

BIOLOGIE MOLÉCULAIRE

PAES

2010-2011

Dr. NAÏMI

Reproduction interdite sans l'accord des auteurs - Faculté de Médecine - UNS

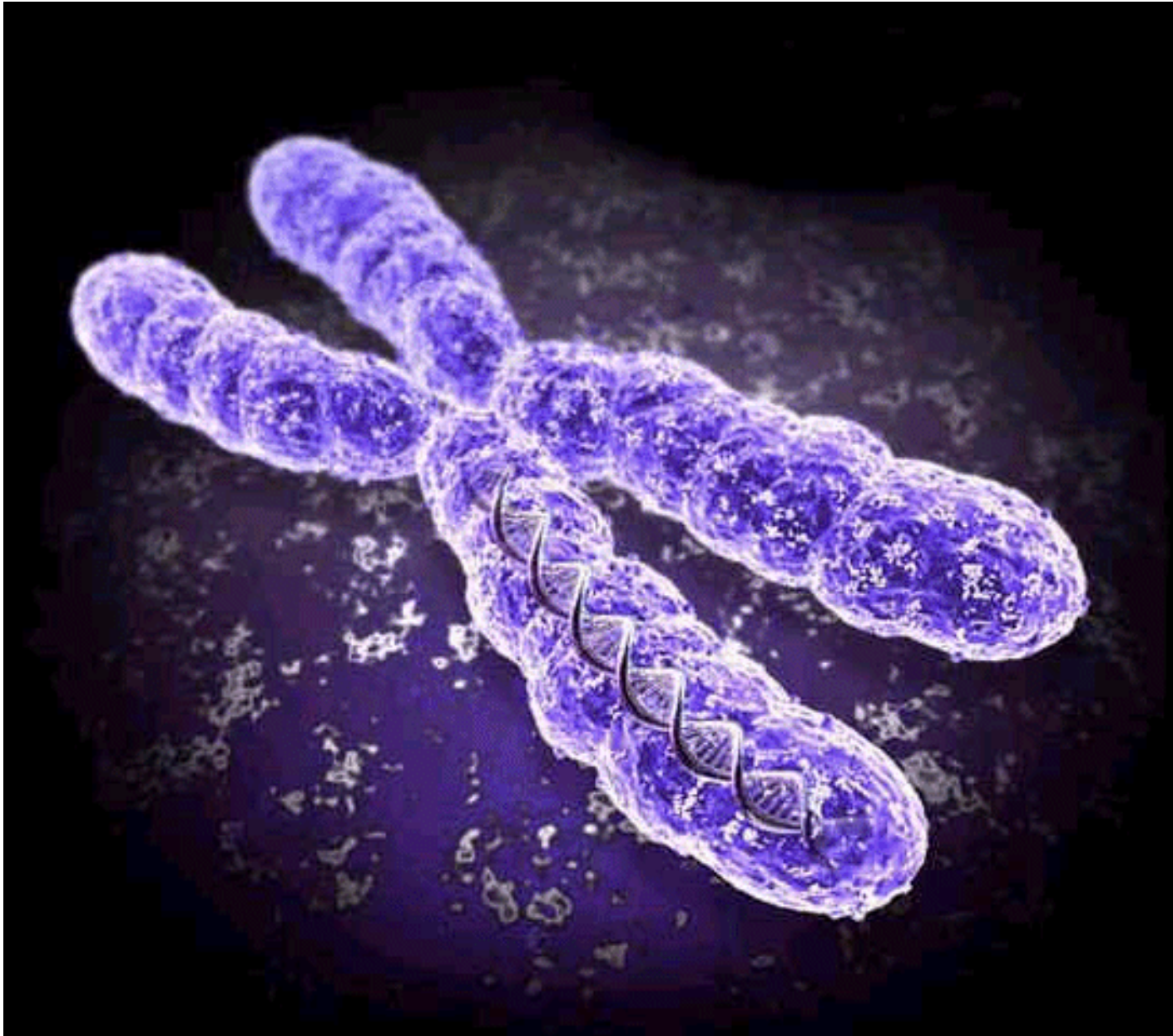
PLAN DU COURS

- Bases essentielles de la notion d'hérédité
- Structure et Organisation du Génome humain
- Réplication du génome
- Expression du génome
 - Transcription et Traduction
 - Régulation de l'expression du génome
- Maintenance du génome
 - Mutations et Systèmes de réparation

Organisation du génome nucléaire

- **Organisation physique/spatiale du génome nucléaire**
 - Différents niveaux de compaction du génome
 - Variabilité de la compaction et rôle dans l'expression génique
 - Régulation épigénétique de la compaction
- **Organisation fonctionnelle du génome nucléaire**
 - Différences entre procaryotes et eucaryotes
 - Nature et rôles des séquences transcrites non codantes
 - Nature et rôles des séquences non transcrites

Le génome nucléaire eucaryote



≠^{ces} entre **Génome eucaryote/procaryote**

- ≠^{ces} **d'organisation physique et spatiale**
 - Le **génome procaryote**: **cytosolique, circulaire et nu** (absence de protéines de structure)
 - Le **génome eucaryote nucléaire**: **linéaire, associé à des protéines de structure** favorisant sa compaction (histones)
 - Deux mètres d'ADN contenus dans un noyau de 10 µm
 - Régions peu compactées et accessibles: **Euchromatine**
 - Régions très compactées et inactives: **Hétérochromatine**
 - **Hétérochromatine**
 - Hétérochromatine constitutive, permanente (centromères, télomères)
 - Hétérochromatine facultative, variable
 - Inactivation de l'X
 - Gènes dont l'expression est réprimée avec la différenciation cellulaire

Généralités sur le noyau

- **Fonctions principales**

- Stockage, Réplication et Transcription du génome

- **Structure du noyau**

- Délimité par une double membrane, interrompue par pores

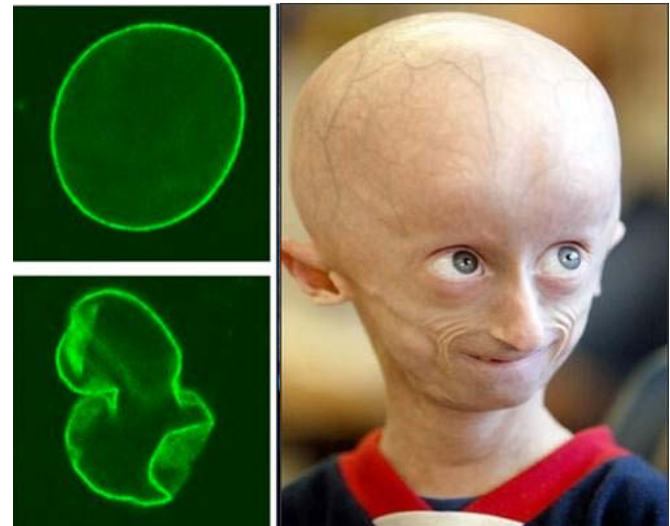
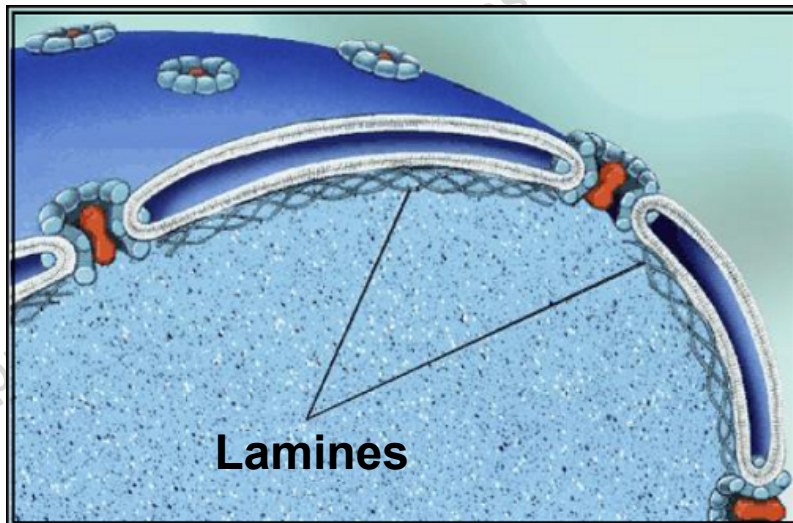
- Contient un ou plusieurs nucléoles

- Site d'expression des gènes d'ARNr (Chr. 13,14,15, 20 et 21)

- Architecture nucléaire maintenue par les lamines

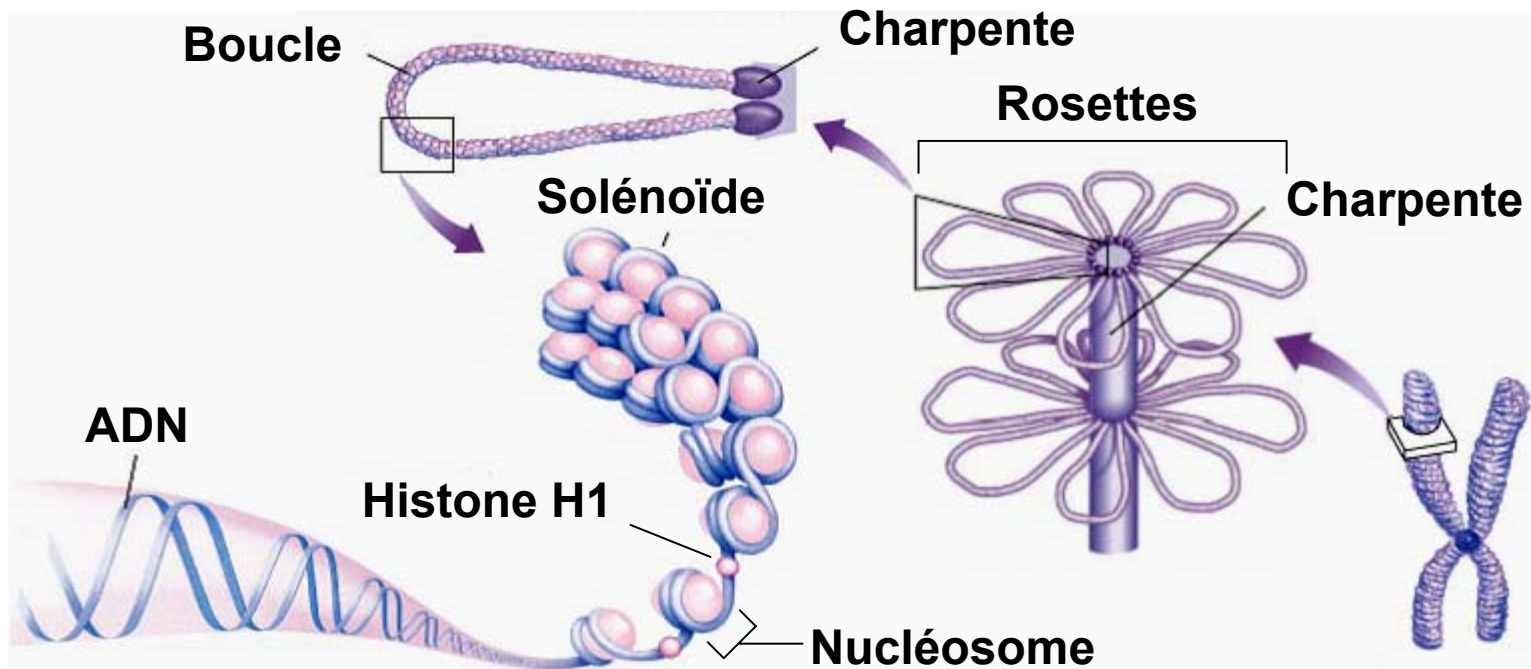
- Lamines, rôle dans l'organisation du génome

- Progéria: Maladie liée anomalie lamine → Vieillesse accélérée



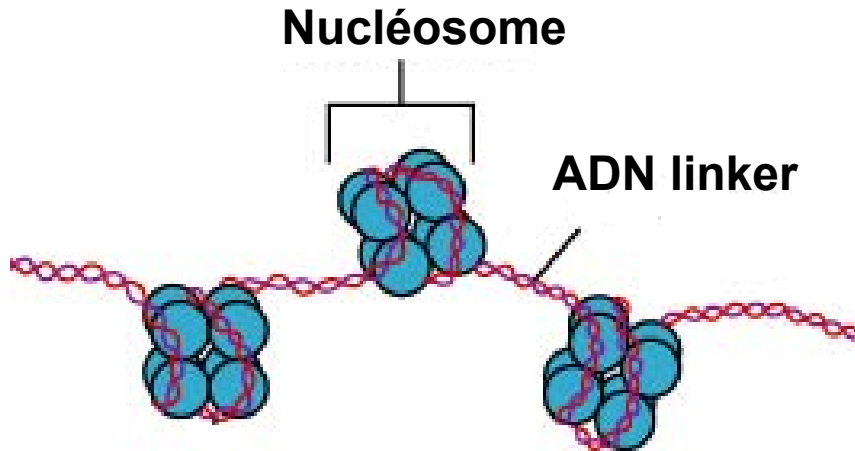
Organisation physique du génome

- Des protéines permettent la compaction de l'ADN
 - Histones: chargées positivement (riches lysine et arginine)
 - Histone H2A, H2B, H3 et H4: Nucléosome
 - Histone H1: Solénoïde
 - Autres protéines = Charpente du chromosome
 - Boucles de chromatine
 - Rosettes de boucles

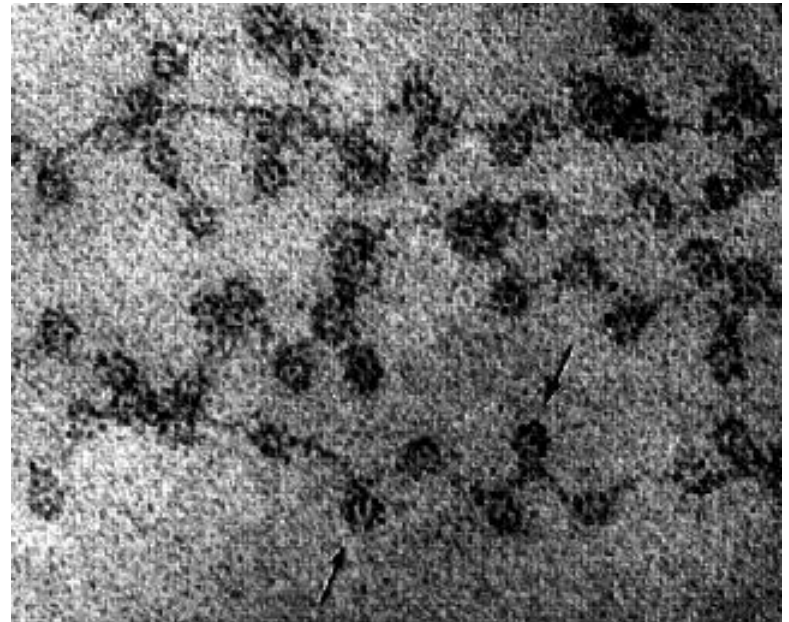


Organisation physique du génome

- **Le nucléosome, premier niveau de compaction**
 - Octamère d'histones: dimère d'H2A, H2B, H3 et H4
 - Deux tours d'ADN: ~145 paires de bases (pb)
 - ADN linker, entre les nucléosomes: ~ 55 pb
- **Forme la fibre de chromatine (10 nm de diamètre)**
 - « collier de perles »

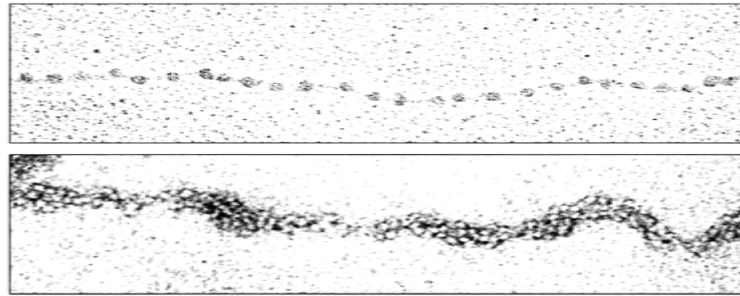


Microscopie électronique



Organisation physique du génome

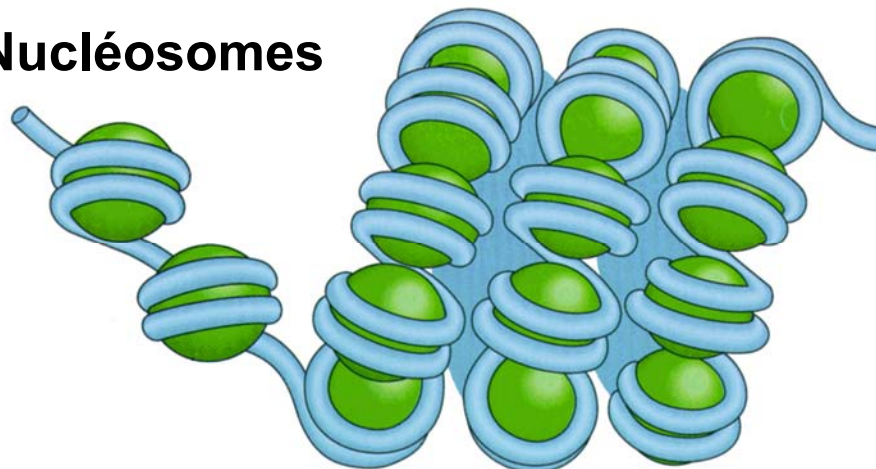
- **Le solénoïde, second niveau de compaction**
 - Hélice de 30nm de diamètre
 - Association de six nucléosomes/tour grâce à l'histone H1



Fibre chromatinienne

Solénoïde

Nucléosomes



Solénoïde

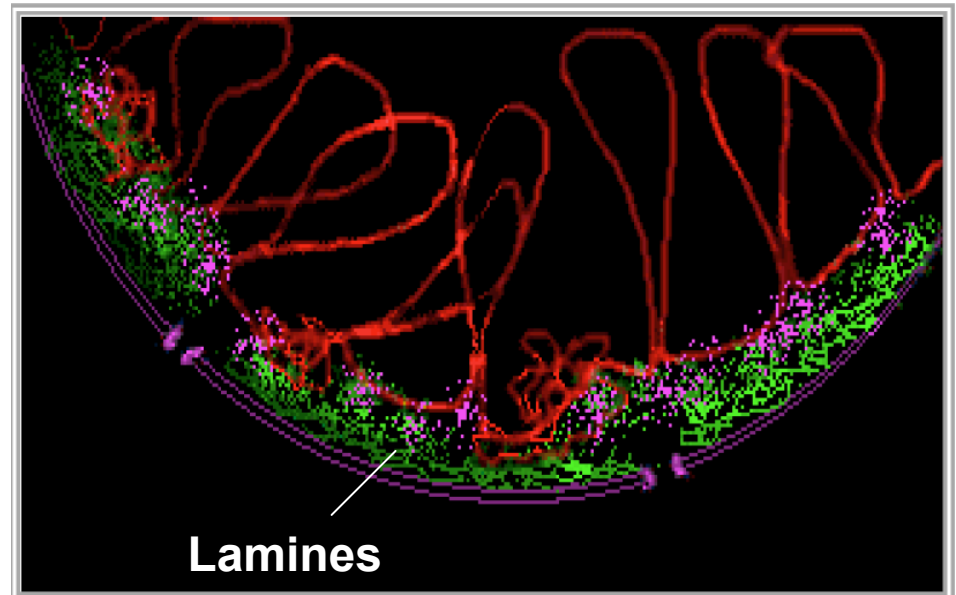
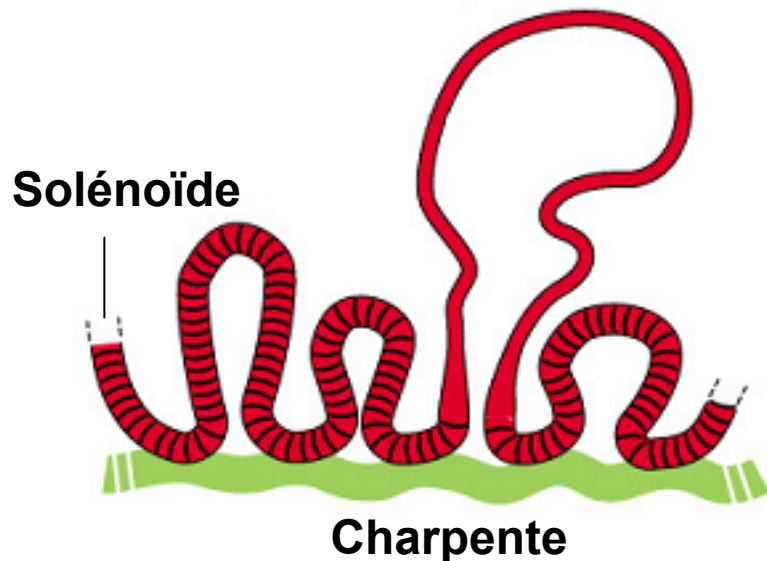
Reproduction in+

