ces organites sont présents en nombre variable.

On appelle le chromosome mitochondrial : ADNmt

Et le chromosome chloroplastique : ADNcp

Dès le début du 20^{ème} siècle les chercheurs ont pensé que les plastes et les mitochondries pouvaient provenir de bactéries. Celles-ci auraient été ingérées par des cellules primitives et vivraient à l'intérieur d'elles en symbiose. Cette théorie endosymbiotique de l'origine des plastes et des mitochondries est devenue parfaitement plausible lorsque l'on a découvert (1950-1960) que ces organites contenaient de l'ADN et des ribosomes.

De nombreux éléments indiquent que les mitochondries et les chloroplastes ont évolué à partir de bactéries ayant subi une endocytose dans les cellules eucaryotes ancestrales, formant des endosymbiotes (théorie d'endosymbiose). Au cours de leur adaptation intracellulaire, ces bactéries ont perdu progressivement une grande partie de leurs gènes qui ont fusionnés avec le génome nucléaire. Néanmoins, les mitochondries et les chloroplastes des eucaryotes actuels, conservent de l'ADN codant des protéines essentielles à la fonction des organites ainsi que l'ARN ribosomal et l'ARN de transfert nécessaires à leur traduction.

I- La mitochondrie :

Le chondriome d'une cellule est l'ensemble des 300 à 800 mitochondries (jusqu'à 2500 dans une cellule du parenchyme du foie) qui constituent jusqu'à 25 % de la masse cellulaire. Les mitochondries sont des organites car elles constituent des unités fonctionnelles de la cellule: elles en sont les "usines énergétiques".

1. Structure :

Décrites pour la première fois par Kölliker et Flemming dès le XIX^{ème} siècle, les mitochondries, organelles cytoplasmiques, apparaissent sous forme de granules ovoïdes plus ou moins allongées avec deux aspects typiques : globulaires et filamenteux, d'où leur appellation avec les mots grecs *mitos* (filaments) et *chondros* (granules), reflétait l'hétérogénéité de leur morphologie. Présente dans toutes les cellules (exception faite des hématies), la morphologie de la mitochondrie est variable d'un tissu à un autre.

Leur nombre variable est proportionnel aux besoins énergétiques du type cellulaire, de quelques dizaines à plusieurs milliers, le compartiment mitochondrial peut occuper une partie conséquente de l'espace cellulaire (20 % dans les cellules hépatiques et jusqu'à 40 % dans les cellules musculaires).

Au microscope électronique, la mitochondrie apparaît constituée d'un système de membrane, deux phases aqueuses distinctes, physiquement séparées, à l'intérieur d'une mitochondrie. Ces deux phases sont délimitées par deux membranes de nature et de composition très différentes : une membrane externe et une membrane interne :

- L'espace inter-membranaire est situé entre la membrane externe de la mitochondrie et la membrane interne, qui projette plusieurs replis vers le centre de l'organite;
- La matrice est le fluide central de la mitochondrie, circonscrit par la membrane interne. Les replis de la membrane interne constituent autant de crêtes mitochondriales à la surface desquelles sont attachées de très nombreuses minuscules structures sphériques baignant dans la matrice.

La matrice contient:

- une centaine d'enzymes
- plusieurs copies identiques d'ADNmt,
- Ribosomes mitochondriaux ARNr
- ARNt et enzymes variées impliquées dans la transcription et la traduction des gènes mitochondriaux

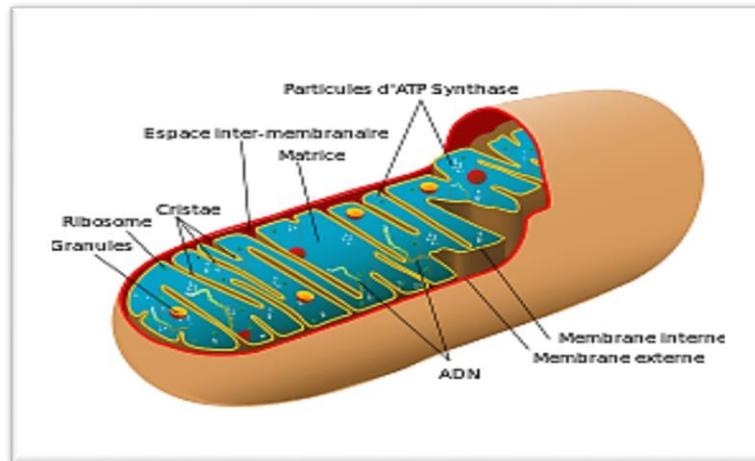


Figure : Structure de la mitochondrie

2. Le génome mitochondrial

Une mitochondrie contient plusieurs copies d'ADN circulaire. L'ADN mitochondrial représente moins de 1% de l'ADN total de la cellule.

L'ADN mitochondrial est compacté en agrégats appelés **nucléoïdes** ou mito-chromosomes. Le composant le plus abondant des nucléoïdes est le facteur de transcription A qui contribue de manière significative à la compaction de l'ADN, de manière similaire à l'action des **histones**.

Le nombre de gènes contenus dans ADNmt est faible comparé à celui ADNcp des chloroplastes dans les plantes vertes. Ce nombre varie de 5 à 94 dans l'ADNmt et il est généralement supérieur à 100 dans l'ADNcp

Le génome mitochondrial des cellules humaines (hérité de la mère) est une molécule d'ADN circulaire, double-brin, de 16569 paires de base (beaucoup plus petit que celui des plantes) et ne comporte pas d'introns, la plupart des séquences sont uniques, non-répétitives. Par contre le génome mitochondrial de la levure est constitué d'exons et d'introns. La séquence d'ADNmt est entièrement connue, avec 435 îlots CpG et 4747 cytosines dans des sites non-CpG. Il est constitué de 37 gènes.