Organisation physique du génome

- **Les boucles de chromatine**

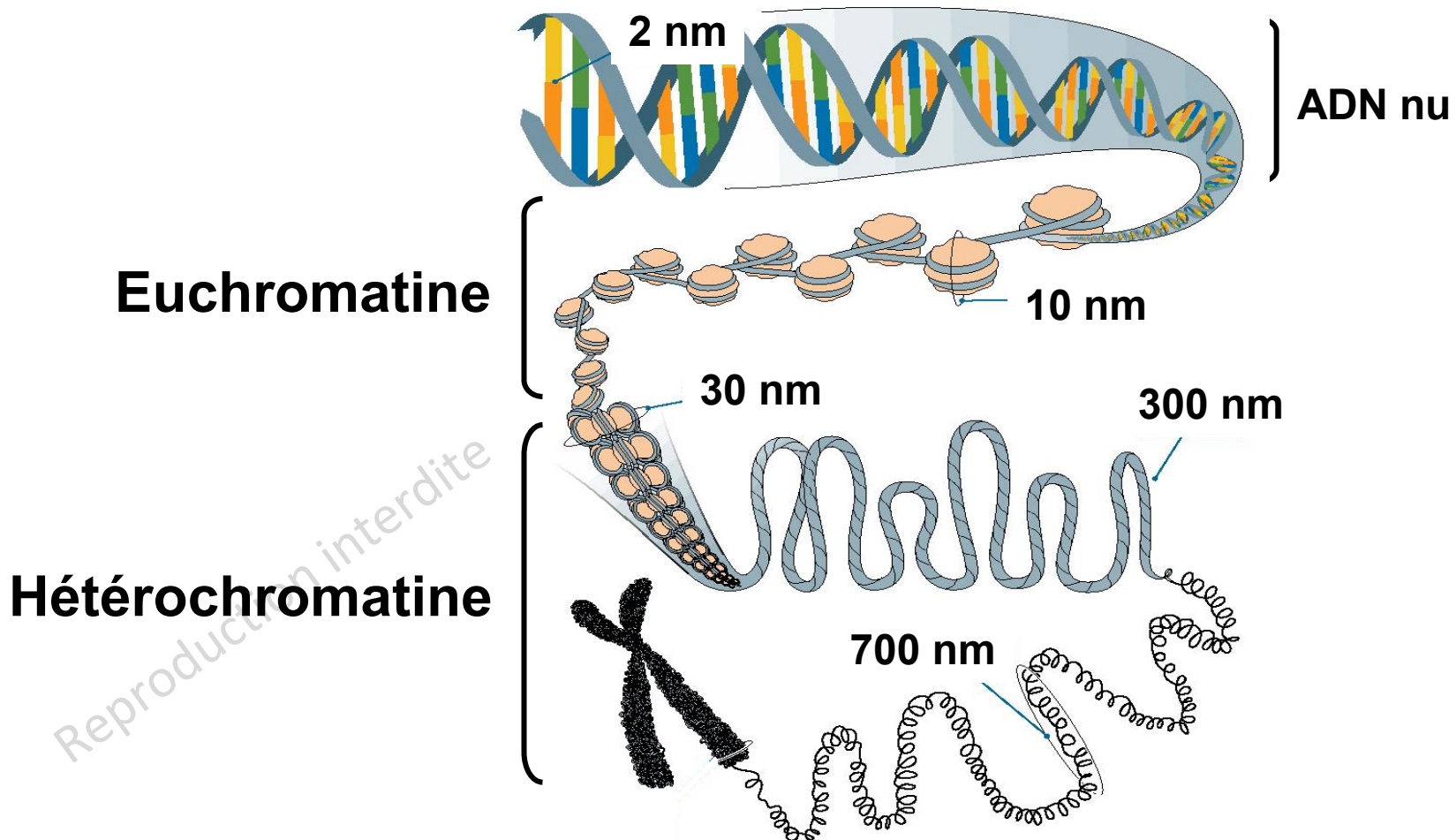
- Le solénoïde s'attache par séquences d'ADN appelées **SAR** (**Scaffold Attachment region**) sur la charpente protéique
- **Chaque boucle** délimite une **région d'ADN autonome** c.à.d. dont les gènes sont régulés de façon autonome
 - Les séquences SARs = **rôle d'insulateurs**, c.a.d qui isolent une région d'ADN du reste du génome
- La charpente est constituée entre autre par les lamines

Boucle d'euchromatine



Organisation physique du génome

- La compaction de l'ADN est variable
 - Euchromatine = Fibre de chromatine de 10nm
 - Hétérochromatine = Niveaux de compaction supérieurs

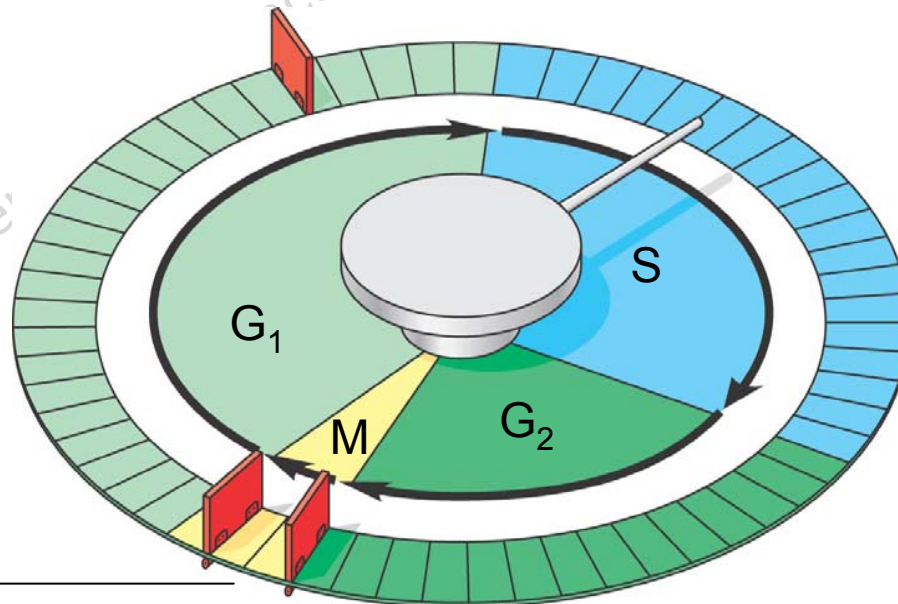
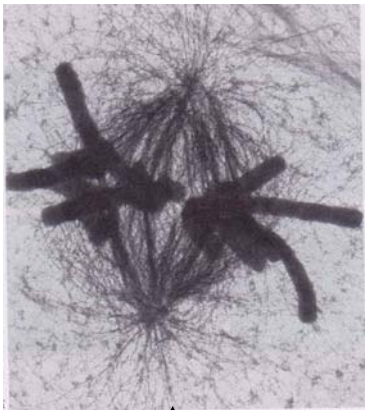


Variabilité de la compaction de l'ADN

- **Au cours du cycle cellulaire**

- **Interphase**: ADN répliqué et transcrit pour préparer division
→ il est sous **forme relâchée (euchromatine)**
 - Phase **G₁**, **S** et **G₂**
- **Mitose (M)**: Chromosomes répartis entre cellules filles →
l'ADN est sous **forme condensée (hétérochromatine)**
 - Prophase, métaphase, anaphase, télophase

Mitose

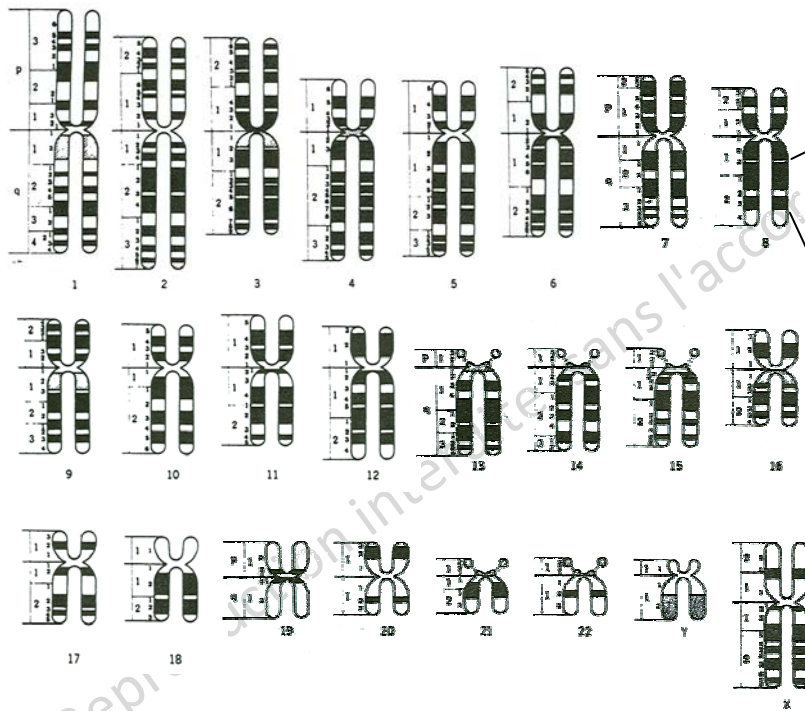


Interphase

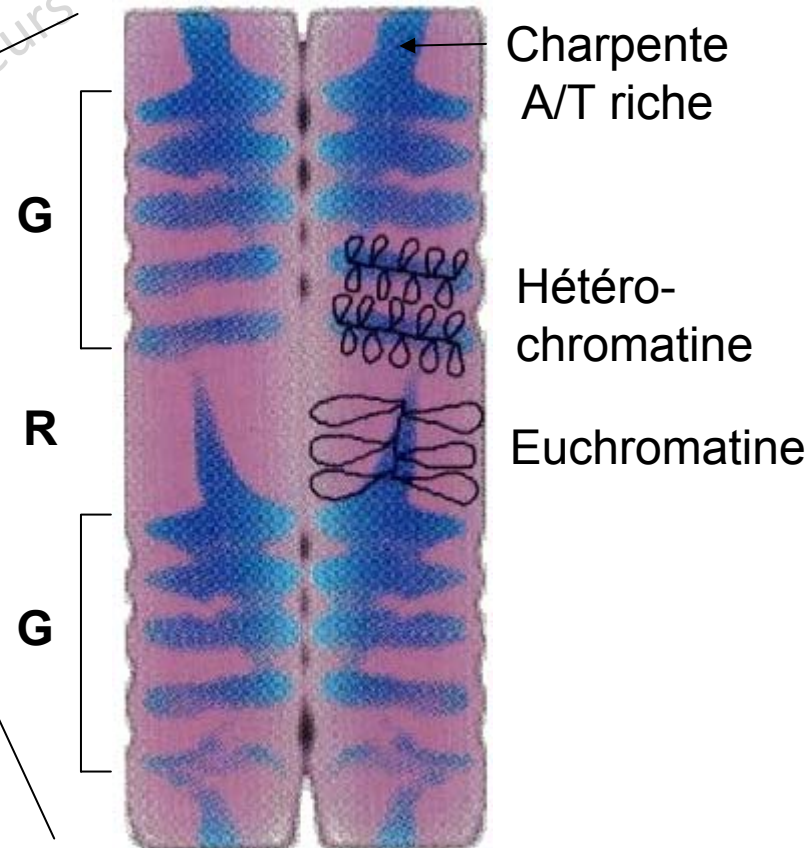


Variabilité de la compaction de l'ADN

- Elle est visible sur le caryotype métaphasique
 - Les chromosomes sont classés grâce à des colorants qui révèlent une **alternance de bandes** caractéristique
 - Bandes G (Giemsa, régions A/T riches) et Bandes R sont en miroir
 - = **Alternance de régions Hétérochromatiques / Euchromatiques**

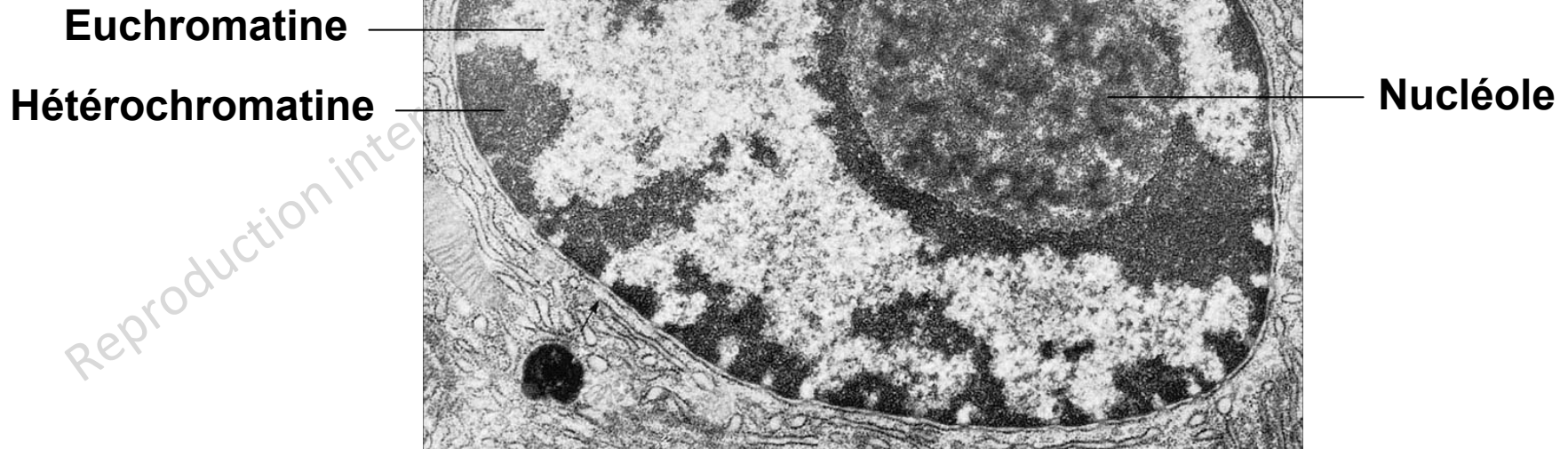


- Bandes G (-), Bandes R (+)
- Bandes G (+), Bandes R (-)



Variabilité de la compaction de l'ADN

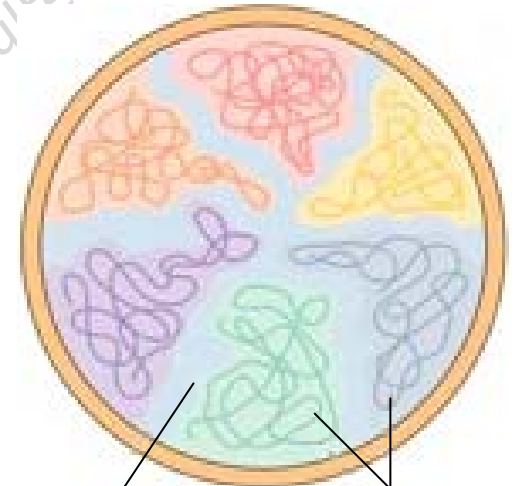
- Elle est visible dans le noyau interphasique
 - L'hétérochromatine est à la périphérie du noyau
 - L'euchromatine est plutôt au centre du noyau
 - Il existerait un compartiment central dédié à l'expression génique
 - → Organisation spatiale du génome



Organisation spatiale du génome

- **Notion de territoire chromosomique**

- Chaque chromosome à territoire défini dans le noyau
- Entre les territoires, domaines interchromosomiques



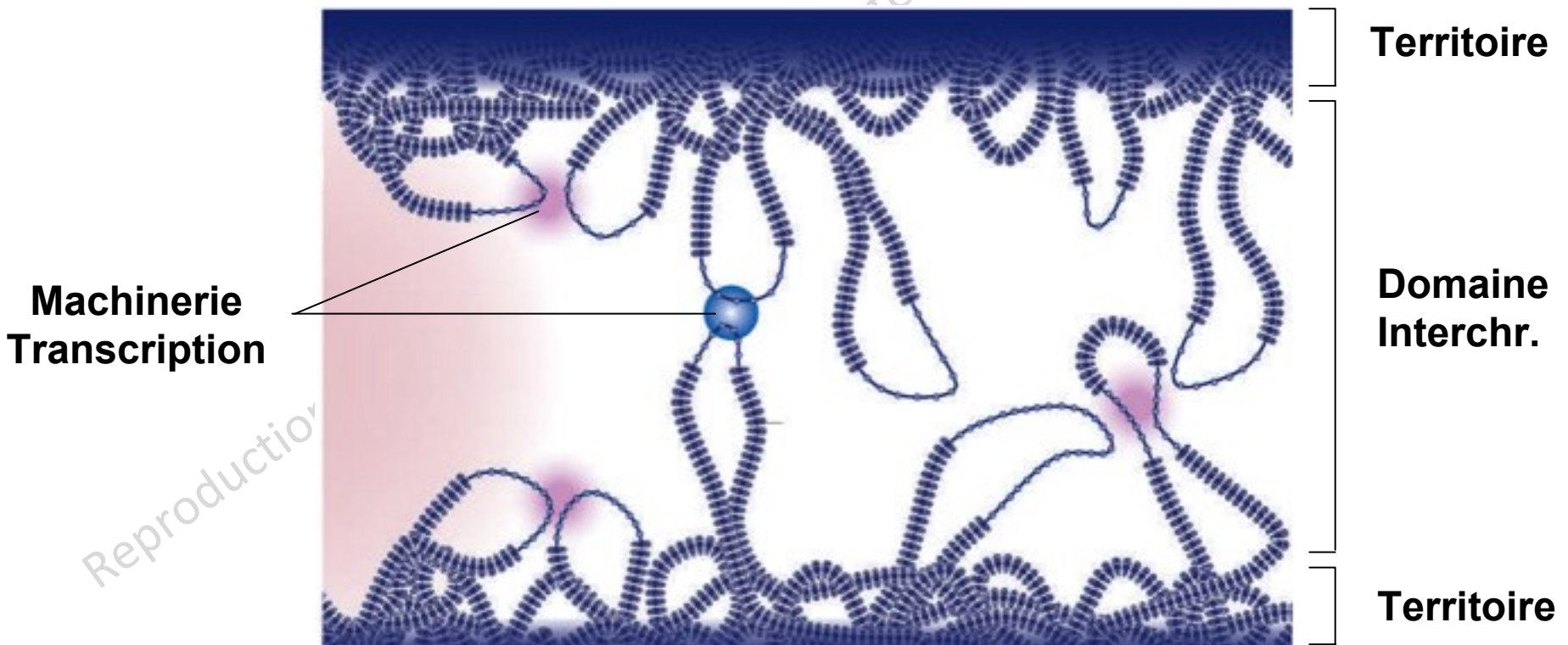
Domaine
Interchromosomique

Territoires
chromosomiques

- Cette organisation joue un rôle dans l'expression génique

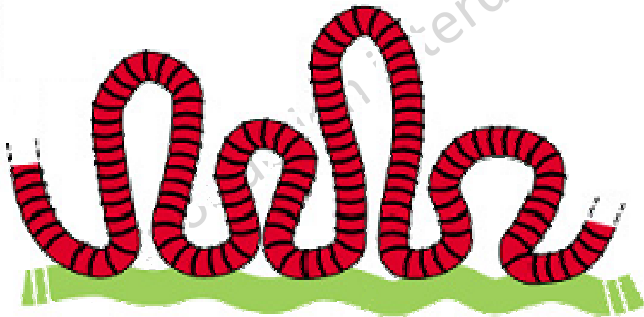
Organisation spatiale du génome

- Elle joue un rôle dans l'expression génique
 - Les **domaines interchromosomiques** contiennent les enzymes de la **machinerie de transcription / maturation**
 - Les **boucle de chromatine** décompactées et riches en gènes (euchromatine) sont à **proximité de ces domaines**

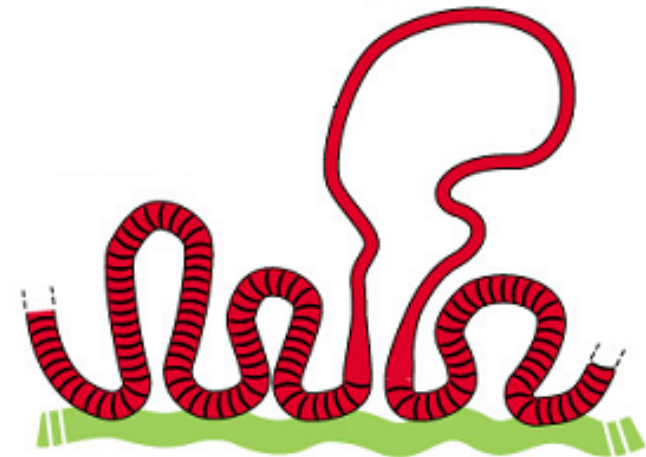
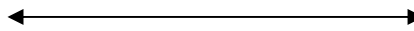


Régulation de la compaction de l'ADN

- Variants d'histones
- **Modifications épigénétiques**
 - Ne changent pas la séquence de l'ADN
 - **Affectent l'expression des gènes**
 - Sont **transmissibles** au cours des divisions cellulaires
 - Font intervenir
 - **Modifications post-traductionnelles des histones**
 - **Méthylation de l'ADN**
 - ARNs non codant

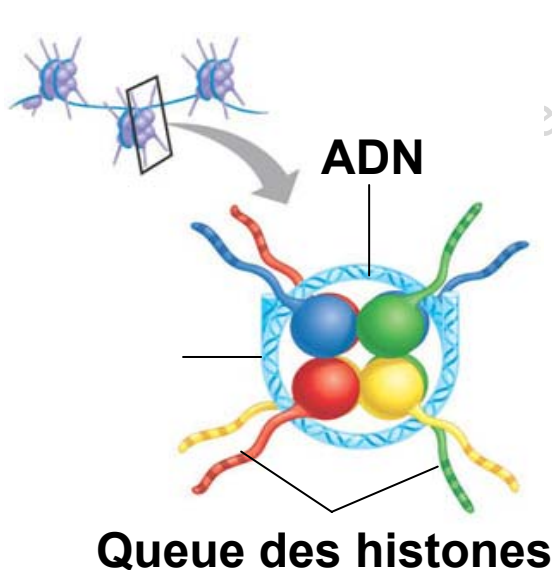


**Modifications
épigénétiques**

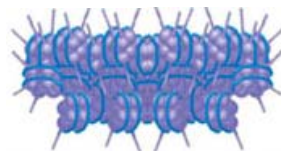


Modifications des histones

- Réversibles, nombreuses: « code » des histones
 - Acétylation (K), méthylation (K/R), phosphorylation...
 - Méthylation ou Acétylation sont mutuellement exclusives
 - Enzymes nombreuses, spécifiques d'un résidu donné
 - Histone acétyltransférases (HAT) ou déacétylases (HDAC)
 - Lysine méthyltransférases et déméthylases...
 - Concernent surtout l'extrémité N-terminale de H3 et H4
- Méthylation / Acétylation ont ~ un rôle opposé

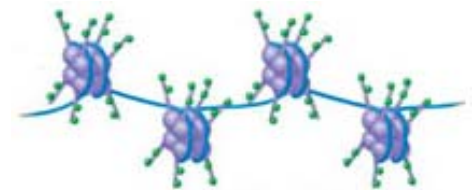


Chromatine fermée



- Déacétylation H3/H4
- Méthylation Lysines (Ex: H3K9)...