L'ADN mitochondrial contient :

- 2 gènes d'ARN ribosomiques (12S et 16S)
- 22 gènes d'ARN de transfert

La plupart des protéines de la mitochondrie sont codées par le génome nucléaire et importées dans la mitochondrie. Cependant, certaines protéines mitochondriales sont codées par l'ADN mitochondrial : 13 gènes codant pour des protéines de la chaîne respiratoire dont 7 du complexe I, 1 du complexe III, 3 du complexe IV et 2 du complexe V

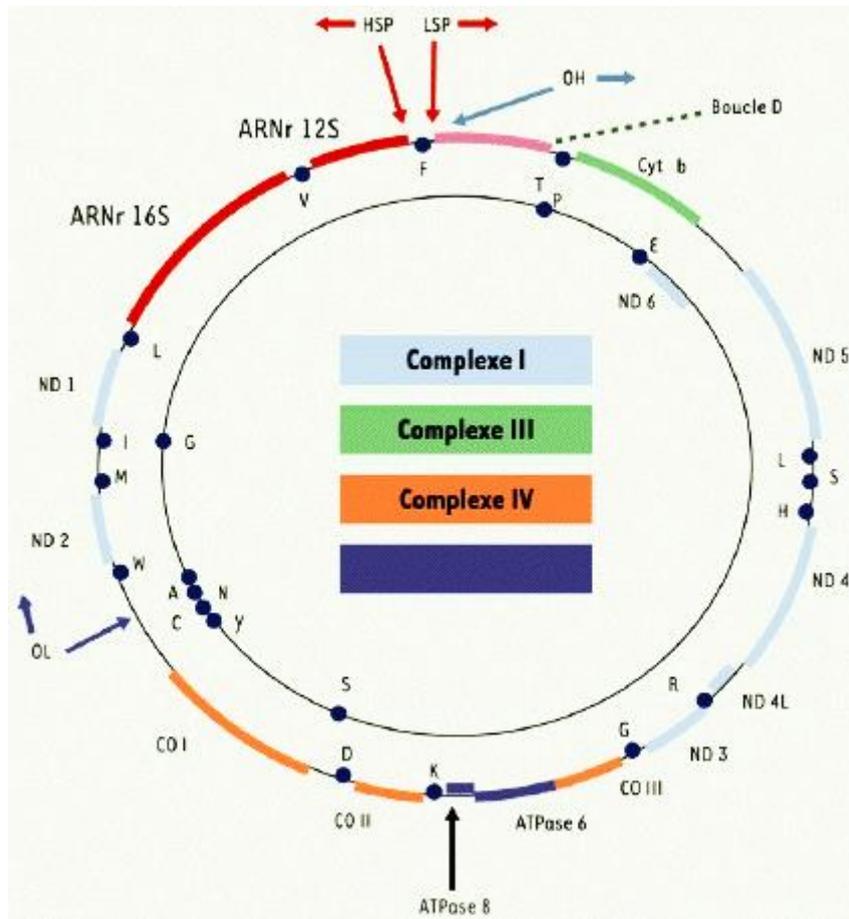


Figure : structure d'ADNmt

L'ADNmt est formé de deux brins: **brin H** (*Heavy*) ou **brin lourd**, riche en bases G (guanine), et, **le brin L** (*Light*) ou **brin léger** riche en bases C (cytosine). Le seul segment non-codant d'ADNmt chez tous les vertébrés est «**la boucle de déplacement**» ou «**D-loop**» ou région de contrôle. Elle contient une origine de réplication du brin H (**OH**) et des sites de régulation pour la transcription du brin L et H. La D-loop a une structure triple-brin, non-codante et c'est une région qui est la plus variable, en séquence, chez différentes espèces. Elle contient deux séquences **HV1 et HV2 (Hyper-Variable)** donc un grand polymorphisme situé respectivement entre les positions nucléotidiques 16024 et 16365, et 73 et 340.

3. Réplication de l'ADN mitochondrial :

Selon le modèle généralement admis, les deux brins H et L de l'ADNmt sont répliqués de manière asynchrone et asymétrique à partir de deux origines distinctes O_H et O_L . La réplication fait intervenir une ADN polymérase γ spécifique de la mitochondrie. La réplication de l'ADNmt est initiée par la synthèse du brin H. La première étape est la production d'une amorce ARN par transcription du brin L qui se fait sous le contrôle du facteur de transcription mitochondrial mtTFA (mitochondrial Transcription Factor A). L'amorce ribonucléotidique forme un hybride ARN/ADN

stable, appelé R-Loop. L'amorce ARN est clivée par une ARNase MRP (Mitochondrial RNA Processing). L'ADN polymérase γ prend alors le relais pour synthétiser le néo-brin H. Différentes enzymes participant à la réplication de l'ADNmt ont été identifiées : une hélicase TWINKLE, des topoisomérases de type I et II et une ADN ligase III.

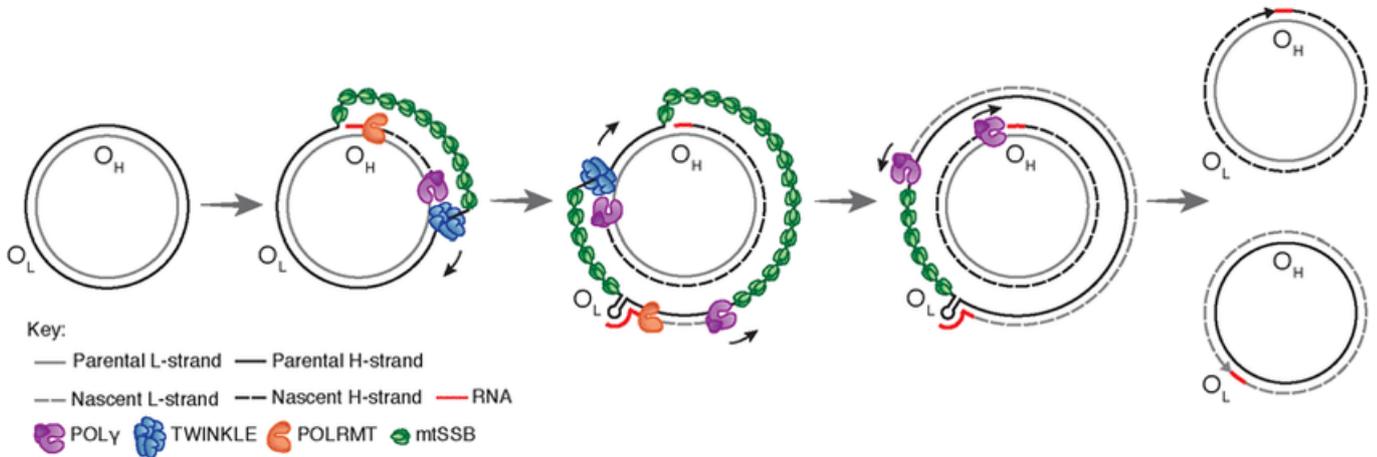


Figure : réplication de l'ADNmt

Généralement, les mitochondries présentes dans les cellules se divisent pendant toute l'interphase, sans synchronisation avec le cycle cellulaire ou avec les autres organites. De même, la réplication de l'ADNmt n'est pas concomitante à la réplication de l'ADN nucléaire lors de la phase S, mais a lieu durant tout le cycle cellulaire.

4. Expression des gènes mitochondriaux :

Chaque brin d'ADNmt est transcrit en un grand ARN poly-cistronique, à la manière des procaryotes. Ce grand transcrit est ensuite clivé par des ARNase spécifiques, au niveau des ARNt, qui ponctuent les gènes structuraux. Puis les Acides Ribo-Nucléiques messagers (ARNm) sont poly-adénylés, permettant la formation de codons STOP pour les ARN se terminant par U ou UA. Ces ARN sont ensuite traduits suivant un code génétique propre à la mitochondrie. L'ensemble de ces étapes est en grande partie assuré par des protéines nucléaires.

Tableau : les modifications du code génétique standard dans les mitochondries

Codon	Code standard : protéine codées par l'ADN nucléaire	Mitochondries			
		Mammifères	Drosophile	<i>Neurospora</i>	Levures
UGA	Stop	Trp	Trp	Trp	Trp
AGA, AGG	Arg	Stop	Ser	Arg	Arg
AUA	Ile	Met	Met	Ile	Met
AUU	Ile	Met	Met	Met	Met
CUU, CUC, CUA, CUG	Leu	Leu	Leu	Leu	Thr

La traduction des ARNm est associée à la membrane interne de la mitochondrie. Bien qu'elle débute par une N-formyl-méthionine et est inhibée par le chloramphénicol (traits typiquement bactériens)

Une autre particularité de la traduction mitochondriale vient de son ribosome. En effet, le ribosome des mitochondries de mammifères est composé d'une petite sous-unité 28S et d'une grande sous-unité 39S. Il est caractérisé par une composition plus faible en ARNr (seulement 33 % contre 66 % dans le ribosome bactérien) et notamment par la perte de certaines régions structurales composées d'ARN

Il est, cependant, intéressant de noter que l'initiation de la traduction ne repose ni sur une séquence de type *Shine-Dalgarno* (bactéries), ni sur la coiffe en 5' (eucaryotes).

II- Les chloroplastes

Les chloroplastes sont des organites présents dans le cytoplasme des cellules eucaryotes photosynthétiques (plantes, algues). Les chloroplastes utilisent l'énergie lumineuse pour synthétiser des hydrates de carbone à partir de CO₂ et H₂O de l'atmosphère. Ils ont leur propre génome. L'ensemble des chloroplastes d'une cellule constitue **le plastidome**

La ressemblance entre un chloroplaste et d'une mitochondrie est confortée par plusieurs caractères :

- l'ADN du chloroplaste est circulaire et non associé à des histones comme chez les mitochondries,
- cet ADN code pour une partie des protéines chloroplastiques (organites semi autonomes),
- une partie de la synthèse de protéines chloroplastiques s'effectue dans le chloroplaste, grâce à la présence de ribosomes
- la division des chloroplastes suit un rythme indépendant de la division du noyau.

1. Structure

Le plaste est en général entouré par deux membranes séparant un espace intermembranaire le tout mesurant 5 à 10µm et régulant les échanges avec le cytosol. La membrane externe est peu sélective car elle contient des porines au contraire de la membrane interne qui observe une régulation stricte des échanges.

Cette double membrane entoure **le stroma** (lumière du plaste) qui est plus acide que le cytosol. Il contient les enzymes de fixation du CO₂ pour la synthèse des molécules organiques; on y observe aussi des grains d'amidon qui sont le résultat de cette synthèse. Il contient aussi de l'ADN et des ribosomes. A l'intérieur du stroma des chloroplastes, se trouve un autre réseau membranaire organisé en sac aplatis, **les thylakoïdes**. Dans certaines régions du chloroplaste, les thylakoïdes sont empilés comme des jetons de poker et forment des structures appelées grana (granum au singulier).