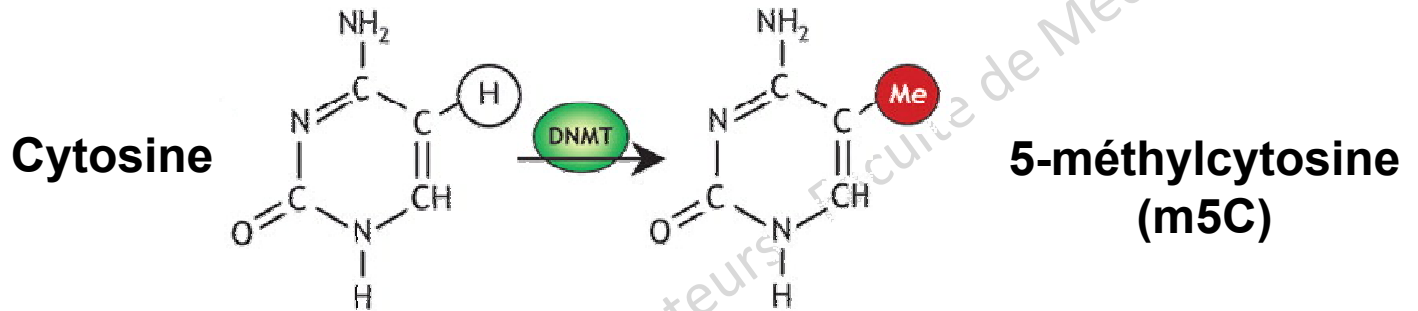
Chromatine ouverte



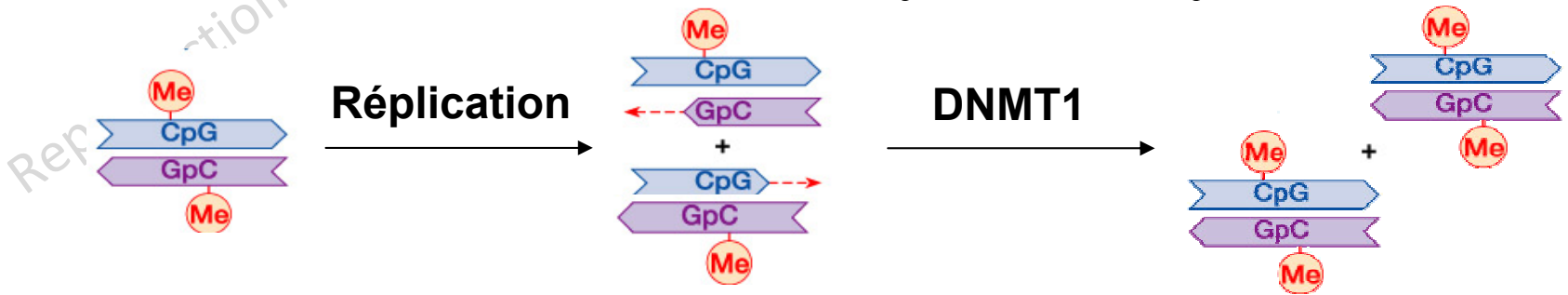
- Acétylation H3/H4
- Déméthylation Lysines (sauf H3K4)...

Méthylation de l'ADN

- **Survient sur les cytosines suivies d'une guanine**
 - Séquence CpG fréquente **promoteur** des gènes: **ilôts CpG**

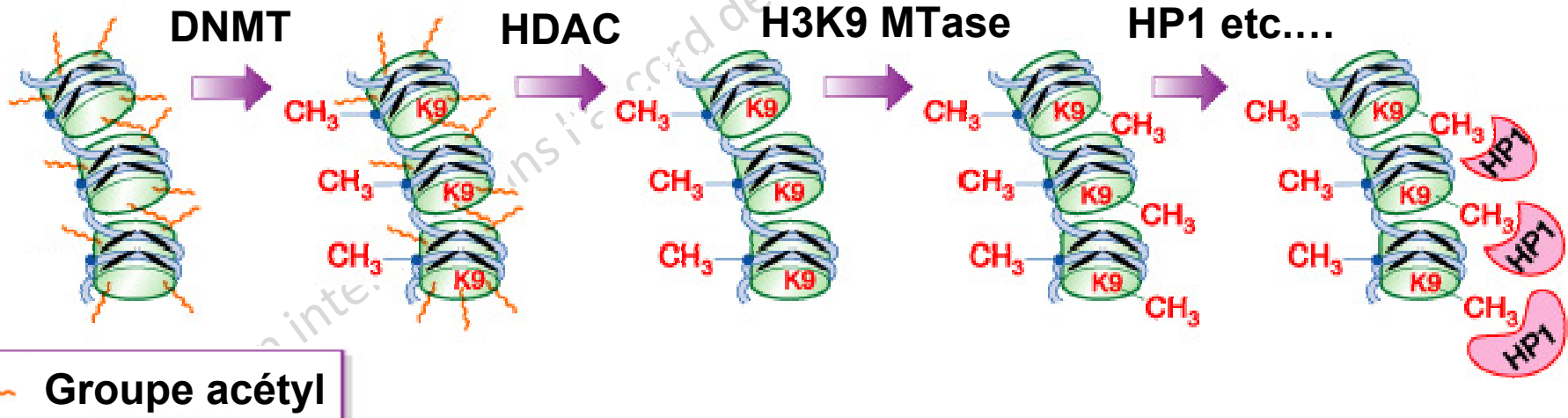


- **Fait intervenir des ADN méthyl-transférases**
 - Méthylation de maintenance (division cellulaire): DNMT1
 - Méthylation *de novo* (développement): DNMT3a, DNMT3b
- **Stable et transmissible lors de la division**
 - DNMT1: Fixe l'ADN hémi-méthylé et méthyle l'autre brin



Méthylation de l'ADN

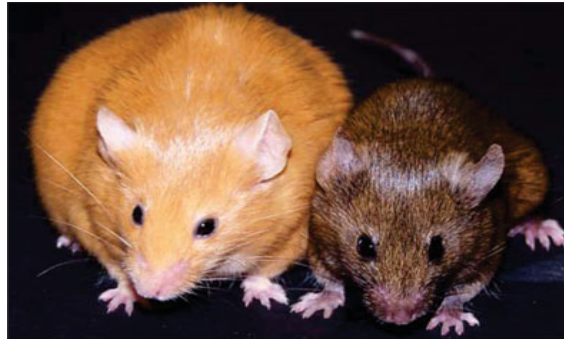
- → **Formation +/- irréversible d'hétérochromatine**
 - Initie une cascade aboutissant au **recrutement de HP1** (*Heterochromatin protein 1*)
 - Fixation de MeCP2 (*Methyl-CpG binding protein 2*)
 - Recrute HDAC et H3K9 MTase: déacétylation et **méthylation H3K9**
 - **Fixation HP1**, recrute HDAC et H3K9 MTase → **Diffusion signal**



- **Son maintien après réplication permet d'assurer la transmission des états chromatinien**

Importance de l'épigénétique

- **Développement et différenciation cellulaire**
 - Expression mono-allélique des gènes soumis à empreinte
 - Gènes dont une seule copie, paternelle ou maternelle, s'exprime
 - Inactivation de l'X par son hyperméthylation
 - Répression des gènes inutiles aux cellules différenciées...
- **Adaptation à l'environnement**
 - Signaux environnementaux modulent l'expression génique notamment par l'intermédiaire de l'épigénétique
- **Anomalies épigénétiques → Maladies**
 - Rares maladies liées à un défaut d'empreinte parentale
 - **Cancer**
 - **Obésité et diabète**
 - **Vieillessement...**



≠^{ces} entre Génome eucaryote/procaryote

- ≠^{ces} d'organisation à l'échelle chromosomique
 - ≠^{ces} d'organisation physique → ≠^{ces} d'expression génique
- ≠^{ces} d'organisation à l'échelle des gènes
 - Génomique comparative d'organismes séquencés

Procaryotes



H. Influenzae (1995)



Levure (1996)

Eucaryotes



C. Elegans (1998)



Homme (2003)

– Questions posées

- ≠^{ces} génétiques fondamentales entre procaryotes et eucaryotes ?
- Information génétique supp. pour coordonner un être multi ζ^R ?

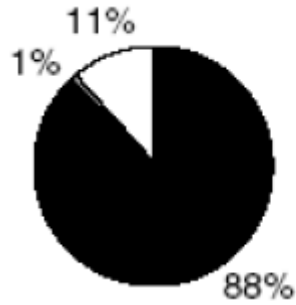
≠^{ces} entre Génome eucaryote/procaryote

- ≠^{ces} d'organisation à l'échelle des gènes

Eucaryotes

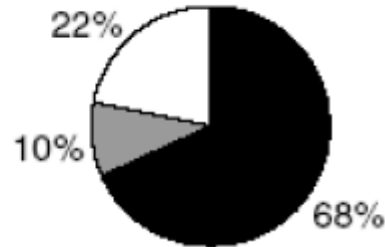
Procaryotes

E. Coli



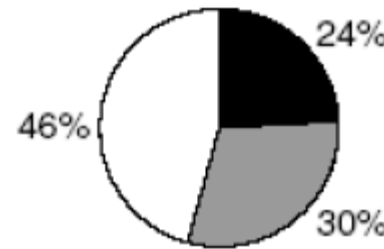
Unicellulaires

S. Cerevisiae

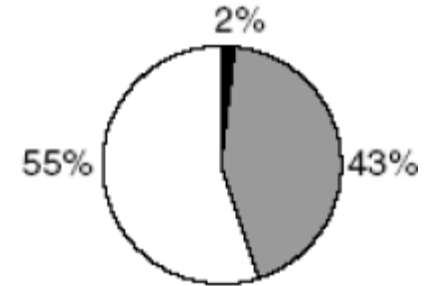


Pluricellulaires

C. Elegans



Homme



■ Régions codantes

■ Régions transcrits non codantes

□ Régions non transcrits

– ≠^{ces} de nombre de gènes ?

– ≠^{ces} proportion séquences codantes / taille du génome ?

≠^{ces} entre **Génome eucaryote/procaryote**

• **Nombre de gènes / taille des génomes**

Organisme	Taille Génome (10⁶ bases)	Nombre Gènes	Rapport
Procaryotes (<i>E. coli</i>)	4.6	~ 4000	~ 1000
Levure (<i>S. cerevisiae</i>)	12	~ 6,000	~ 2000
Ver (<i>C. elegans</i>)	100	~ 14,000	~ 7000
Homme	3 000	~ 30,000	~ 100 000

- Nombre de gènes ~ entre Procaryotes et Eucaryotes
- Rapport Gènes/Taille génome ↑ Eucaryotes multicellulaires

• ➔ ≠^{ces} **proportion génome codant / taille génome**

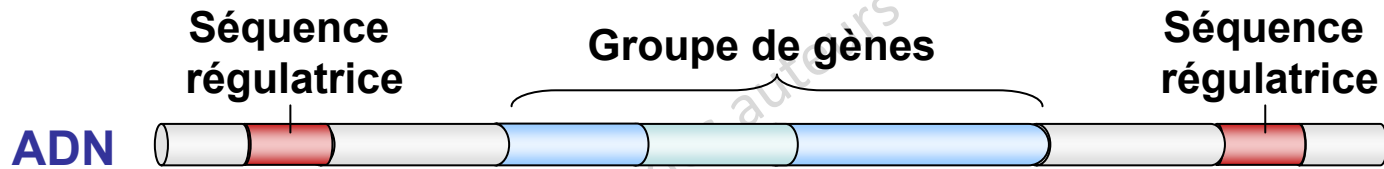
- **Procaryotes: majoritaire / Eucaryotes: minoritaire**
 - En proportion, plus de **séquences transcrites non codantes**
 - Et plus de **séquences non transcrites, intergéniques**
 - Procaryotes: Un gène toutes les 1000 bases
 - Homme: Un gène toutes les 100 000 bases

≠^{ces} entre **Génome eucaryote/procaryote**

- Les séquences transcrites non codantes reflètent les ≠^{ces} d'organisation des gènes

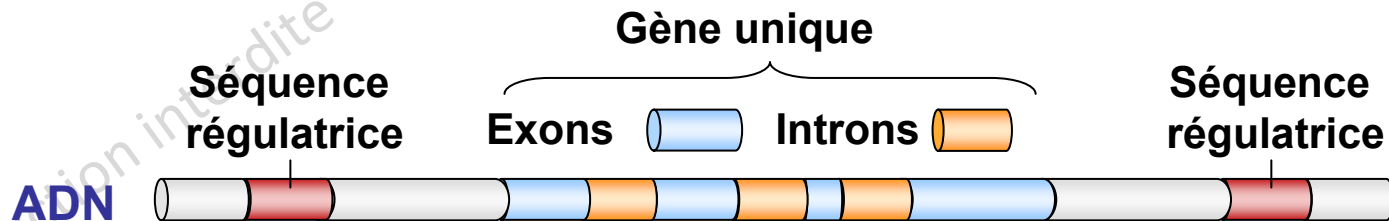
– Le génome procaryote:

- Gènes **compacts** (absence d'introns) et **regroupés** en unités de régulation (cf. modèle de l'opéron lactose)



– Le génome eucaryote:

- Gènes **morcelés** (présence d'introns), **régulés individuellement**



- Les introns sont transcrits (ARNm prémature) puis éliminés
- **Introns → Plus de séquences transcrites non codantes par gène**

≠^{ces} entre Génome eucaryote/procaryote

- Les séquences non transcrites reflètent les ≠^{ces} de taille des régions intergéniques

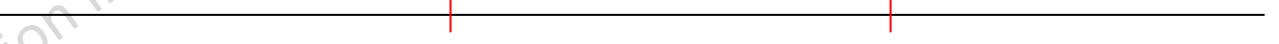
– Le génome procaryote:

- Peu de régions intergéniques
- Densité élevée de gènes (un gène toutes les 1000 bases)

ADN  200Kb

– Le génome eucaryote:

- Vastes régions intergéniques
- Densité faible des gènes (un gène toutes les 100 000 bases)
- Gène noyés dans un « désert »

ADN  200Kb

- Régions intergéniques = ADN « poubelle » ?

≠^{ces} entre **Génome eucaryote/procaryote**

- **Régions intergéniques = Séquences répétées (50% du génome humain)**
 - **Séquences répétées dispersées (45%) = Transposons**
 - Capables de se déplacer dans le génome = gènes « sauteurs »
 - Auraient favorisé l'évolution, plupart inactifs aujourd'hui
 - Découverts initialement chez le maïs (Barbara Mc Clintock, 1947) 🏆



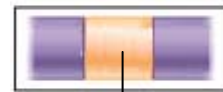
Gène normal



→
Pigment



Gène muté



→
✗

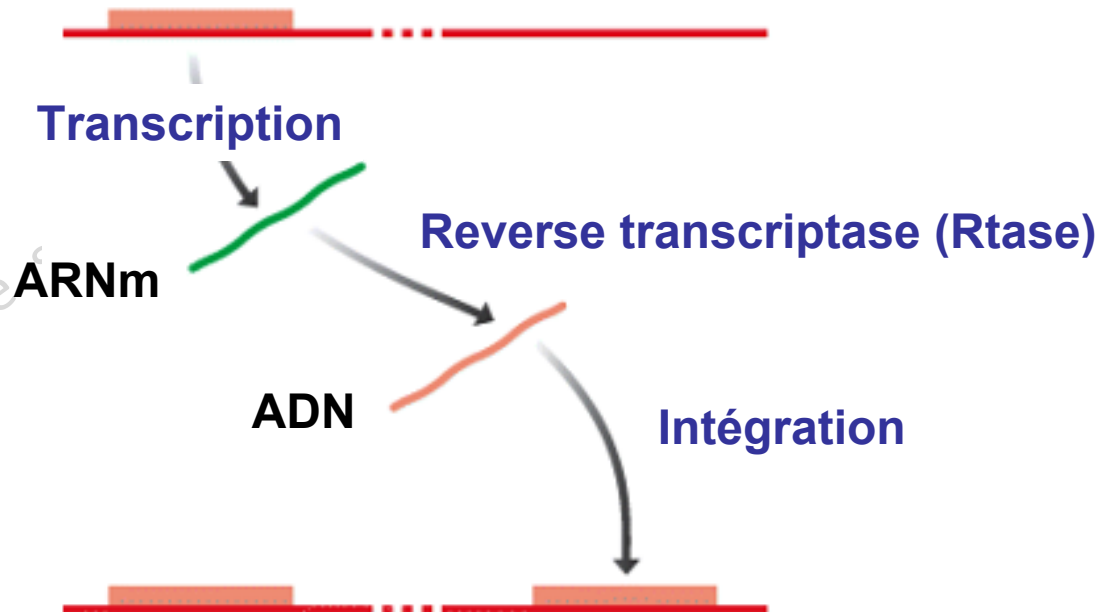


Transposon

- **Les transposons peuvent inactiver un gène au point d'insertion**

≠^{ces} entre **Génome eucaryote/procaryote**

- **Régions intergéniques = Séquences répétées**
 - **Séquences répétées dispersées (45%) = Transposons**
 - Transposons à ADN: Séquence d'ADN coupée et intégrée ailleurs
 - Rétrotransposons = Se déplacent par l'intermédiaire d'un ARN
 - L'ARN est rétro-transcrit en ADN avant l'intégration dans l'ADN cible (le même mécanisme est utilisé par les virus à ARN)
 - Aboutit à une multiplication du nombre de copies du transposon



- **Rétrotransposon autonome** = Code sa propre reverse transcriptase
- **Rétrotransposon non autonome** = Utilise celle d'un autre transposon

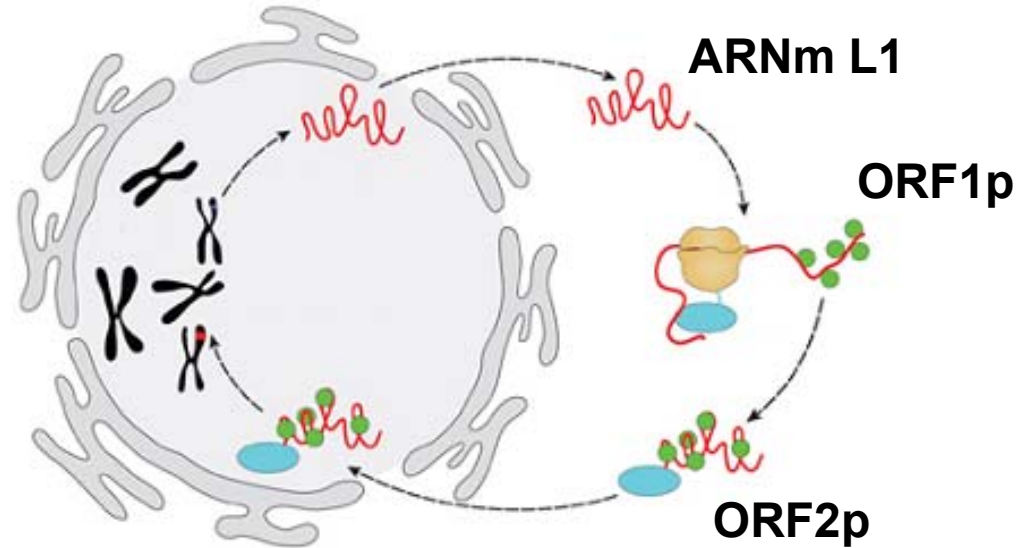
≠ ces entre **Génome eucaryote/procaryote**

- **Régions intergéniques = Séquences répétées**
 - **Séquences LINE** (*Long Interspersed Repeated Sequence*)
 - Rétrotransposons autonomes
 - 850 000 copies de 6-8 kb → soit 20% génome
 - Prototype, **séquences L1, seules séquences encore actives**



ORF1p: Se lie à l'ARNm

ORF2p: Rtase et Endonucléase



- **Séquences SINE** (*Short Interspersed Repeated Sequence*)
 - Rétrotransposons non autonomes
 - 1 500 000 copies de 100-300 b → 15 % génome
 - Prototype, **séquences Alu, abondantes dans introns des gènes**

≠^{ces} entre **Génome eucaryote/procaryote**

- **Régions intergéniques = Séquences répétées**
 - **Séquences répétées en tandem: (5%)**
 - Classées par la taille totale de la séquence, motif répété ~ 1-100pb
 - **Séquences satellites** (100 000-10 millions pb):
 - Retrouvées au niveau de chaque **centromère** (rôle structural)
 - **Minisatellites** (100-100 000 pb), deux types:
 - Séquences polymorphes (nombre de répétitions variant entre individus), appelées VNTR (*Variable Number of Tandem Repeats*)
 - » Considérées comme **points chauds de recombinaison**
 - Séquences des **télomères**: motif TTAGGG répété sur 3-20 kb
 - » Protègent l'extrémité des chromosomes
 - **Microsatellites** ou STR (*Simple Tandem Repeats*) (10-100pb)
 - » Répétition de di-, tri- ou tétra-nucléotides
 - » Erreurs de réplication fréquentes → **points chauds de mutation**
 - » Instabilité → Maladies (Ex: maladies par expansion de triplets)