Les thylacoïdes sont le siège de la phase photochimique de la photosynthèse. Elle contient les chlorophylles A et B essentielles à la photolyse de l'eau : $H_2O \rightarrow O + H^+ + e^-$. L'O issu de cette réaction est rejeté sous forme d'O₂. Les thylacoïdes peuvent s'entasser (thylacoïdes granaires) ou rester seuls (intergranaire).

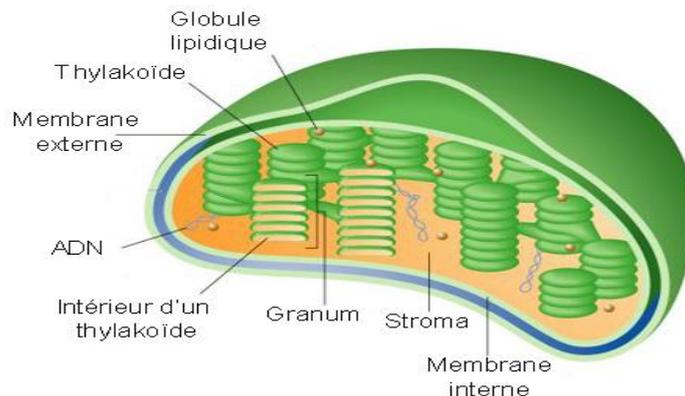


Figure : structure d'un chloroplaste

2. Le génome chloroplastique

En règle générale, la taille des génomes chloroplastiques varie peu d'une plante à l'autre. Aujourd'hui, une proportion seulement de l'ADN chloroplastique total est considérée comme circulaire, les autres formes étant linéaires, ramifiées ou plus complexes. Il est admis que l'ADN chloroplastique a une structure complexe et multigénomique dont la quantité et la structure varient

selon les tissus et les stades de développement des chloroplastes. Néanmoins, l'organisation et la structure des génomes chloroplastiques restent généralement stables et conservées entre espèces végétales.

La structure et la quantité d'ADN chloroplastique peuvent varier en fonction des tissus et des stades de développement des chloroplastes

En général, L'ADN chloroplastique, d'une longueur de 120-160kb, est circulaire et ne contient pas d'histones et (3 à 5) code pour environ 120 à 130 gènes dont les photosystèmes I et II, l'ARNpol, l'ARN ribosomique et L'ARNt (environ 30). Chez la plupart des organismes, le génome chloroplastique comporte 2 régions répétées et inversées (IR, Inverted Repeat) encadrant une région longue (LSC) et une région courte (SSC). Les séquences IR chloroplastiques portent les gènes pour les ARNr. Le génome chloroplastique comporte 120.

Les introns présents dans les gènes chloroplastiques sont classés en deux familles : les introns du groupe I et ceux du groupe II. Cette classification se fait en fonction de la structure primaire et secondaire de l'intron, mais aussi en fonction du mécanisme d'épissage. Les génomes chloroplastiques des végétaux supérieurs contiennent une vingtaine d'introns du groupe II et un intron du groupe I

3. Expression des gènes chloroplastiques

Les gènes des chloroplastes sont souvent organisés en «cluster», et sont transcrits en pré-ARN polycistroniques qui sont ensuite maturés en ARN plus petits. La transcription dépend de deux ARN polymérase d'origine et de fonctionnement différent, ARN polymérase chloroplastique et ARN polymérase nucléaire

L'expression des gènes chloroplastiques peut aussi nécessiter la fixation de facteurs de transcription, codés par le noyau et adressés à des séquences promotrices spécifiques dans le génome chloroplastique.

Exemple :

La séquence comprise entre -16 et -102 nucléotides en amont du site d'initiation de la transcription du gène *rbcL* (gène qui code la grande sous-unité de la Rubisco) est un site de fixation pour un facteur de transcription chloroplastique (CDF1) codé par le génome nucléaire.

Suite à la transcription, différents processus post-transcriptionnels interviennent pour maturer les transcrits. Parmi eux existent l'édition, l'épissage des introns et la maturation des extrémités 5' et 3'

Chapitre 6 : l'hérédité humaine

Introduction

Une maladie congénitale est présente à la naissance. Les maladies qui se développent pendant l'enfance et la vie adulte ne sont pas congénitales.

Une maladie acquise résulte de l'action d'une cause extérieure comme une infection (bactérie, virus, parasite), un empoisonnement ou un accident.

Une maladie génétique résulte du dysfonctionnement d'un ou plusieurs gènes. Quand une maladie résulte du dysfonctionnement d'un seul gène, elle est dite **monofactorielle** ou **monogénique** (ces deux termes sont équivalents).

Une maladie génétique peut ne pas être héréditaire : par exemple, la plupart des cancers qui résultent de mutations affectant des gènes dans les cellules tumorales, cellules somatiques qui ne participent pas à la reproduction sexuée.

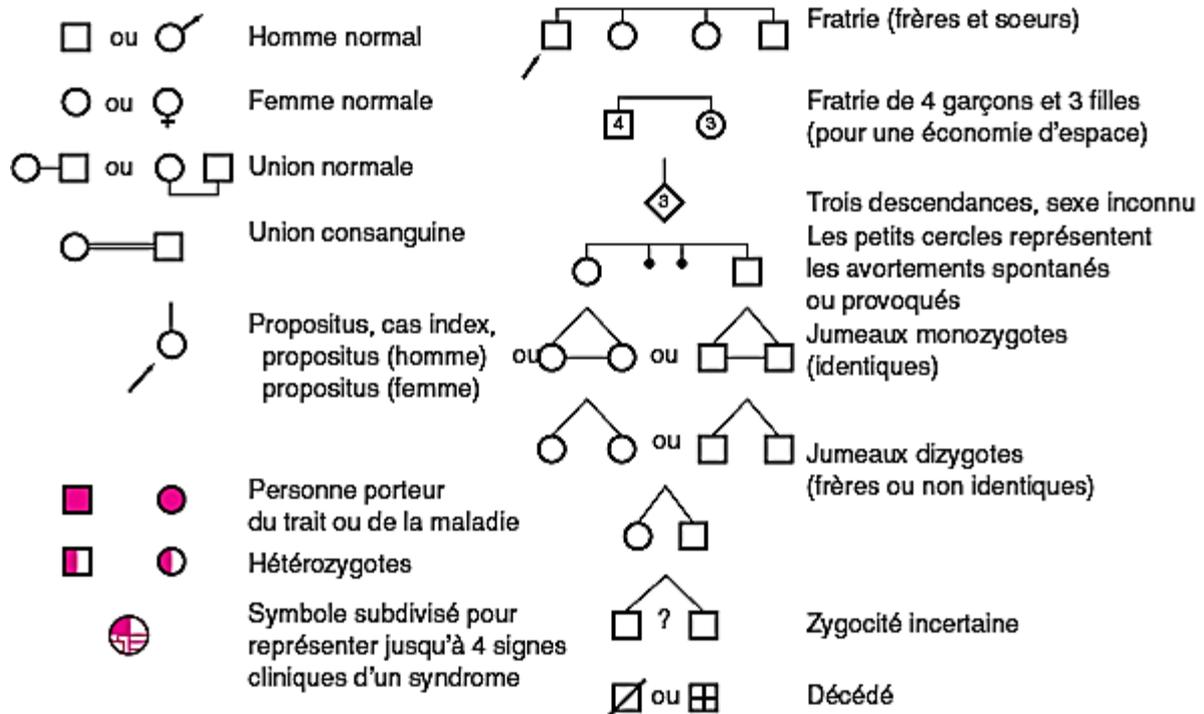
Le terme de **maladie héréditaire** est aujourd'hui réservé aux **maladies génétiques** et on préfère utiliser le terme de **maladie transmissible** quand la cause n'est pas génétique, par exemple les maladies sexuellement transmissibles ou liées à un facteur constant du milieu (insuffisance en iode conduisant au crétinisme).

Ces définitions ne sont pas obligatoirement contradictoires : certaines maladies génétiques sont congénitales et d'autres ne le sont pas, certaines maladies génétiques sont héréditaires et d'autres ne le sont pas.

Pour étudier la transmission de caractère héréditaire, il faut étudier un arbre généalogique : **une représentation schématique des liens familiaux. Un arbre généalogique** est une représentation graphique de la généalogie ascendante ou descendante d'un individu (cas index : celui sur lequel porte la généalogie).

La généalogie d'une famille (arbre généalogique) peut être utilisée pour tracer des défauts génétiques transmissibles. Il est également couramment utilisé dans le conseil génétique. L'arbre généalogique fait appel à des symboles conventionnels pour représenter les membres de la famille et les informations pertinentes concernant leur santé (Symboles pour construire l'arbre généalogique d'une famille.). Certaines pathologies familiales ayant des phénotypes identiques ont de multiples modes de transmission différents.

Dans l'arbre généalogique, des symboles pour chaque génération dans la famille sont placés sur une colonne et numérotés en chiffres romains, en commençant par la génération la plus ancienne au sommet et s'achevant avec les plus récentes en bas. Dans chaque génération, les personnes sont numérotées de gauche à droite en chiffres arabes. La fratrie est listée habituellement selon l'âge avec le plus âgé à gauche. Ainsi, chaque membre de l'arbre généalogique peut être identifié par 2 numéros (ex : II 4). Au conjoint est également attribué un nombre d'identification.



Hérédité mendélienne : Le mode de transmission d'une maladie génétique **monofactorielle, ou monogénique**, suit les lois de Mendel, ce qui explique l'usage du terme "maladie mendélienne".

Les modes de transmission d'une maladie mendélienne se distinguent par des probabilités de risque très caractéristiques. On définit pour les maladies mendéliennes quatre modes de transmission, autosomique ou lié à l'X, selon que le gène impliqué est localisé sur un autosome ou sur le chromosome X ; dominant ou récessif, selon que la maladie est dominante ou récessive.

1- Maladies autosomiques dominantes :

1-1- Définition

Une maladie est transmise selon le mode autosomique dominant si le gène en cause est porté par un autosome et si la présence d'un seul allèle muté suffit pour que la maladie se manifeste. Les individus hétérozygotes (A/a) pour le gène en cause sont malades. Généralement, les individus

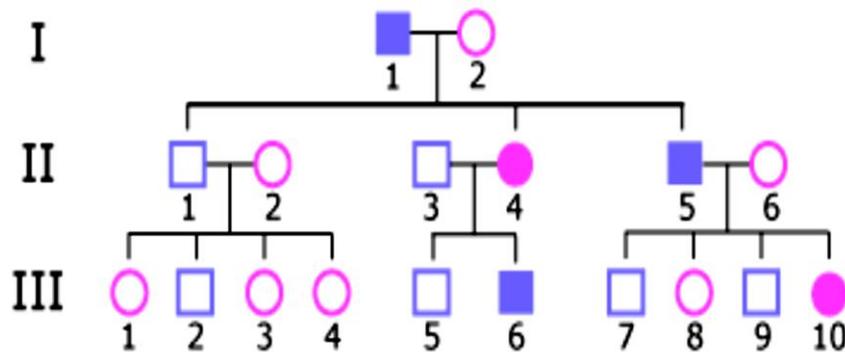
homozygotes (A/A), s'ils sont viables, sont plus sévèrement atteints par la maladie; ils sont si rares qu'on peut considérer que tous les atteints sont, sauf exception, hétérozygotes.