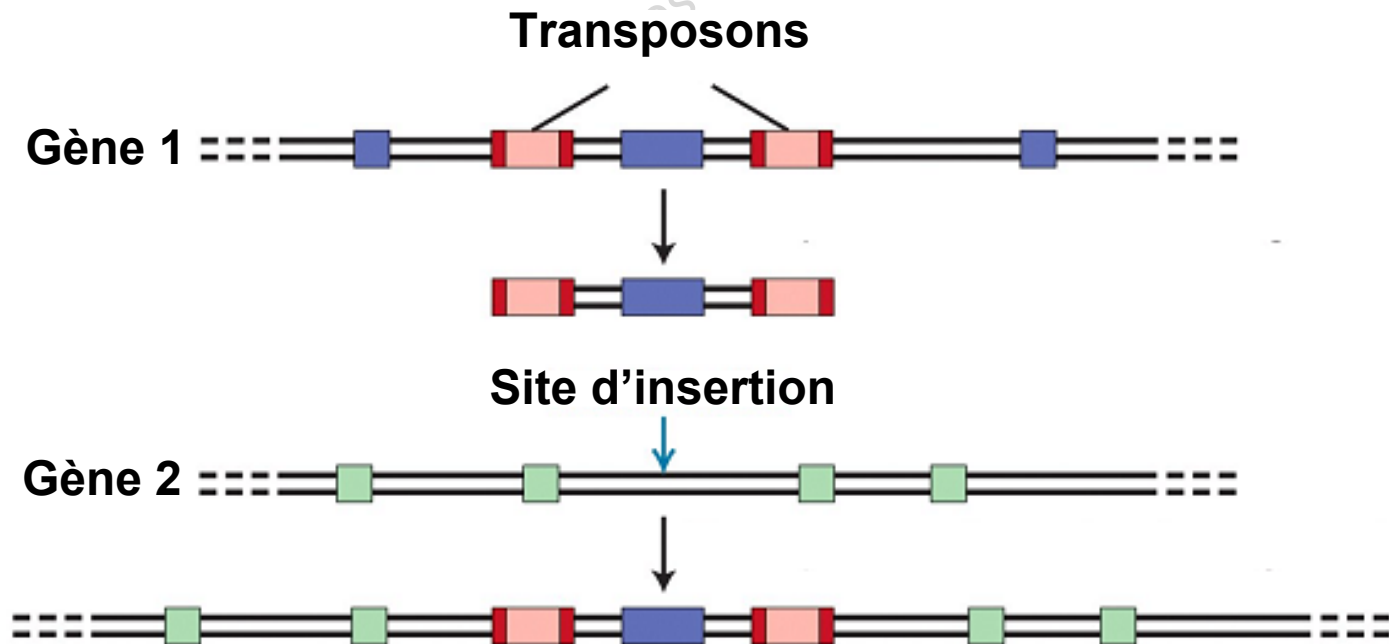
≠^{ces} entre Génome eucaryote/procaryote

- Rôle des régions intergéniques = Évolution

→ Création de nouveaux gènes / brassage d'exons
(*Exon Shuffling*)

– Par transposition d'exons d'un gène à un autre

- Un exon peut-être mobilisé par les transposons qui l'encadrent



Reprod'

≠^{ces} entre Génome eucaryote/procaryote

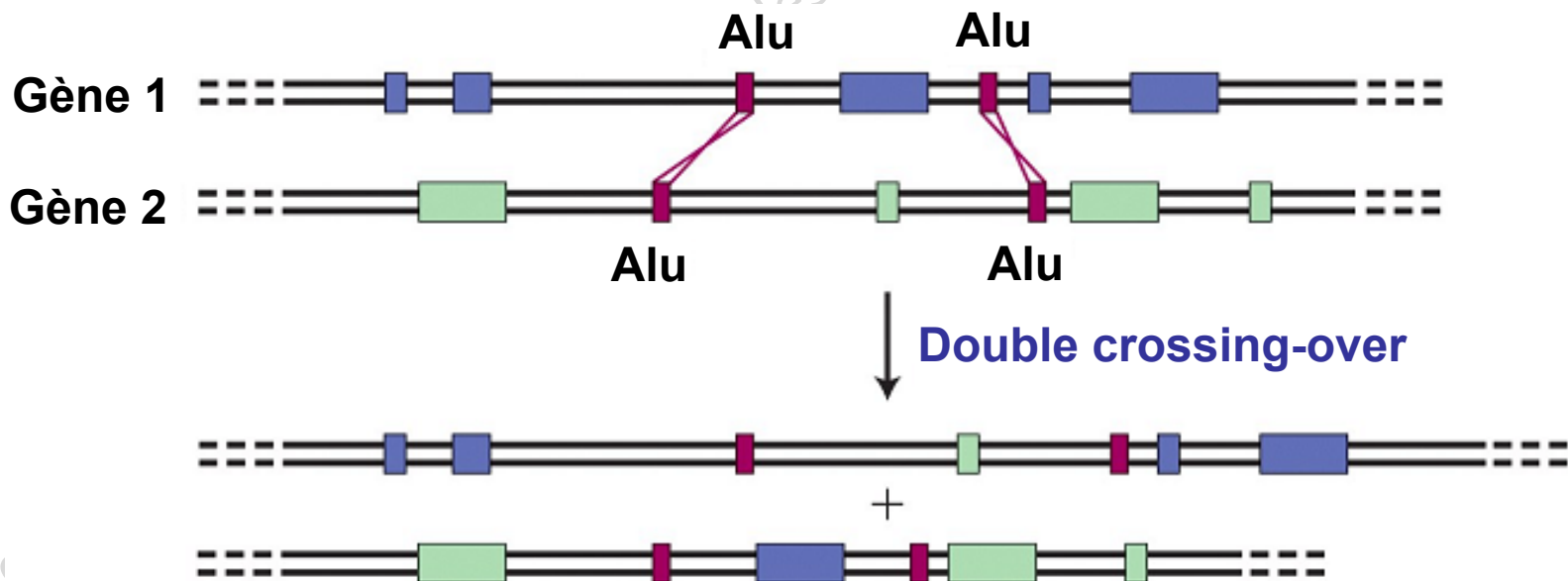
- Rôle des régions intergéniques = Évolution

→ Création de nouveaux gènes / brassage d'exons

(*Exon Shuffling*)

– Par recombinaison entre séquences répétées

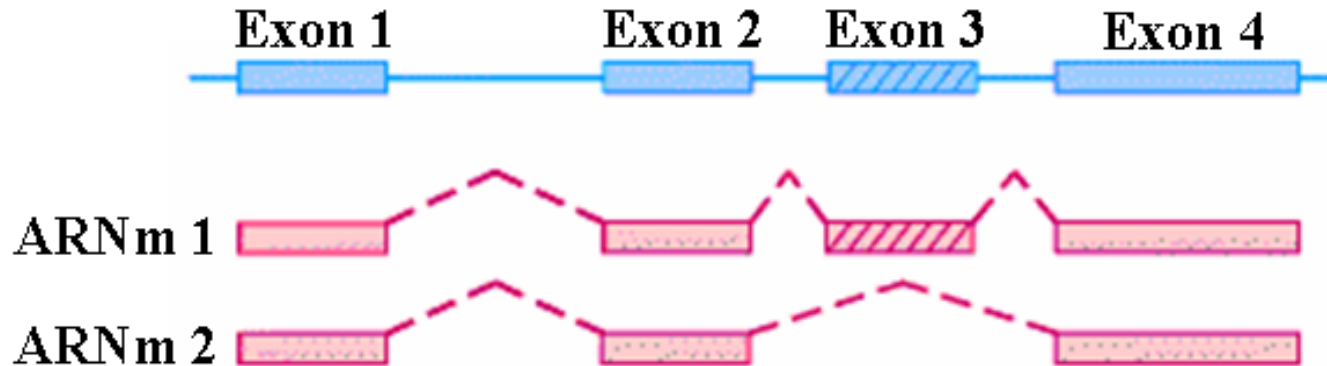
- Séquences Alu contenues dans les introns de gènes ≠



≠^{ces} entre Génome eucaryote/procaryote

• Rôle actuel des Introns

- **Augmenter le nombre de protéines** issues d'un seul gène
 - Choix des exons inclus dans l'ARNm peut varier (épissage alternatif)



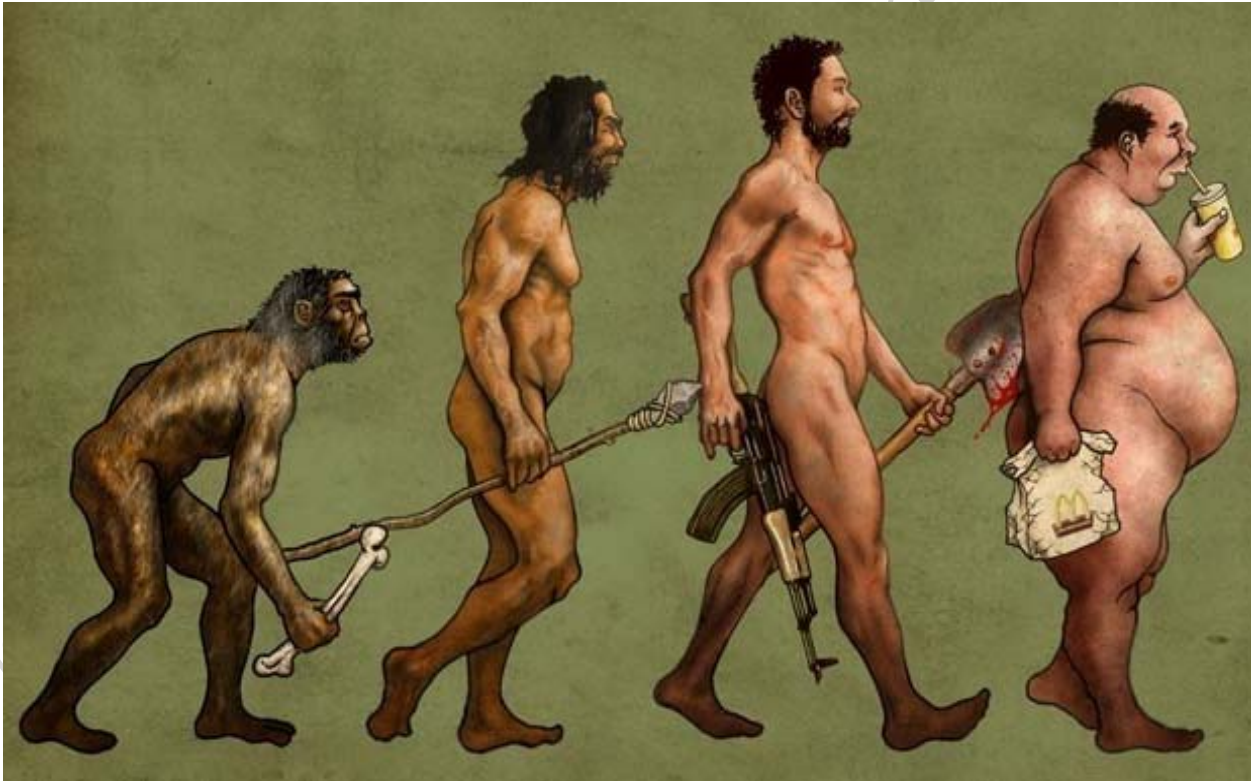
- Un exon code souvent pour un domaine protéique
- Un gène → Protéines ≠ par combinaison de ≠ domaines
- **Chez l'homme, 30 000 gènes → 200 000 protéines ≠**
 - C'est nombre de protéines ≠ qui reflèterait la complexité des eucaryotes, et notamment de l'homme

≠^{ces} entre Génome eucaryote/procaryote

- Séquences répétées intergéniques

- Sont à la base de l'évolution des espèces

- Grâce aux transpositions, aux recombinaisons ou aux mutations



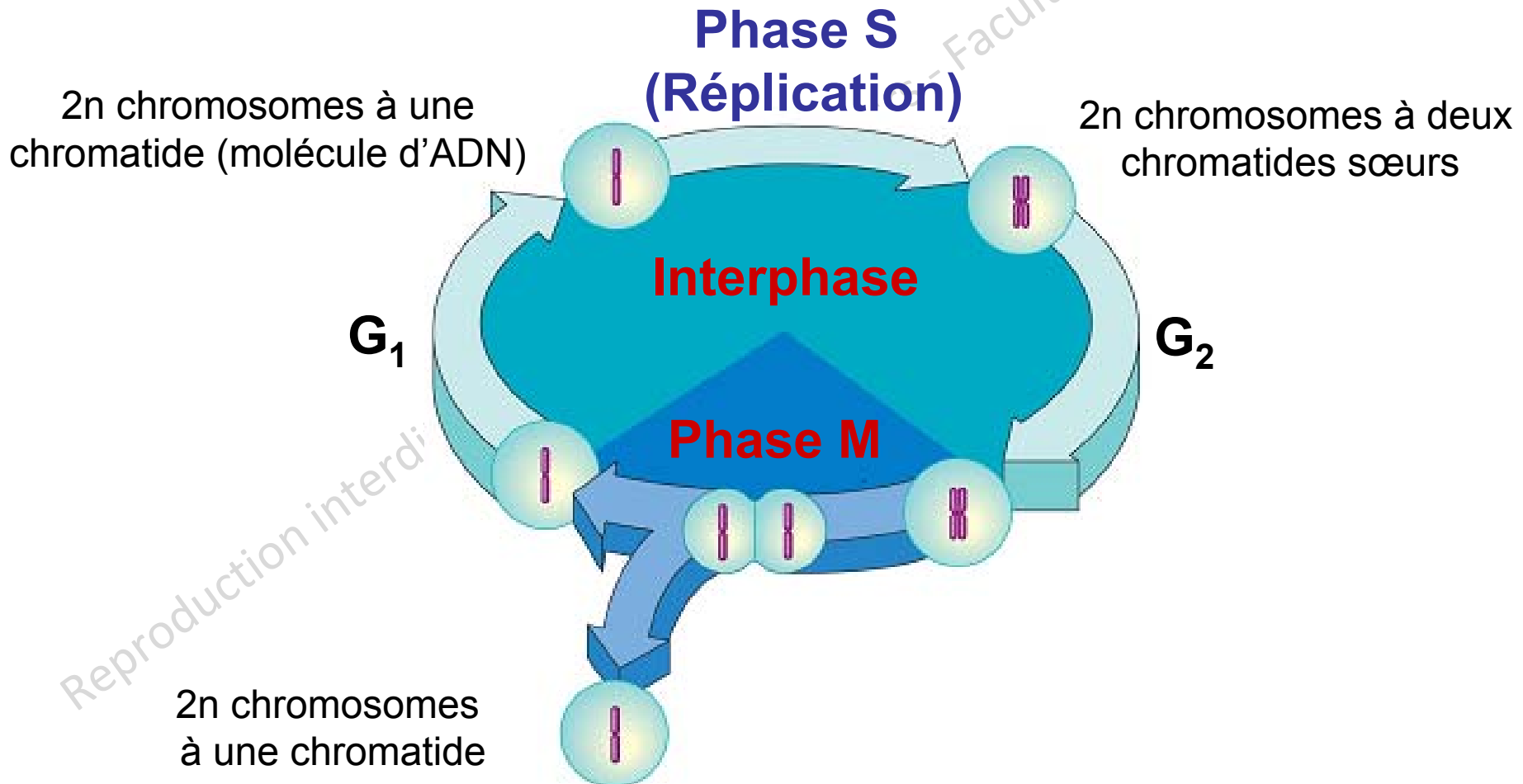
- Peuvent favoriser l'apparition de maladies génétiques

Réplication du génome eucaryote

- **Caractéristiques générales de la réplication**
- **Mécanismes biochimiques de la réplication**
 - L'initiation de la réplication
 - L'élongation de la réplication
 - La terminaison de la réplication
- **La fidélité de la réplication**
 - Mécanismes assurant la fidélité de la réplication
 - Accumulation de Mutations et Cancer

Caractéristiques de la réplication

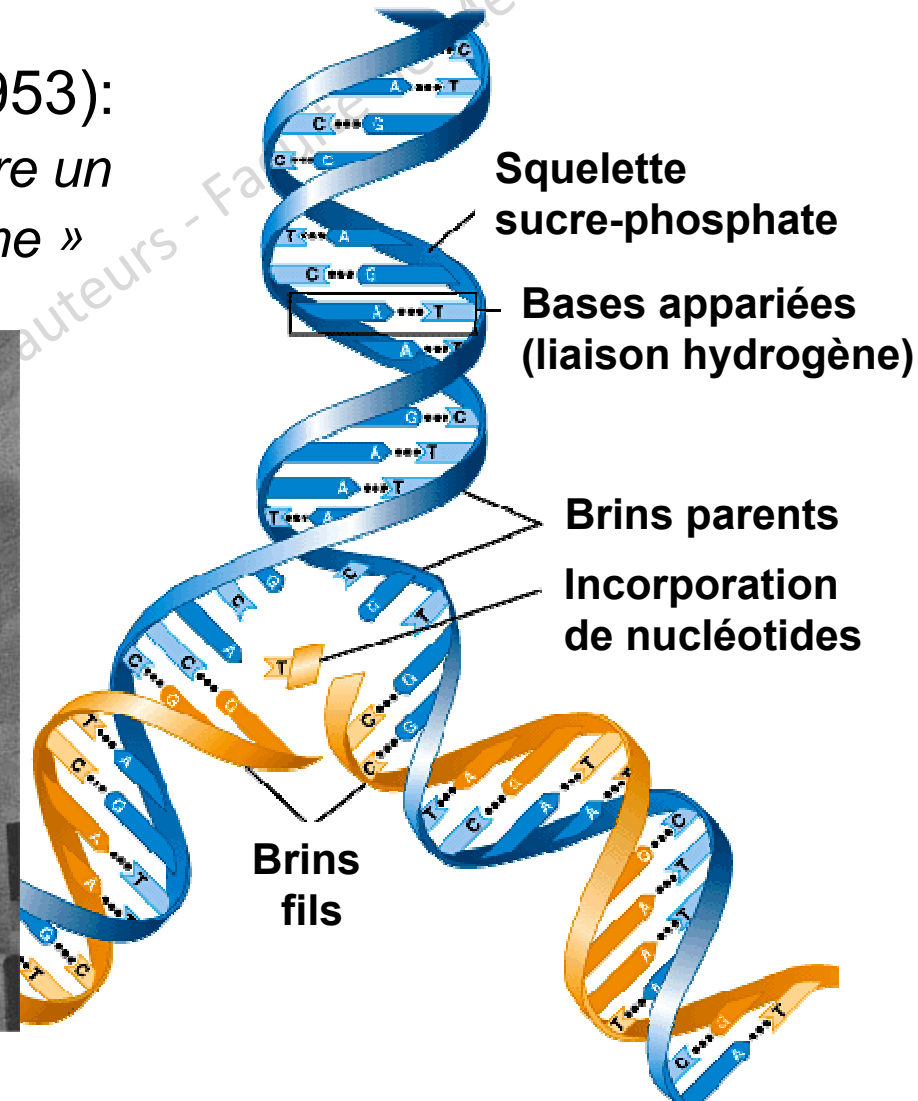
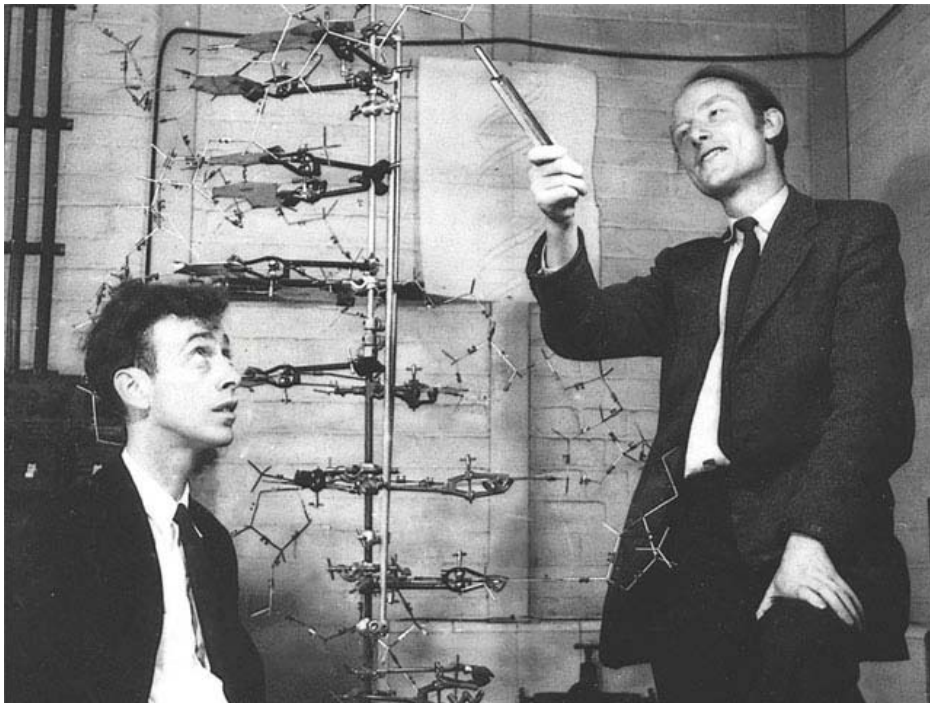
- Elle est couplée au cycle cellulaire
 - Une cellule qui se divise doit répliquer son ADN
 - Chaque cellule fille hérite d'une copie d'ADN identique



Caractéristiques de la réplication

- Elle doit être parfaitement fidèle
 - Cette fidélité repose sur la complémentarité des bases

James Watson et Francis Crick (1953):
« *La complémentarité des bases suggère un possible mécanisme de copie du génome* »



Caractéristiques de la réplication

• Modèles possibles de réplication

– Modèle semi-conservatif (Watson, Crick) :