1-2- Caractéristiques

- ✓ Il y a autant de filles que de garçons atteints parce que le gène impliqué est sur un autosome qui peut aussi bien se trouver chez un garçon que chez une fille.
- ✓ Une personne malade a un de ses deux parents atteint.
- ✓ Une personne malade a un risque de 50% de "transmettre" la maladie à chacun de ses enfants. Le fait d'avoir un premier enfant sain ne signifie pas que le deuxième enfant sera malade. A chaque grossesse, le risque que l'enfant soit malade est de 50%, parce qu'à chaque fois, l'individu atteint a un risque sur deux de transmettre l'allèle muté (A) et une chance sur deux de transmettre l'allèle normal (a).
- ✓ La transmission des maladies autosomiques dominantes s'effectue sans saut de génération (transmission verticale)
- ✓ Une personne qui n'a pas l'allèle pathologique ne peut pas "transmettre" la maladie à ses enfants.

Exemple : Diabète familial du jeune adulte (MODY)

Le MODY (Maturity Onset Diabetes of the Young) est une forme de diabète familial, à transmission autosomique dominante et à début précoce, associé à des anomalies primaires de l'insulinosécrétion.



Analyse de l'arbre

- ✓ Un des deux parents d'un individu malade est atteint.
- ✓ La transmission est verticale (pas de saut de génération).
- ✓ Hommes et femmes sont malades.
- ✓ La transmission est identique quel que soit le sexe du parent atteint.
- ✓ Une personne saine ne peut pas "transmettre" la maladie.

Ces observations sont conformes au mode autosomique dominant.

Risques pour la descendance

- ✓ Chacun des individus malades a un risque de 50% de "transmettre" la maladie à chaque nouvelle grossesse.
- ✓ Chacun des individus sains a un risque de 0% de "transmettre" la maladie à chaque nouvelle grossesse.

2- Maladies autosomiques récessives

2-1- Définition

Une maladie est transmise selon le mode autosomique récessif si le gène en cause est porté par un autosome et si la présence de deux allèles mutés du gène est nécessaire pour que la maladie se manifeste. Les malades sont tous des homozygotes pour le gène en cause.

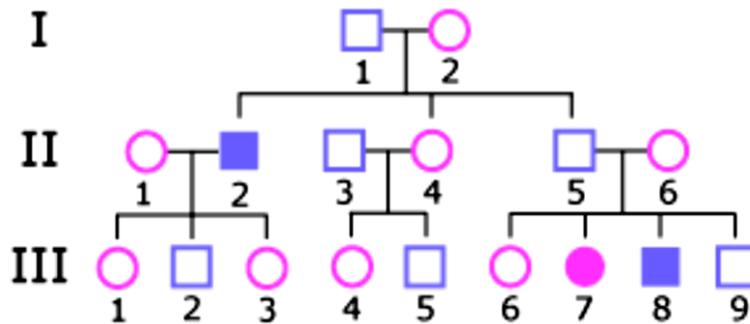
2-2- Caractéristiques

- Il y a autant de filles que de garçons atteints parce que le gène impliqué est sur un autosome.
- Il n'y a pas de personnes malades à toutes les générations, car la plupart du temps, les sujets atteints naissent de parents hétérozygotes, porteurs sains (A/a).
- Un couple à risque est formé par deux conjoints porteurs sains hétérozygotes (A/a). Il a, à chaque grossesse :
 - ❖ un risque de 25% d'avoir un enfant atteint (homozygote a/a).
 - ❖ une probabilité de 50% d'avoir un enfant porteur sain (hétérozygote A/a) qui peut avoir un enfant atteint si, et seulement si, son conjoint est lui-même porteur sain (avec un risque de 1/4).
 - ❖ une probabilité de 25% de donner naissance à un enfant sain (homozygote A/A) qui ne peut pas avoir d'enfant atteint.
- ✓ Un sujet malade qui se marie à un sujet normal donne habituellement naissance à des enfants normaux car les porteurs sains (A/a) sont beaucoup plus rares que les individus sains homozygotes (A/A).
- ✓ La maladie, du fait de la faible dimension des familles humaines, peut ne toucher qu'une personne dans une famille. Le cas isolé ne signifie donc pas nécessairement cas de novo (mutation survenue dans la lignée germinale de l'un des parents).

Exemple : Mucoviscidose

Dans la famille ci-dessous, III-7 et III-8 sont atteints de mucoviscidose, la maladie autosomique récessive la plus fréquente en Europe (un nouveau-né sur 3 000 environ). Ils sont porteurs de deux allèles pathologiques, et leurs parents II-5 et II-6 sont hétérozygotes porteurs sains.

Dans la plupart des cas, on observe très rarement d'autres individus atteints parmi les ascendants (comme ici II-2), aussi bien pour la mucoviscidose que pour toutes les autres maladies récessives qui sont plus rares qu'elle.



Analyse de l'arbre

- ✓ Aucun des parents d'un malade n'est atteint par la maladie.
- ✓ Il n'y a pas de personnes malades à toutes les générations.
- ✓ Les hommes et les femmes sont également atteints.

Ces observations sont conformes au mode autosomique récessif.

Risques pour la descendance

- ✓ Le risque de "transmettre" la maladie pour une personne homozygote malade (a/a) ou pour une personne apparenté sain hétérozygote (A/a) dépend du statut de son conjoint.
- ✓ Le risque de "transmettre" la maladie pour une personne homozygote saine (A/A) est nul.

3- Maladies liées à l'X

3-1- Particularités de l'hérédité liée au chromosome X

- ✓ Les individus de sexe masculin n'ont qu'un seul chromosome X : ils sont hémizygotés et ne possèdent qu'un seul exemplaire des gènes du chromosome X.
- ✓ Par contre, les femmes possèdent deux chromosomes X et deux exemplaires de chacun des gènes localisés sur le chromosome X.

3-2- Conséquence pour les maladies liées à un gène du chromosome X

La question de la dominance ou de la récessivité ne se pose pas chez les individus de sexe masculin :

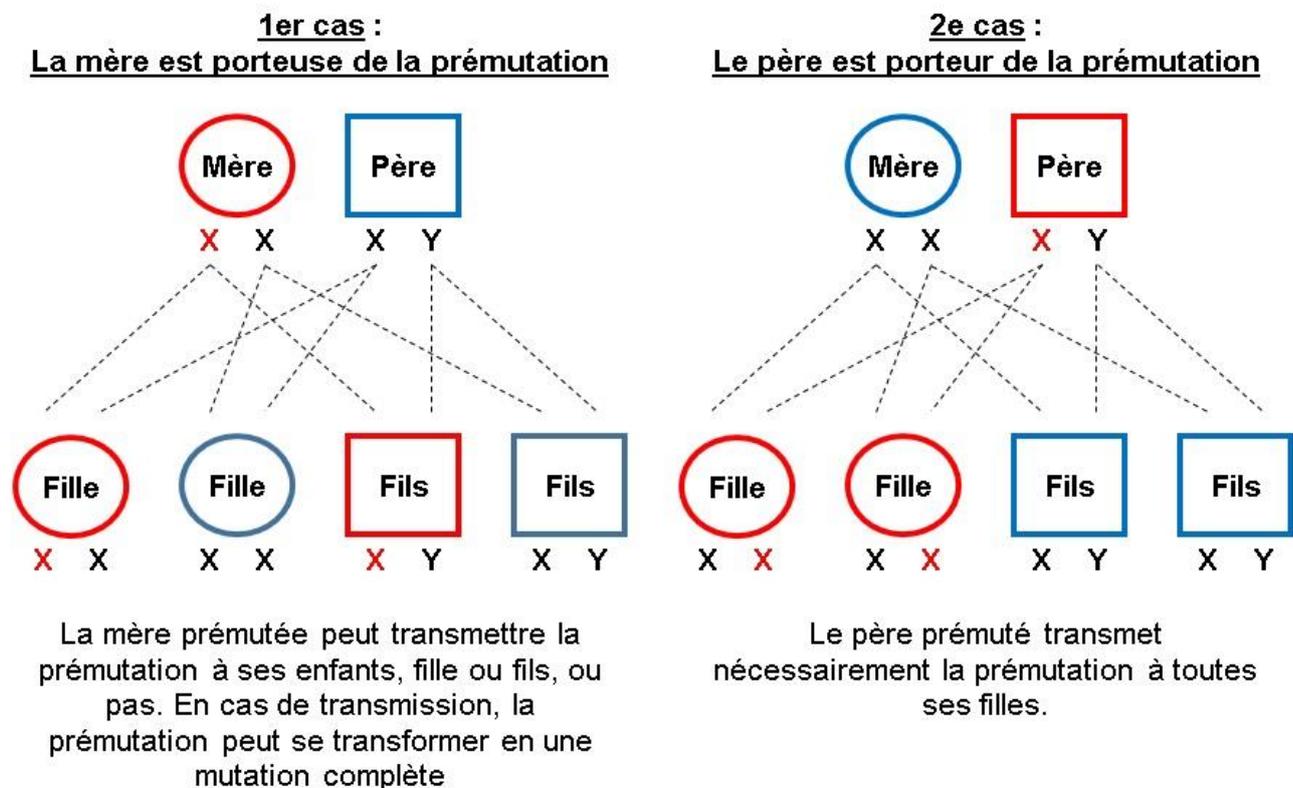
- ❖ Soit le gène est muté : ils sont atteints,
- ❖ Soit le gène est normal : ils sont sains.

La question de la dominance ou de la récessivité ne se pose que chez les individus de sexe féminin :

- ❖ Si la maladie survient quand un seul gène est muté, elle est dominante,
- ❖ Si la maladie survient seulement quand les deux exemplaires sont mutés, elle est récessive.

3-3- Risques de transmission

La transmission des maladies dominantes liées à l'X est différente suivant que le parent atteint est le père ou la mère



3-4- Maladies dominantes liées à l'X

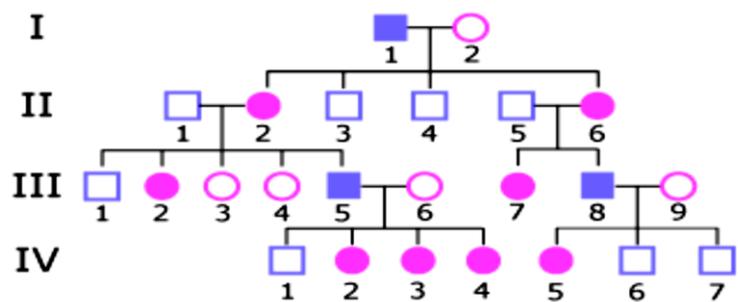
Caractéristiques

- ✓ Les deux sexes peuvent être touchés par la maladie. Mais en général, les filles hétérozygotes sont moins sévèrement malades que les garçons hémizygotés.