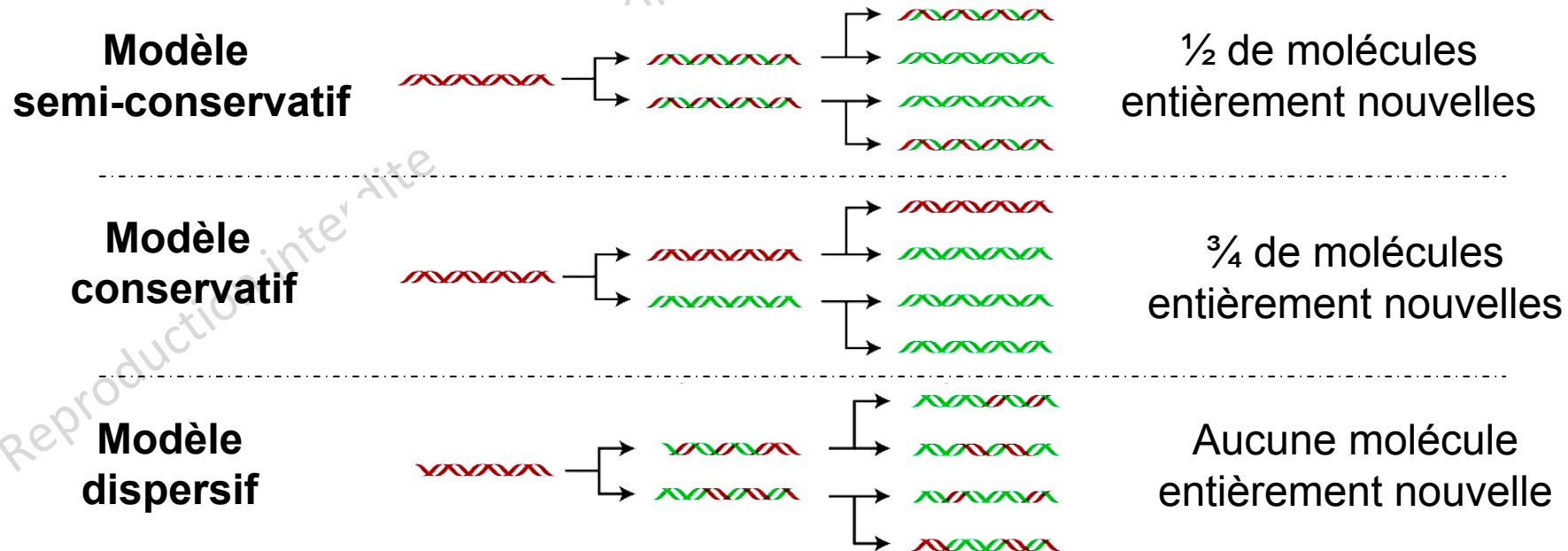
- Chaque molécule contient un brin parent et un nouveau brin

– Modèle conservatif:

- Une molécule intacte, une molécule contenant deux nouveaux brins

– Modèle dispersif:

- Chaque molécule = mélange de fragments originaux et néosynthétisés

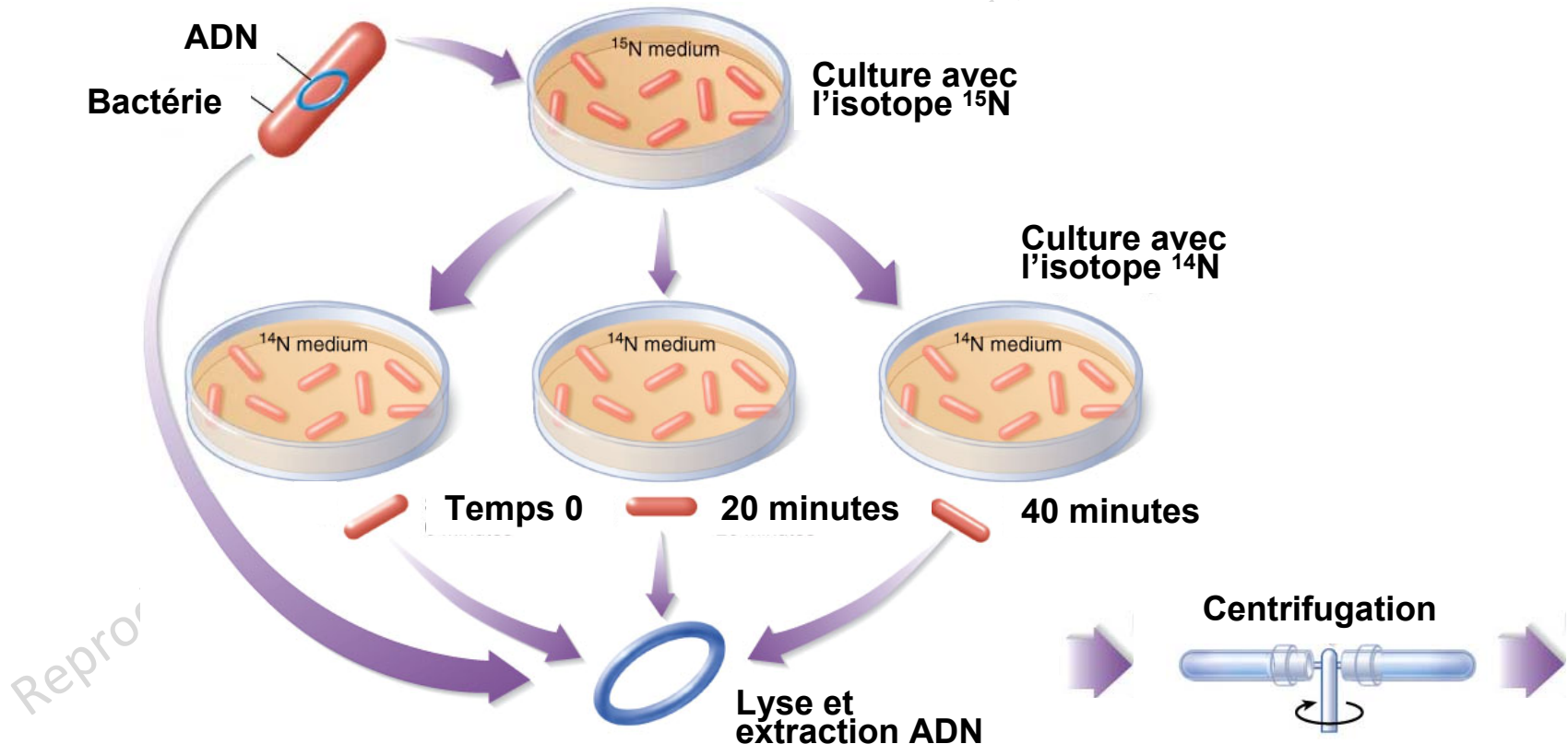


Caractéristiques de la réplication

• Expérience de Meselson et Stahl (1958)

– Afin de permettre la distinction entre brins parents et fils:

- Bactéries cultivées en présence de l'isotope lourd ^{15}N puis léger ^{14}N
- Extraction d'ADN à chaque division (20mn) et analyse de sa densité



Caractéristiques de la réplication

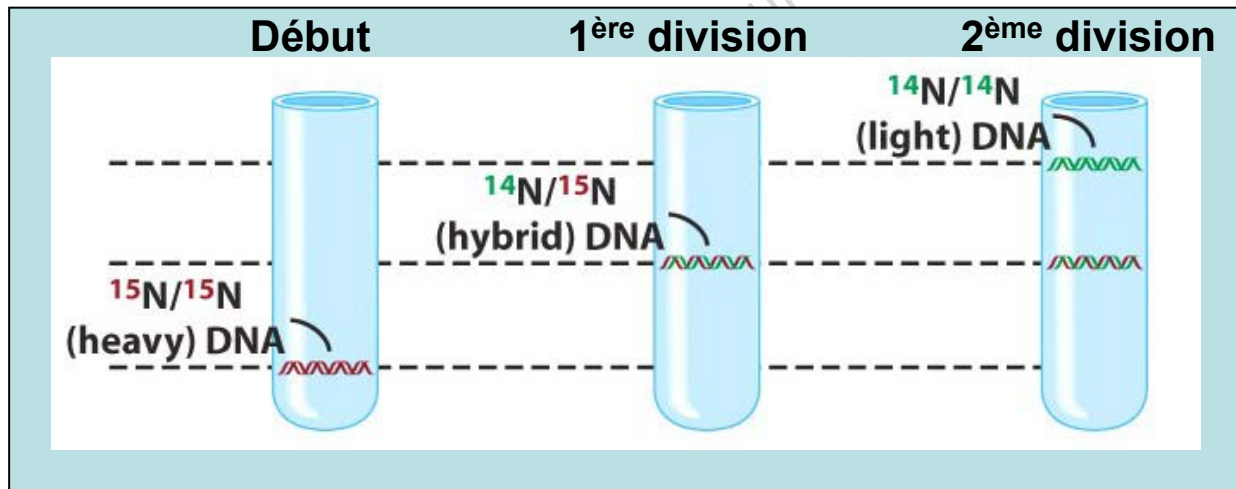
- Les résultats prouvent le modèle semi-conservatif:

- 1^{ère} division

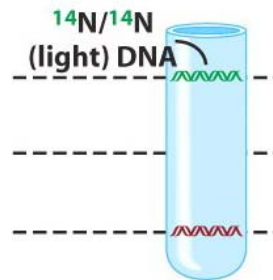
- Bande intermédiaire = molécules constituées pour moitié de ^{14}N et ^{15}N

- 2^{ème} division

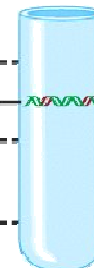
- Nouvelle bande = molécules constituées uniquement de ^{14}N



≠ Modèle conservatif



$^{14}\text{N}/^{15}\text{N}$ (hybrid) DNA



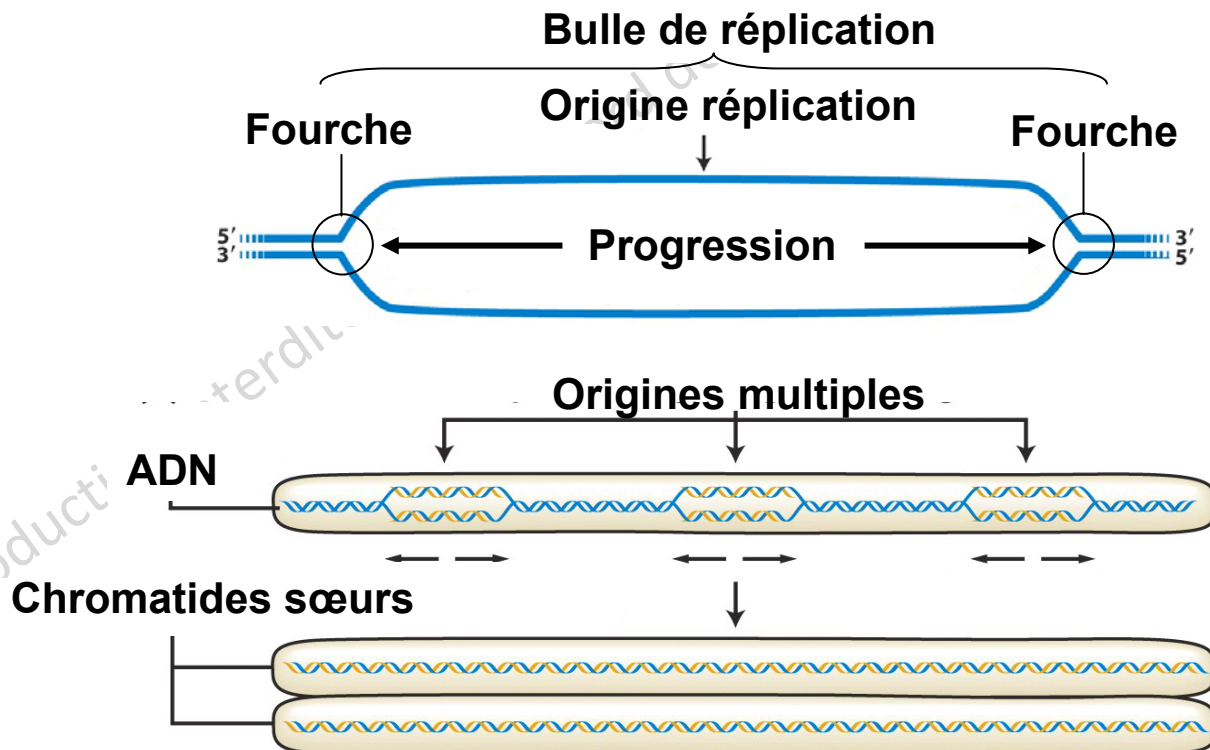
≠ Modèle dispersif

Étapes de la réplication

• Étape d'initiation:

– Formation des bulles de réplication

- La réplication initiée en de multiples sites: **Origines de réplication**
- Des hélicases séparent les brins parents: **Fourches de réplication**
- Elles progressent dans les deux sens en déplaçant les nucléosomes
- Les brins séparés sont stabilisés par RPA (*Replication Protein A*)

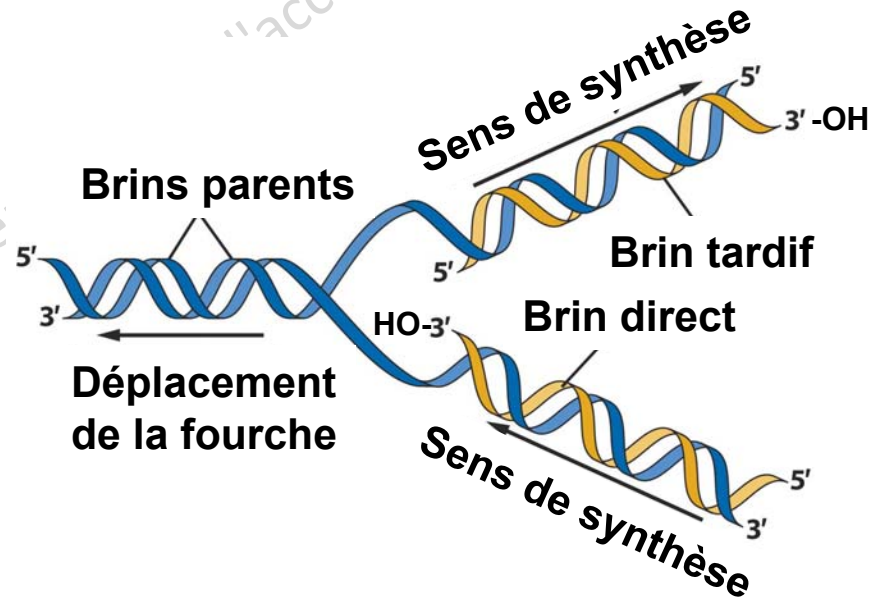


Étapes de la réplication

- **Étape d'initiation:**

- **Synthèse des amorces fournissant les extrémités 3'-OH**

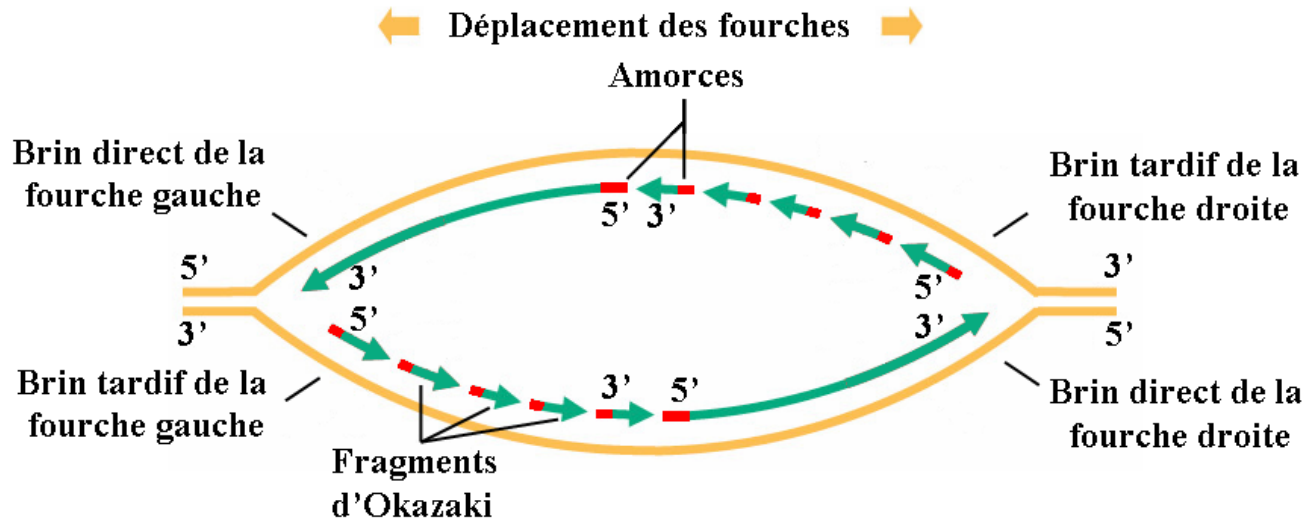
- Hybrides ARN/ADN, complémentaires des brins parents
- Synthétisées dans le sens 5'-3' par l'**ADN polymérase α** (primase)
- La réplication se fait dans le sens 5'-3' donc, au niveau d'une fourche
 - Un brin sera synthétisé dans le sens de la fourche: **brin direct**
 - L'autre sera synthétisé en sens opposé: **brin tardif**



Étapes de la réplication

• Étape d'élongation:

- Les dNTPs (dATP, dGTP, dCTP et dTTP) correctement appariés sont reliés un à un à l'extrémité 3'-OH des amorces
 - → liaisons 3'-5' phosphodiester par les **ADN polymérase** δ et ϵ
 - Rôles respectifs inconnus, forment un dimère à chaque fourche
 - Synthèse du brin direct continue à partir d'une amorce unique
 - Synthèse du brin tardif par fragments discontinus (Okazaki, 200pb)
 - Amorces produites au fur et à mesure de l'avancée de la fourche

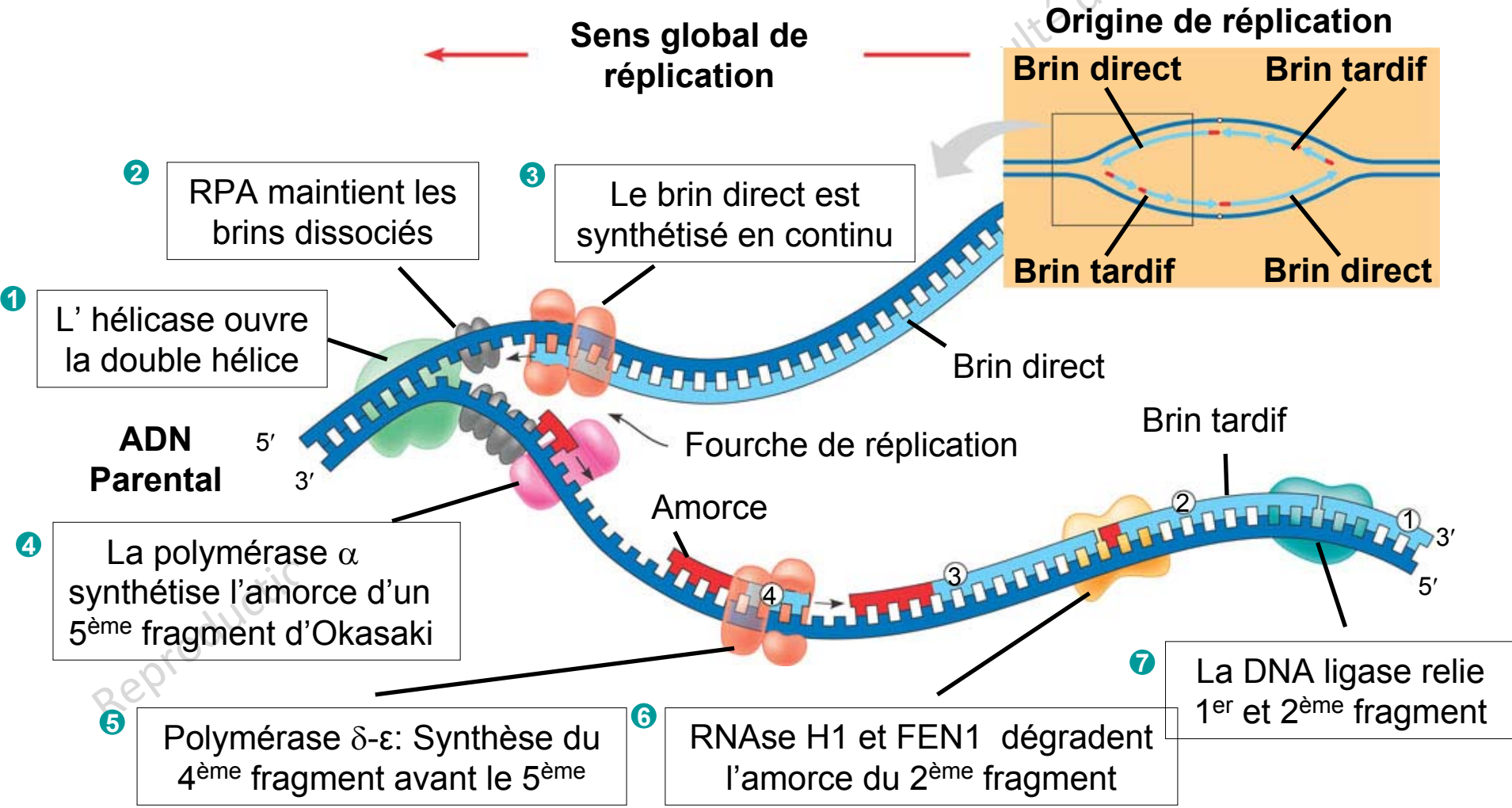


- Brin direct d'une fourche = Brin tardif de l'autre fourche, et inversement

Étapes de la réplication

Étape d'élongation:

- Amorces dégradées (RNase H₁ et FEN₁), brèches comblées (polymérase δ/ϵ) et les fragments reliés (ADN ligase)

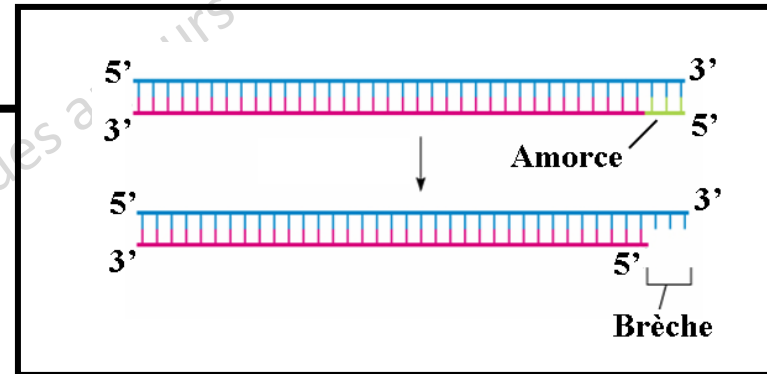


Étapes de la réplication

- **Étape de terminaison:**

- **La réplication des télomères pose un problème**

- **A l'extrémité 5' du brin fils de chaque chromatide, la dégradation de l'amorce ARN/ADN la plus distale laisse persister une brèche**