- ✓ Il y a plus de femmes atteintes que d'hommes.
- ✓ La transmission est verticale (si la maladie est compatible avec la fécondité, il y a des malades à toutes les générations).
- ✓ Elle diffère de l'hérédité autosomique dominante car il n'y a jamais de transmission père-fils.
- ✓ Comme pour l'hérédité autosomique dominante, la pénétrance peut être incomplète et l'expressivité peut être variable

Exemple : Rachitisme vitamino-dépendant

Dans la famille suivante, plusieurs personnes sont atteintes de rachitisme vitamino-résistant, une maladie qui se traduit par un déficit en phosphate à l'origine de fragilités osseuses.



Analyse de l'arbre

- On constate que toutes les filles d'un homme atteint sont atteintes, mais qu'il n'y a pas de transmission père-fils.
- Par contre, tous les enfants d'une femme atteinte ne sont pas malades. Il y a plus de femmes atteintes que d'hommes atteints.

Ces observations sont conformes au mode dominant lié à l'X.

Risques pour la descendance

- Les hommes atteints "transmettront" la maladie à toutes leurs filles mais à aucun de leurs garçons.
- A chaque grossesse d'une mère atteinte, le risque que l'enfant, fille ou garçon, soit malade est de 50%.
- Les enfants sains ne "transmettent" pas la maladie.

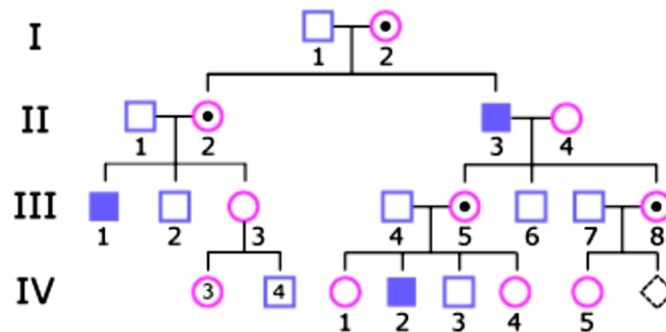
4- Maladies récessives liées à l'X

4-1- Caractéristiques

- Les homozygotes pour l'allèle pathologique étant très rares dans le sexe féminin, les individus atteints sont essentiellement, voire exclusivement des hommes.

Exemple : Hémophilie

Quatre membres de la famille suivante sont atteints d'hémophilie, une maladie caractérisée par une absence de coagulation sanguine due à la mutation d'un gène codant pour un facteur de coagulation.



Analyse de l'arbre

- ✓ Dans cette famille, on constate que seuls les hommes sont atteints. Il n'y a aucune transmission père-fils. Toutes les filles d'un homme malade sont conductrices.
- ✓ La moitié environ des fils d'une femme conductrice sont malades. Toutes les filles d'une femme conductrice ne sont pas conductrices.

Ces observations sont conformes au mode récessif lié à l'X.

Risques pour la descendance

- ✓ Un père malade ne "transmettra" la maladie à aucun de ses fils mais toutes ses filles seront conductrices.
- ✓ Une femme hétérozygote conductrice a un risque de 50% de transmettre son chromosome X muté à chacun de ses enfants. Dans ce cas, la moitié des garçons seront malades et la moitié des filles seront conductrices.

5- Cas particuliers

On dit qu'une maladie mendélienne présente :

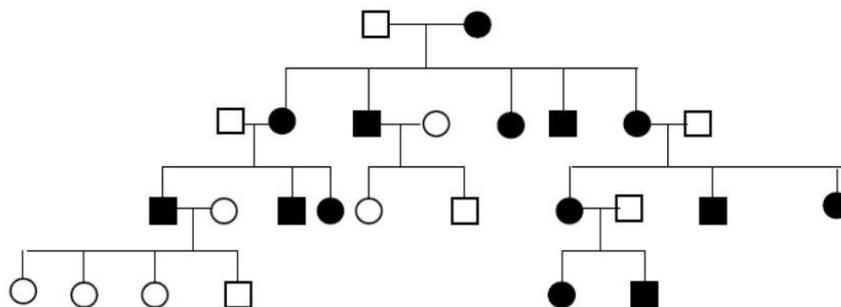
- **une pénétrance incomplète** quand on peut avoir le génotype à risque sans être atteint de la maladie,
- **une expressivité variable** quand, pour un même génotype à risque, la maladie peut prendre différentes formes,
- **une empreinte parentale** quand la maladie dépend du fait que la mutation responsable a été transmise par le père ou par la mère.

✚ Hérité mitochondriale

Lors de la fécondation, le spermatozoïde apporte un noyau d'origine paternelle qui va fusionner avec le noyau de l'ovule, d'origine maternelle, mais le cytoplasme de l'œuf ainsi réalisé est exclusivement d'origine maternelle. Par conséquent les maladies mitochondriales sont des maladies à transmission maternelle exclusive (mode de transmission non mendélien).

Exemple :

Hérité mitochondriale



- Les individus sont atteints pendant de nombreuses générations
- Hommes et femmes sont également atteints
- Un père atteint n'a pas d'enfants atteints
- Une mère atteinte transmet la maladie à tous ses enfants (mais la pénétrance peut être incomplète...)

✚ Hérité multifactorielle

Une maladie multifactorielle n'est pas mendélienne parce qu'elle dépend à la fois de plusieurs gènes simultanément et de facteurs de l'environnement. Sa transmission ne présente donc pas les probabilités de risque observables chez les maladies mendéliennes.

Dans une maladie multifactorielle, c'est la combinaison particulière d'allèles « normaux » de certains gènes qui est pathologique (comme peut être pathologique l'association de plusieurs médicaments, sans danger pris isolément).