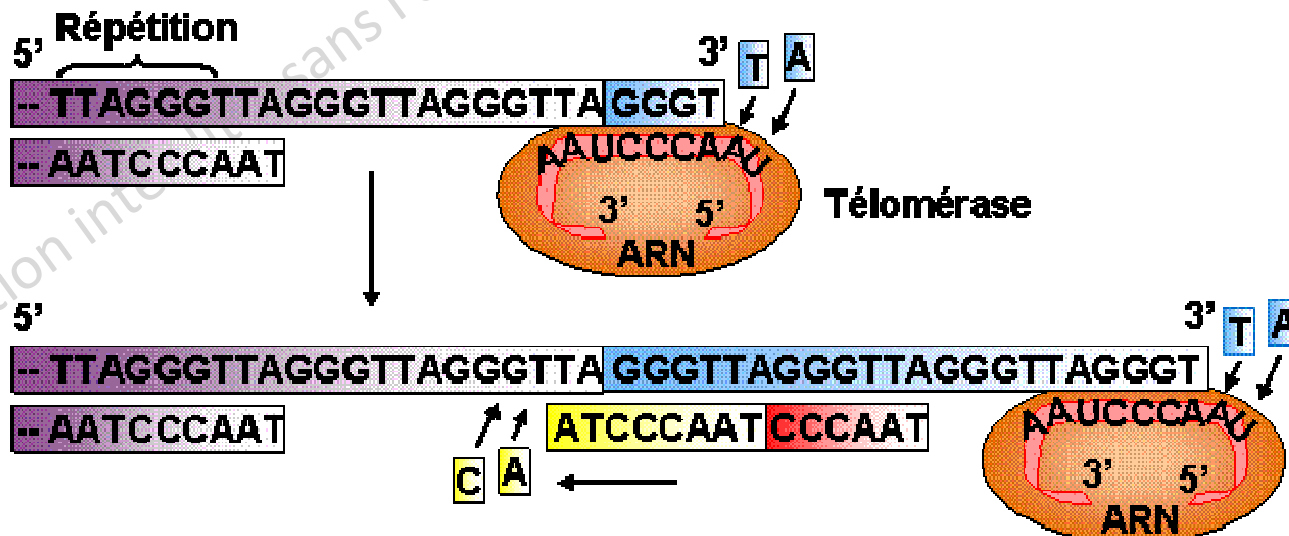
- **Si elle n'est pas comblée,**

- → **raccourcissement progressif des télomères à chaque division**
- Lorsque ce raccourcissement atteint un **seuil critique**, la cellule arrête de se diviser et devient sénescence
 - » **Le nombre de divisions** qu'une cellule effectue est **limité** (limite de Hayflick) et est **lié à la longueur de ses télomères**

Étapes de la réplication

- **Étape de terminaison:**

- Elle nécessite une enzyme appelée **téломérase**
 - Possède une activité de type **reverse transcriptase**
 - Lui permet la synthèse d'ADN à partir d'ARN
 - Et un **ARN matrice** complémentaire des répétitions télomériques
- La télomérase **va s'apparier au brin parent et l'allonger**
 - Une amorce plus distale pourra s'apparier au brin parent allongé
 - La brèche du brin fils sera comblée (polymérase δ/ϵ et ligase)

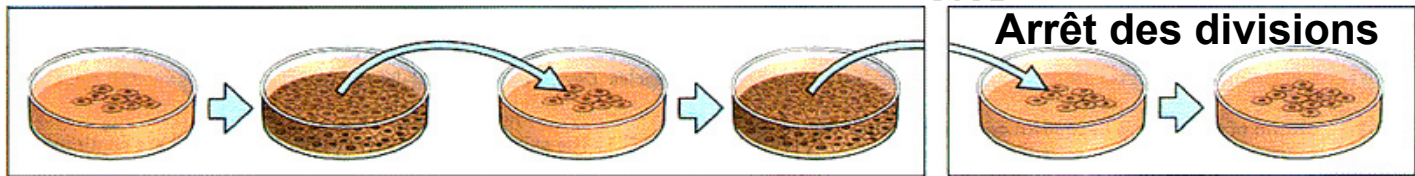


Étapes de la réplication

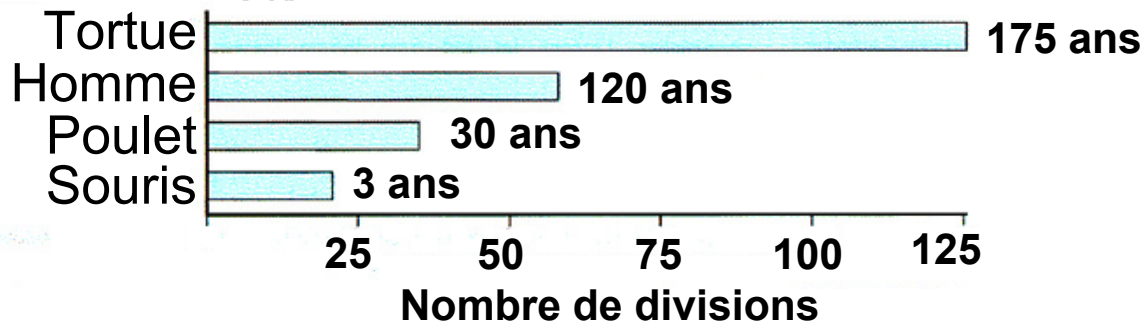
• Étape de terminaison:

– Elle nécessite une enzyme appelée **téломérase**

- La plupart des cellules de l'organisme n'en possède pas et leurs télomères se raccourcissent jusqu'à arrêt des divisions



- Les cellules mortes ne sont plus remplacées, et l'organisme vieillit
- La durée de vie d'un organisme serait liée à la longueur des télomères
 - » Corrélation entre longévité et nombre de divisions max. des cellules

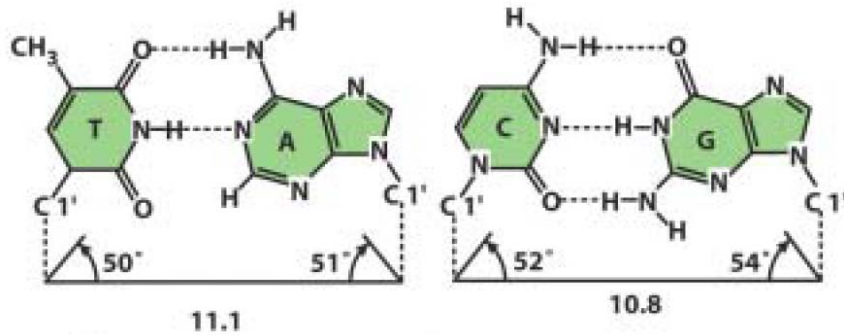


- Les cellules souches, germinales ou cancéreuses ont une activité télomérase et se divisent à l'infini (~ cellules immortelles)

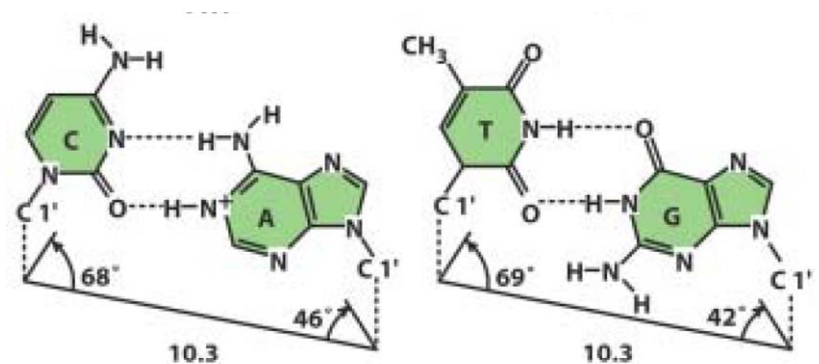
Fidélité de la réplication

- **Trois mécanismes agissant séquentiellement:**
 - **La sélection des bases complémentaires de la matrice**
 - Assurée par le **site actif des ADN polymérases α et δ/ϵ**
 - Il ne fonctionne que lorsque la **géométrie de la paire de base** formée est **conforme au principe de complémentarité**
 - **L'adénine avec la thymine, la guanine avec la cytosine**
 - Sinon, les angles et la géométrie des bases sont modifiés et le site actif de l'enzyme ne fonctionne pas

Association purine-pyrimidine correcte



Association purine-pyrimidine incorrecte



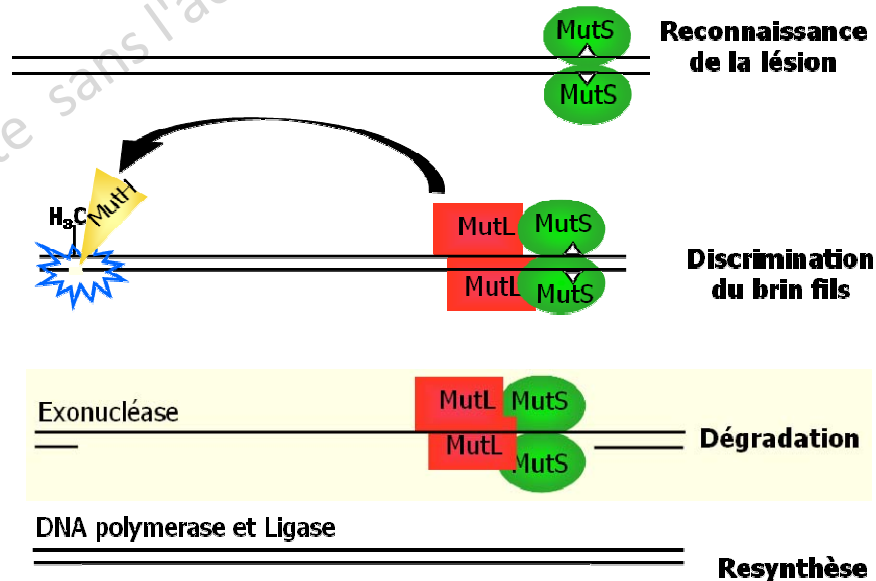
REV

Fidélité de la réplication

- **Trois mécanismes agissant séquentiellement:**
 - **L'activité de correction d'épreuve (*proofreading*)**
 - Uniquement assurée par **l'ADN polymérase δ/ϵ**
 - Et par la polymérase répliquant l'ADN mitochondrial (Pol γ)
 - La polymérase α ne possède pas ces activités: les amorces synthétisées seront dégradées et peuvent contenir des erreurs
 - La correction d'épreuve est liée à **l'activité 3'-5' exonucléasique**
 - Lui permet de «revenir sur ses pas» pour exciser un nucléotide erroné
 - Celui-ci est ensuite remplacé selon la complémentarité des bases

Fidélité de la réplication

- **Trois mécanismes agissant séquentiellement:**
 - **Le système MMR** (*Mutation Mismatch Repair*)
 - Système de détection et réparation des mésappariements
 - Assure la **correction des erreurs échappant à la polymérase**
 - **Procaryotes**, constitué de **trois protéines** formant des homodimères
 - **MutS** reconnaît la lésion et recrute **MutL** qui active **MutH**
 - » MutH est une **endonucléase** qui clive spécifiquement le **brin fils**
 - » L'ADN bactérien est méthylé mais le brin fils ne l'a pas encore été
 - Le brin est dégradé jusqu'à la lésion (exonucléase) puis resynthétisé



Fidélité de la réplication

- **Trois mécanismes agissant séquentiellement:**
 - **Le système MMR** (*Mutation Mismatch Repair*)
 - **Chez les Eucaryotes**, le système s'est diversifié
 - On connaît 6 homologues de MutS (MSH1 à MSH6), 4 homologues de MutL (MLH1, MLH3, PMS1, PMS2) mais pas d'homologue de MutH
 - » MSH2-MSH6 reconnaît substitutions et insertions / délétions 1-2nt
 - » MSH2-MSH3 reconnaît insertions / délétions de taille supérieure
 - Reconnaissance du brin fils liée à la présence des fragments d'Okazaki
- **Grâce a ces trois mécanismes séquentiels,**
 - **La réplication du génome ($6 \cdot 10^9$ nt) se fait ~ sans erreur**
 - **Sélection bases:** Laisse échapper 1 erreur tous les 10^5 nt
 - **Proofreading:** En corrige 99%. Il reste 1 erreur tous les 10^7 nt
 - **Système MMR:** En corrige 99,9%. Il reste 1 erreur tous les 10^{10} nt

Fidélité de la réplication

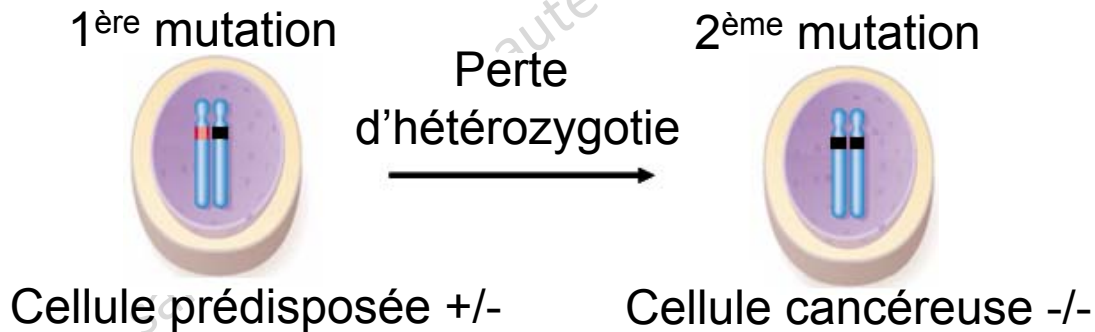
- L'inactivation du système MMR

- → Accumulation de mutations et apparition de cancers

- Gènes système MMR = Gènes suppresseur de tumeurs

- Ce type de gènes fonctionne sur un mode récessif

- Les deux copies doivent être inactivées pour l'apparition des KC



- La 1^{ère} mutation peut-être héritée (*MSH2* ou *MLH1* surtout)

- → Prédisposition au **syndrome de Lynch** ou **syndrome HNPCC**

(*Hereditary non polyposis colon cancer*): colon, estomac, ovaires...

- Toutes les cellules de l'individu sont hétérozygotes

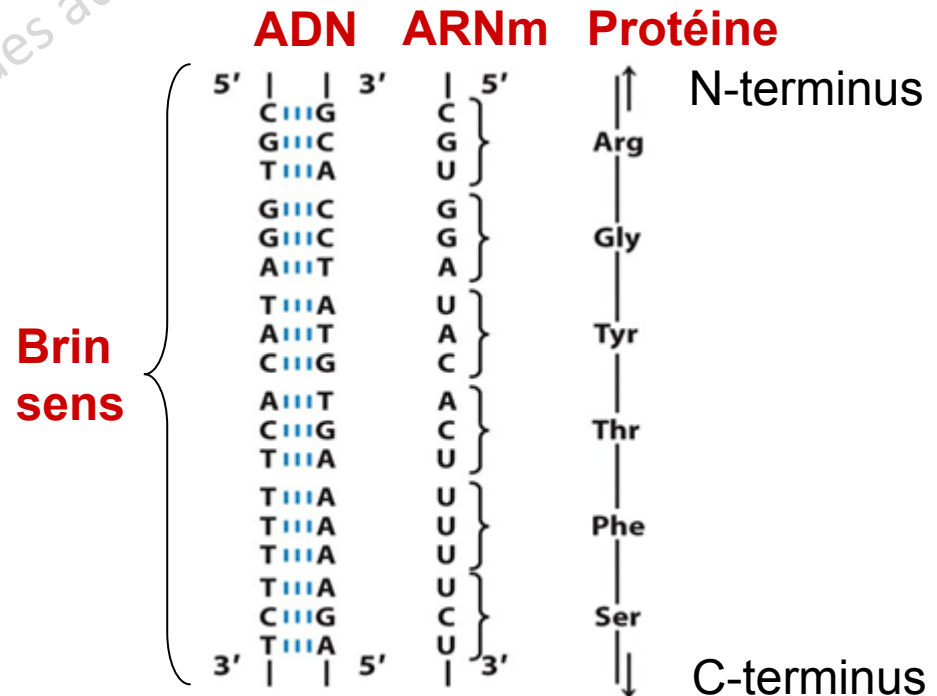
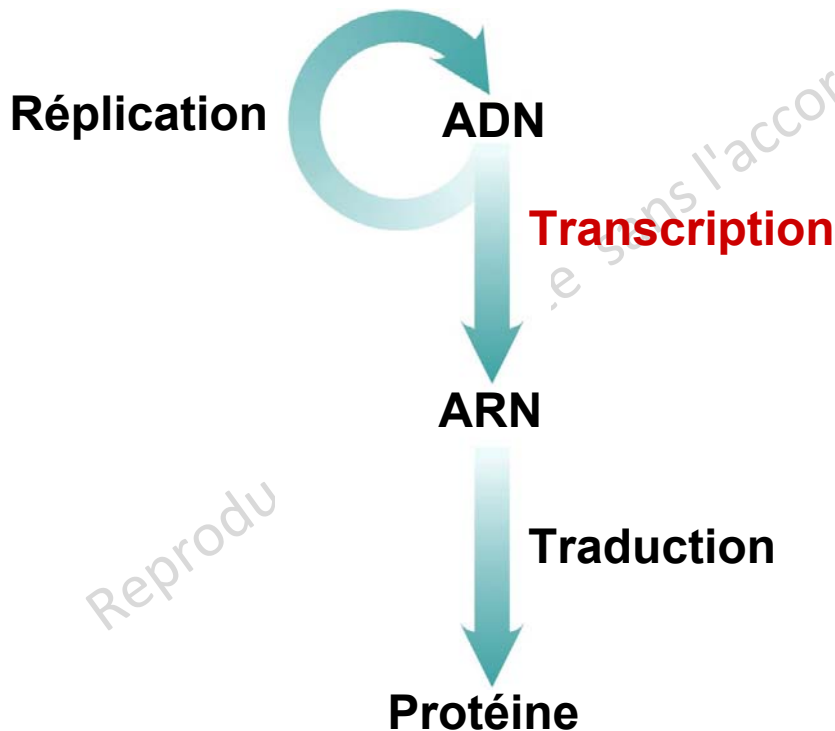
- La 2^{ème} mutation, constante, ne survient que dans certains tissus

Transcription des gènes eucaryotes

- **Vue générale de la transcription**
- **Transcription des gènes codant les protéines**
 - **Structure des gènes codant les protéines**
 - **Les éléments d'un promoteur**
 - **La machinerie basale de transcription**
 - **Initiation, élongation et terminaison de la transcription**
 - **La maturation des transcrits primaires**

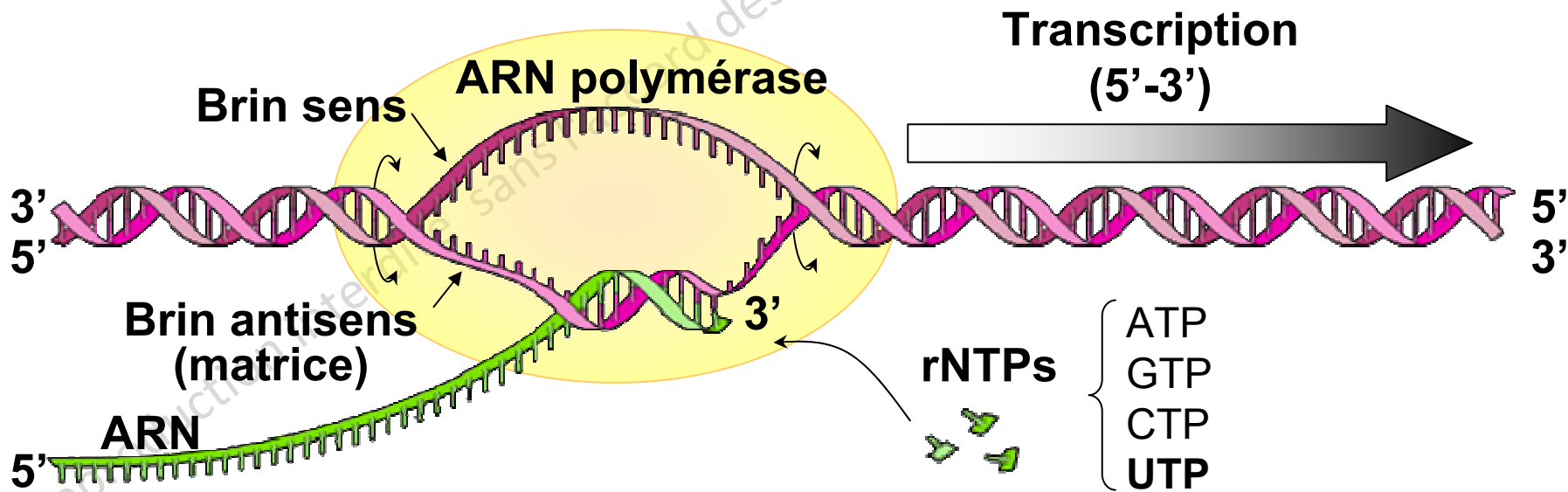
Vue générale de la transcription

- **Les gènes = Unités fonctionnelles du génome**
 - Au niveau d'un gène, le **brin sens** est celui portant l'information, le **brin antisens** son brin complémentaire
- **Transcription = Synthèse d'ARN ~ au brin sens**
 - Utilise le principe de **complémentarité des bases** (F. Crick)



Vue générale de la transcription

- **Transcription = Synthèse d'ARN ~ au brin sens**
 - **Une ARN polymérase**
 - Utilise le **brin antisens** comme matrice, lu dans le sens **3'-5'**
 - **Fonctionne dans le sens 5'-3'**
 - Relie les rNTPs qui y sont appariés (→ liaisons 3'-5' phosphodiester)
 - Ne nécessite pas d'amorce



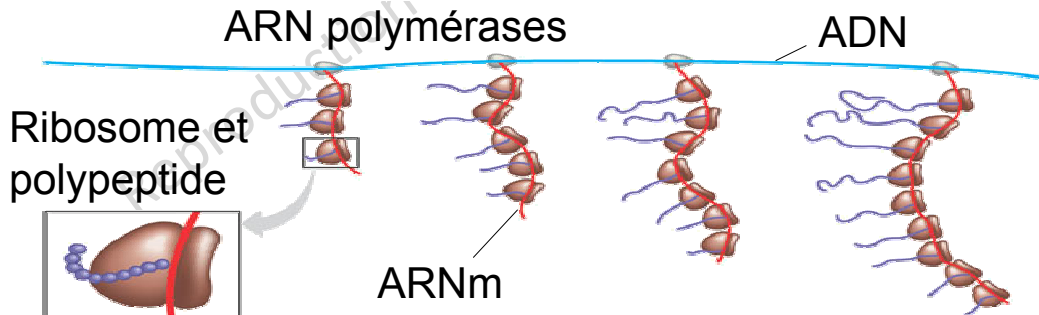
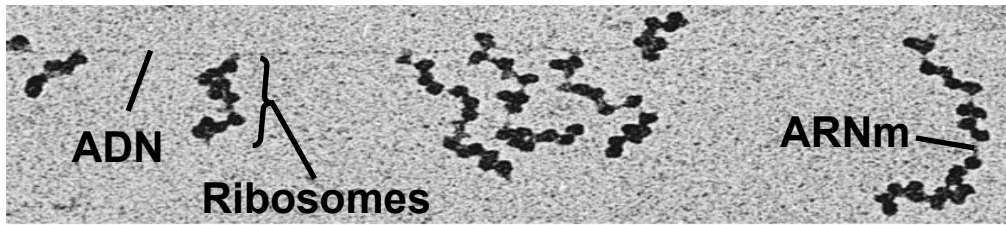
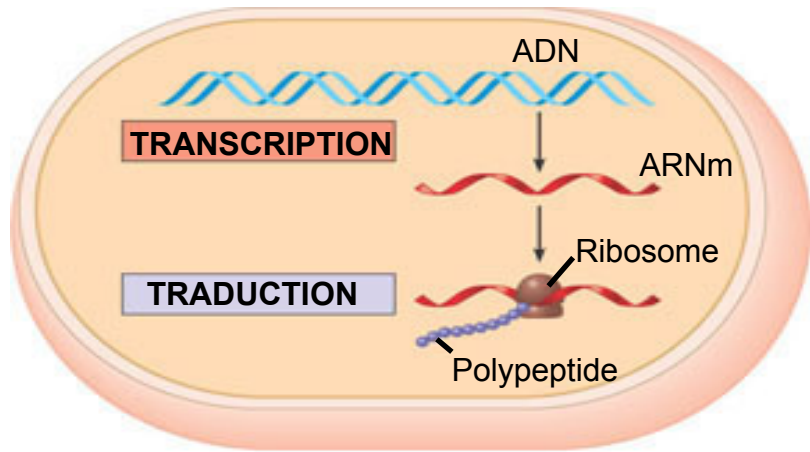
Vue générale de la transcription

- **Transcription = Synthèse d'ARN ~ au brin sens**
 - **Trois ARNs polymérases, trois catégories de gènes**
 - **Gènes de classe I transcrits par l'ARN POL I**
 - Les ARNs ribosomiaux (ARNr) 5.8S, 18S et 28S
 - **Gènes de classe II transcrits par l'ARN POL II**
 - **Les ARNs messagers**
 - Les petits ARNs nucléolaires (snoARNs, small nucleolar)
 - » snoARNs : maturation des ARNr
 - Les petits ARNs nucléaires (snARNs, small nuclear) sauf U6
 - » snARNs : maturation des ARNm
 - Les microARNs (régulation post-transcriptionnelle)
 - **Gènes de classe III transcrits par l'ARN POL III**
 - Les ARNs de transfert (ARNt)
 - L'ARNr 5S
 - Le snRNA U6

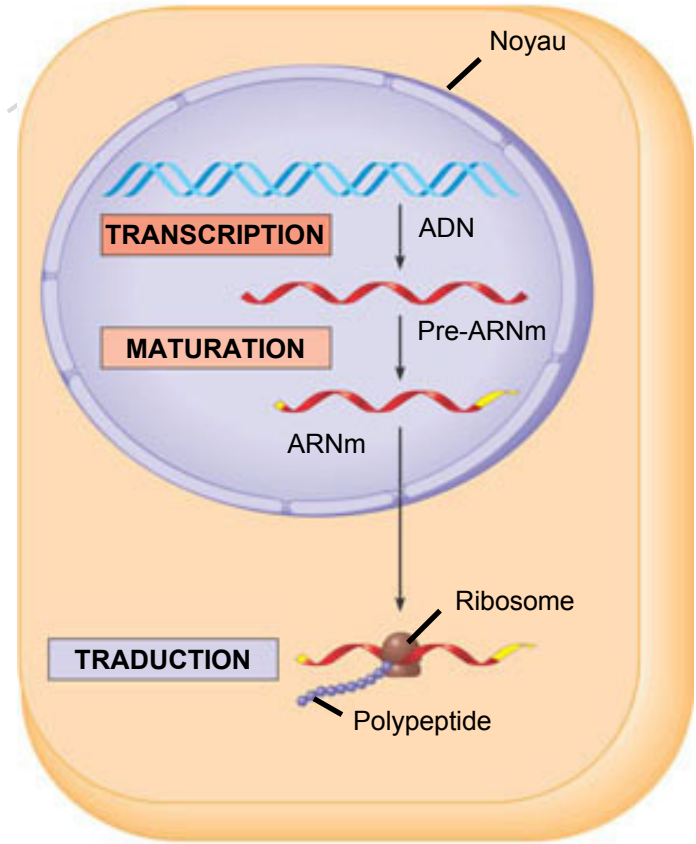
Vue générale de la transcription

• Différences entre procaryotes et eucaryotes

Cellule procaryote. En l'absence de noyau, transcription et traduction sont simultanées



Cellule eucaryote. L'ARN subit une maturation avant d'être traduit dans le cytosol



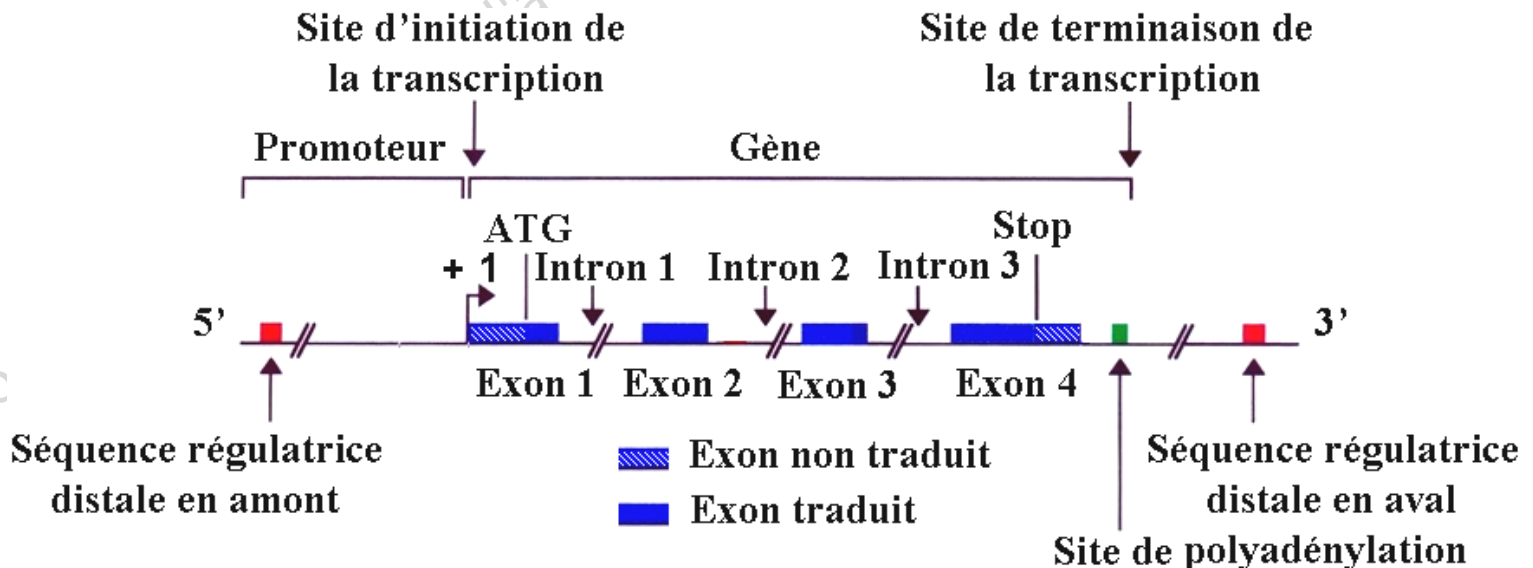
Structure d'un gène → une protéine

• L'unité de transcription

- Ensemble de séquences transcrites, pas forcément traduites
 - Débute au 1^{er} exon (nt noté +1), s'arrête ~ au site de polyadénylation
 - Les signaux ATG et Stop délimitent la séquence qui sera traduite
 - Amont et aval de ces signaux: 5'-UTR et 3'-UTR (*UnTranslated Region*)

• Le promoteur

- Séquences en amont ou en aval de l'unité de transcription
 - Nécessaires à l'initiation de la transcription et à sa régulation

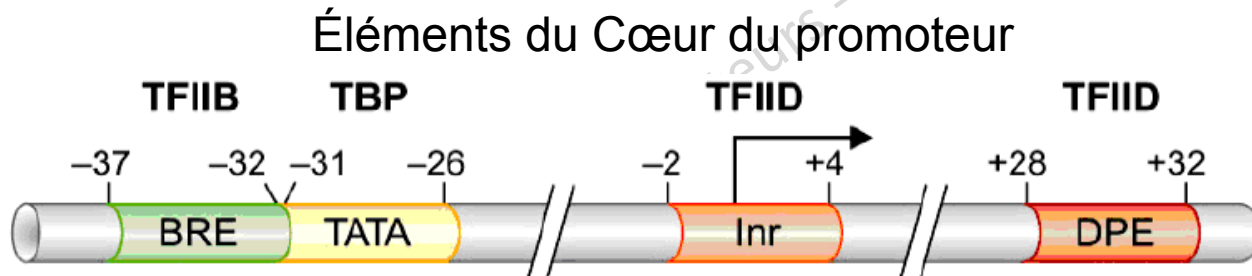


Structure d'un gène → une protéine

• Le promoteur

– Le cœur du promoteur

- Situé à proximité immédiate du site d'initiation de la transcription
- Il comprend diverses **séquences quasi constantes** (TATA box,..)
→ **Assemblage de la machinerie basale de transcription**



BRE: *TFIIB Recognition Element*; Inr: *Initiator*; DPE: *Downstream Promoter Element*

– Le promoteur proximal et distal

- Respectivement situés entre les positions -50 à -200 et plusieurs kb en amont voire en aval du gène
- **Combinaison de séquences variable** d'un gène à l'autre
- Fixent des **facteurs de transcription (FT) spécifiques**
 - Qui facilitent ou réduisent l'assemblage de la machinerie basale
 - Et déterminent **quand et à quel taux un gène doit être transcrit**

Machinerie basale de transcription

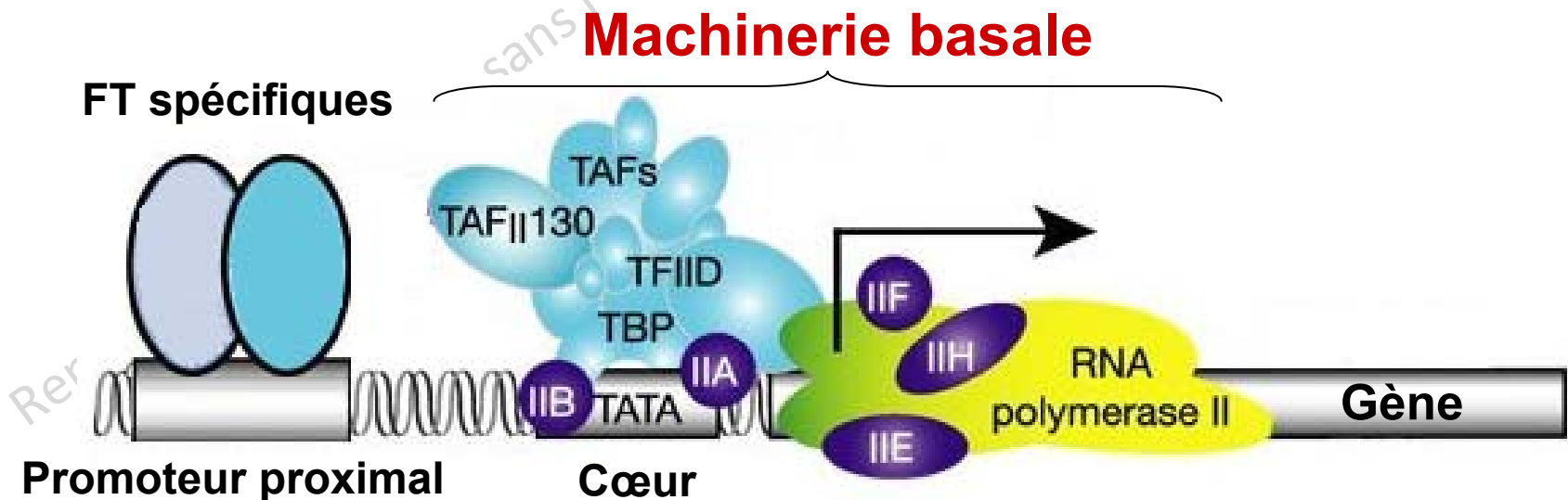
- **Constituée par:**

- **L'ARN Polymérase II**

- Complexe multiprotéique, dizaine de sous-unités chez les eucaryotes
 - Le domaine CTD (*Carboxy-Terminal Domain*), spécifique de l'ARNPII
 - » Motif de phosphorylation (Tyr-**Ser**-Pro-Thr-**Ser**-Pro-Ser) répété 50x
- Incapable seule de se fixer à l'ADN et d'initier la transcription

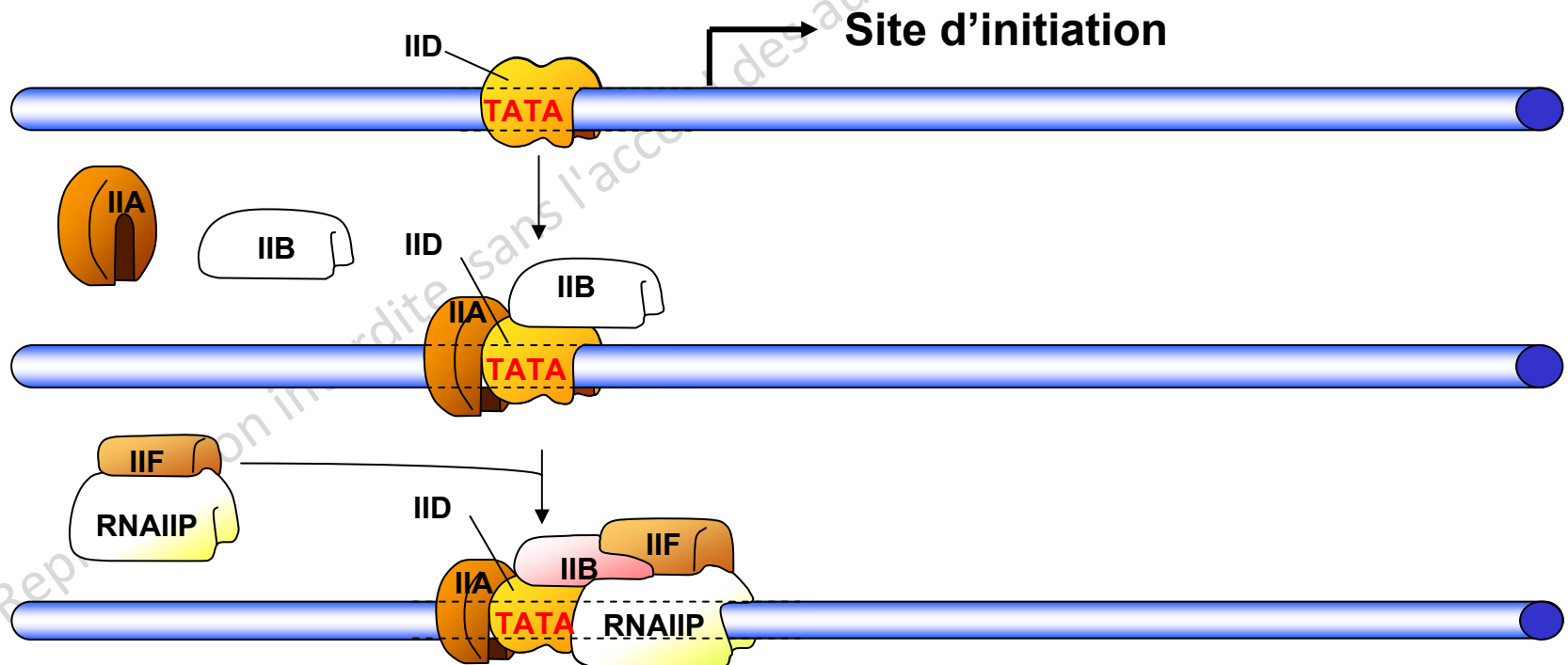
- **Assistée des facteurs généraux de transcription**

- TFII A, B, D, E, F et H, chacun constitué de plusieurs sous-unités
- Ils permettent notamment la liaison à l'ADN de l'ARNPII et son activation



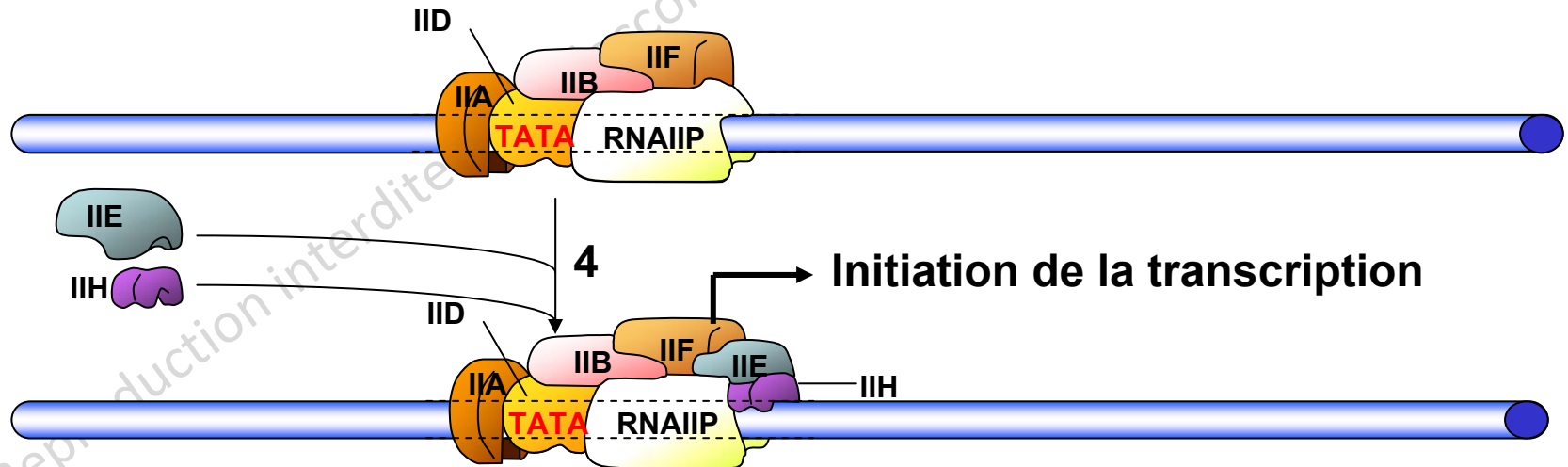
Initiation de la transcription

- **Assemblage et activation de la machinerie basale**
 - **Début par la fixation de TFIID sur la boîte TATA**
 - S'y fixe par sa sous-unité TBP (*TATA Binding Protein*)
 - Autres composants de TFIID appelés TAF (*TBP-Associated Factors*)
 - **Les autres complexes sont ensuite recrutés**
 - TFIIA et TFIIB qui recrute TFIIIF et l'ARN Pol II



Initiation de la transcription

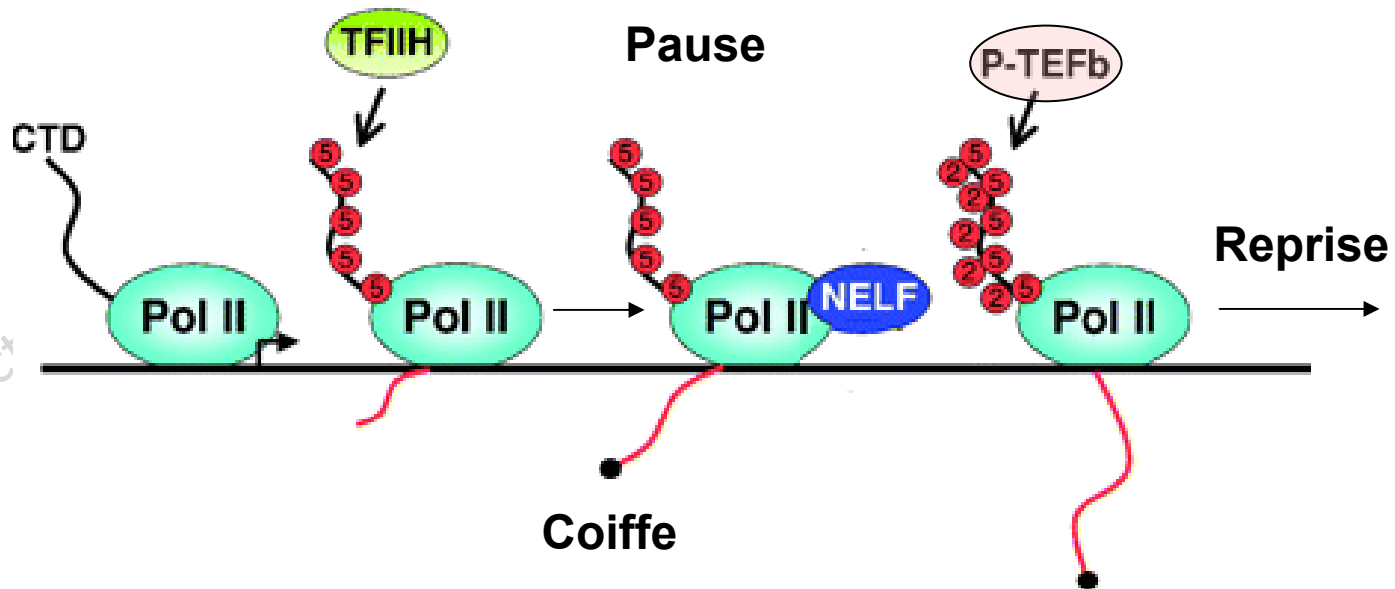
- **Assemblage et activation de la machinerie basale**
 - **Les autres complexes sont ensuite recrutés**
 - TFII E et TFII F sont enfin recrutés
 - L'ensemble forme le **complexe de préinitiation (PIC)**
 - **La transcription débute grâce à TFII H**
 - Son activité hélicase ouvre la double hélice (→ bulle de transcription)
 - Son activité kinase phosphoryle la sérine en position 5 du CTD de l'ARNPII: le PIC est activé (synthèse des premiers nucléotides)



Complexe de préinitiation (PIC)

Élongation de la transcription

- Des facteurs d'élongation sont recrutés
 - Le facteur **NELF** (*Negative Elongation Factor*)
 - Il met l'ARN Pol II qui vient juste de quitter le promoteur **en pause**
 - Ceci permet la **mise en place de la coiffe** sur l'ARNm naissant
 - Coiffe = Modification qui protège l'ARNm de la dégradation
 - La maturation de l'ARNm est ainsi est couplée à l'élongation
 - Le facteur **P-TEFb** (*Positive Transcription Elongation Factor*)
 - Il phosphoryle la sérine en position 2 du CTD → L'ARN Pol II repart



Terminaison de la transcription

- **À peu près au niveau du signal de polyadénylation**
 - Chez les eucaryotes, il n'y pas de séquence signalant à la polymérase où arrêter précisément la transcription
 - Elle s'arrête lorsque les enzymes de clivage du transcrit primaire qui l'accompagnent rencontrent les signaux de polyadénylation
 - Pour un gène donné, l'arrêt se produit de 0,5 à 2kb après ces signaux (longueur variable des transcrits primaires)

Maturation des transcrits primaires

- Doivent tous être modifiés pour être fonctionnels (sauf les ARNm des histones)

- La coiffe (modifications de l'extrémité 5')

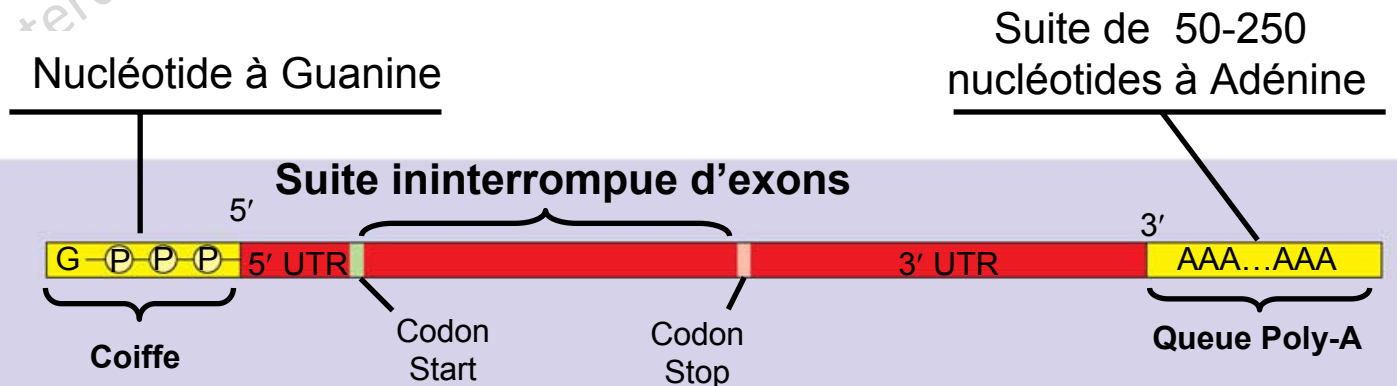
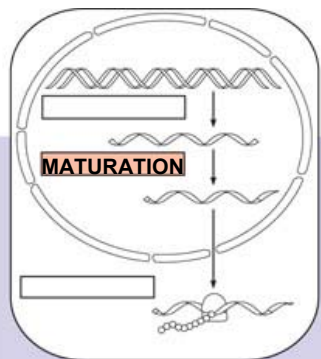
- Protège le transcrit de la dégradation, augmentant sa durée de vie
- Nécessaire à sa reconnaissance par la machinerie traductionnelle

- La polyadénylation (ajout d'une queue polyA en 3')

- Elle ralentit aussi sa dégradation par des exonucléases

- L'épissage (excision des introns)

- Met bout à bout la séquence codante des différents exons

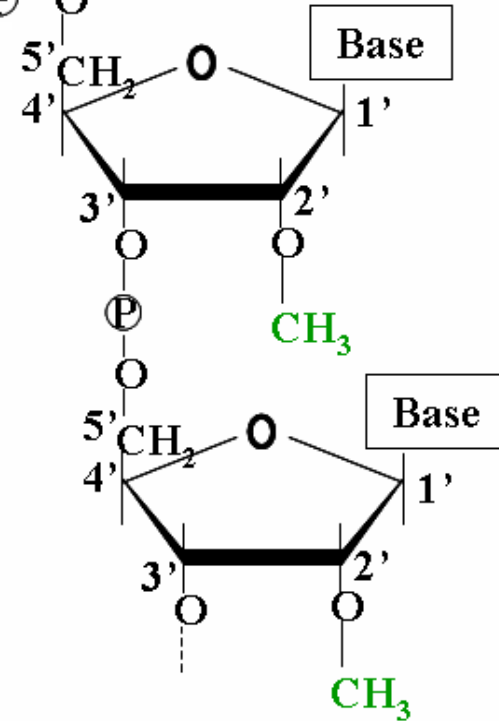
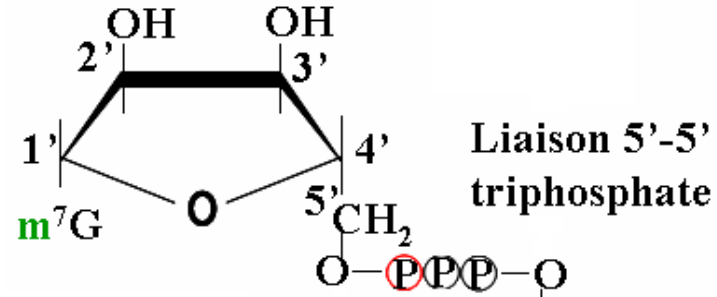
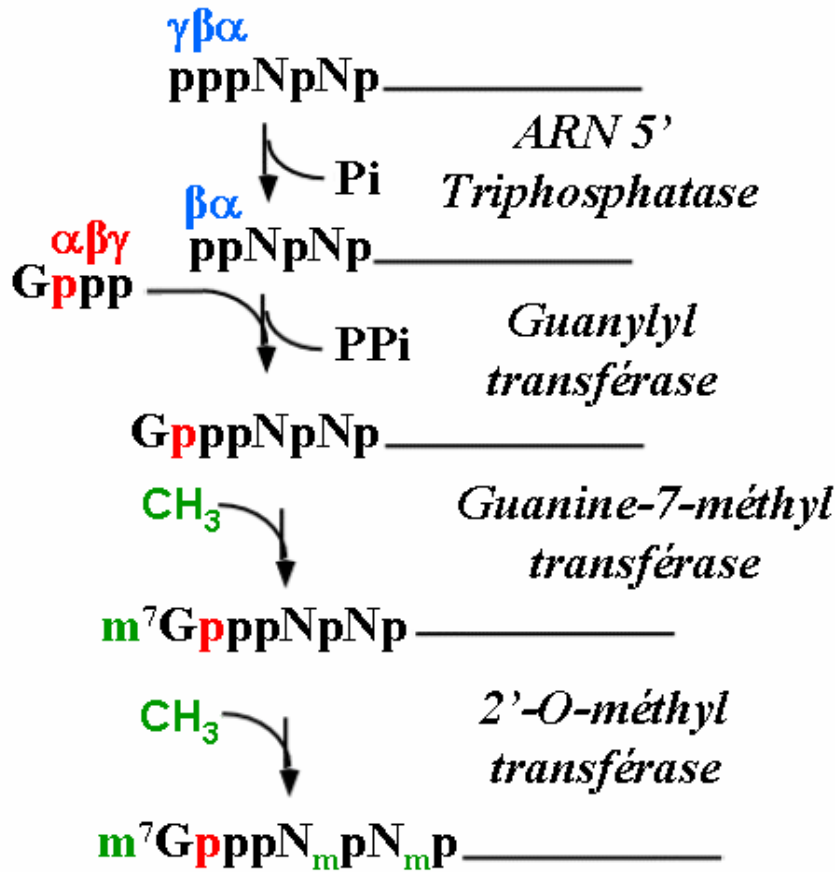


La coiffe de l'ARNm

- **Ajout d'un nucléotide à guanine à l'extrémité 5'**
 - **Deux activités enzymatiques liées à la même protéine**
 - L'activité ARN-5'-triphosphatase élimine le phosphate γ du transcrit
 - L'activité Guanylyl-transférase relie un GTP par son phosphate α au phosphate β du transcrit (liaison 5'-5' triphosphate)
- **Ajout de groupes méthyle**
 - **Deux enzymes différentes**
 - Méthylase de l'azote N7 du nucléotide ajouté
 - Et 2'-O-méthylase du ribose des deux premiers nucléotides

La coiffe de l'ARNm

ine - UNS

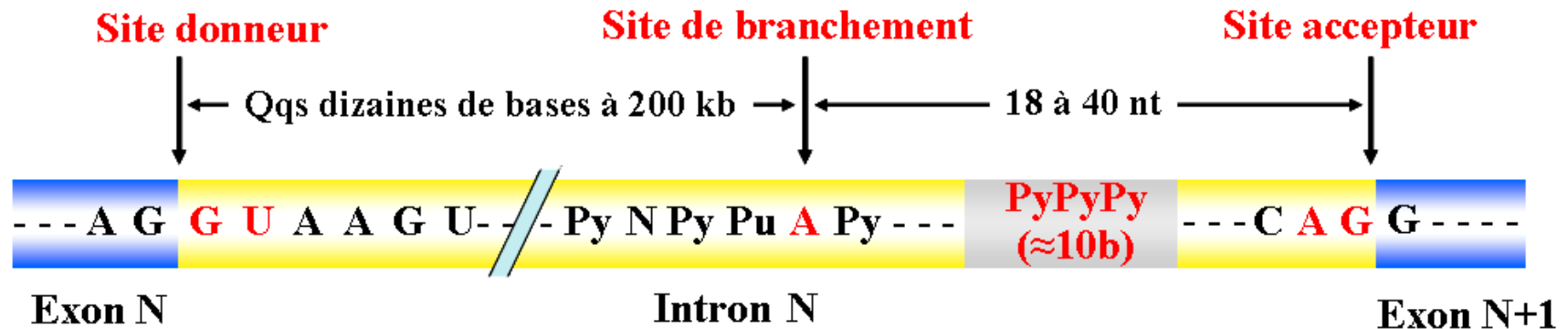


REV

L'épissage de l'ARNm

- Il nécessite:

- Des séquences introniques ~ constantes (= consensus)
 - Site donneur (GU) et accepteur (AG) au début et à la fin de l'intron
 - Avant la fin de l'intron, suite de pyrimidine et site de branchement (= Adénine de la séquence PyNPyPu**A**Py)

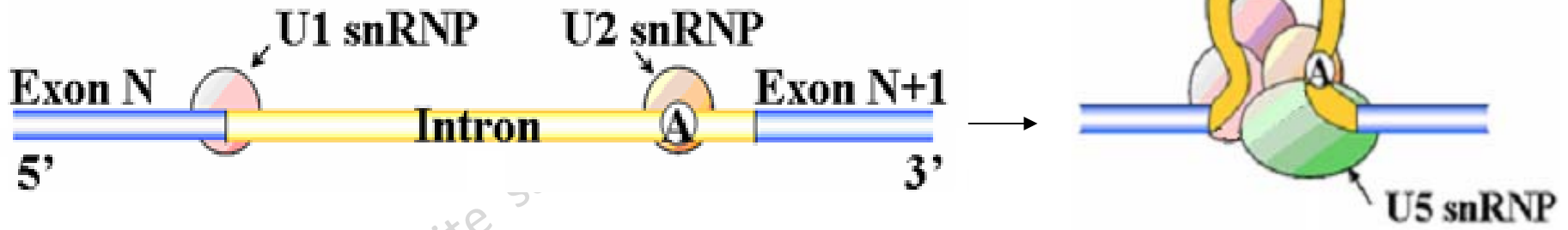


- et les ribonucléoprotéines U1, U2, U4, U5 et U6

- Complexes enzymatiques qui assurent l'épissage
- Constituées de protéines communes et spécifiques
- Et des snRNAs correspondants

L'épissage de l'ARNm

- **Rôle des ribonucléoprotéines**
 - U1 et U2 se fixent au site donneur et de branchement
 - U4, U5 et U6 interagissent avec U1 et U2
 - Ce qui rapproche les exons



Reproduction interdite

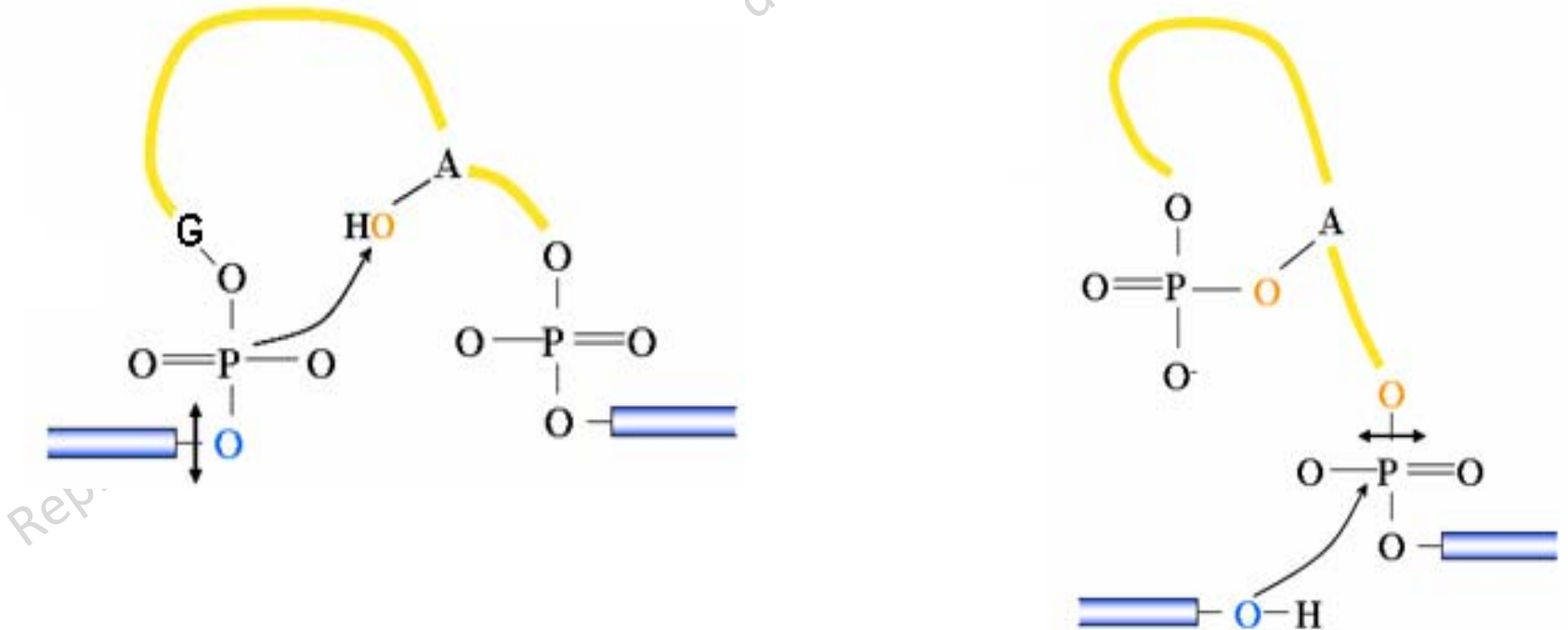
auteurs - Faculté de Médecine - UNS

L'épissage de l'ARNm

- **Rôle des ribonucléoprotéines**

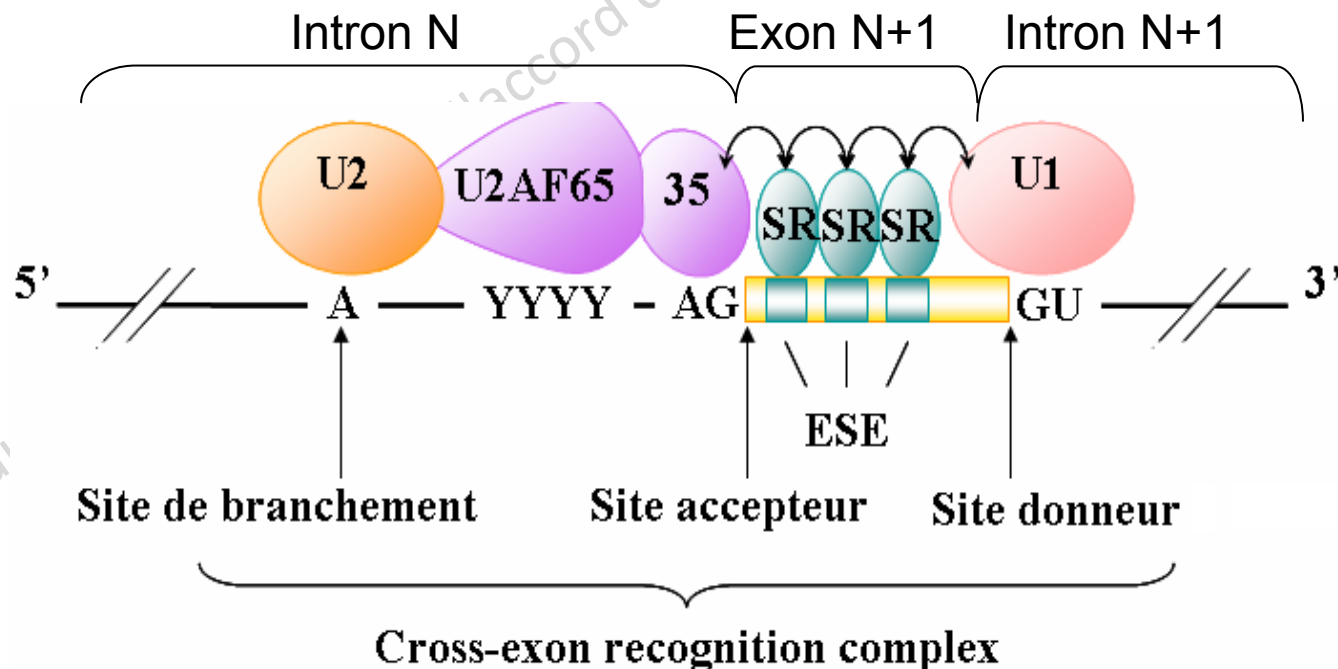
- **Deux réactions de transestérification éliminent l'intron**

- Clivage exon-intron en 5', liaison 5'-2' phosphodiester entre la guanine du site donneur et l'adénine du site de branchement (l'intron forme un lasso)
- Clivage intron-exon en 3', **liaison 3'-5' phosphodiester** entre exons



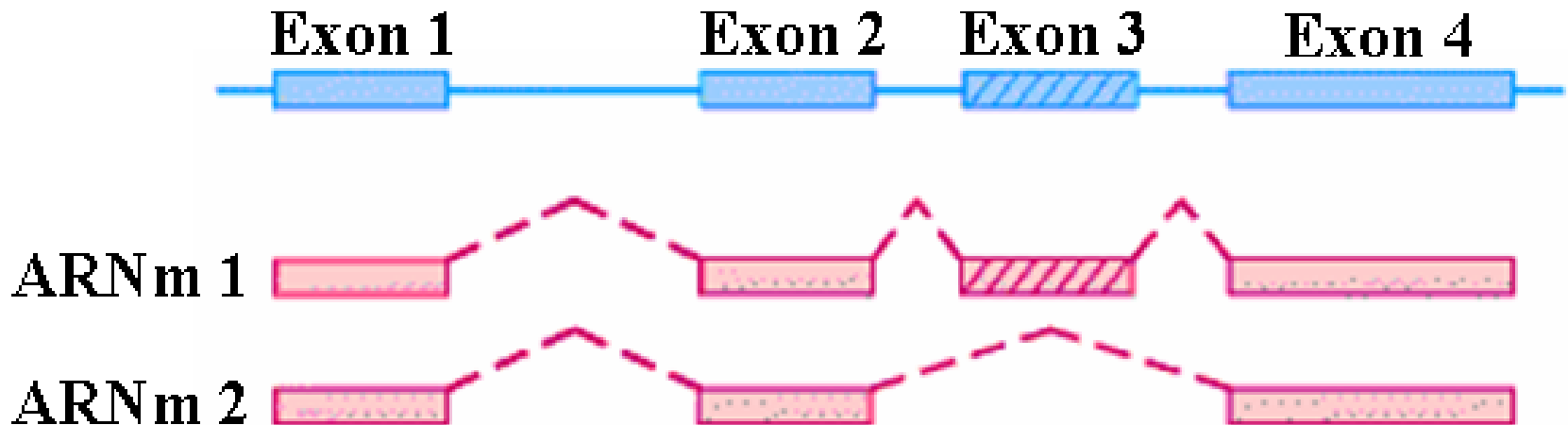
L'épissage de l'ARNm

- Un autre complexe aide à la détection des exons
 - Les séquences des introns peuvent être très longues
 - « **Cross-exon recognition complex** »
 - Établit un pont entre U1 et U2 de part et d'autre de l'exon
 - Comprend des protéines liées au site accepteur, à la suite de pyrimidines et aux séquences ESE (*Exonic Splice Enhancer*)



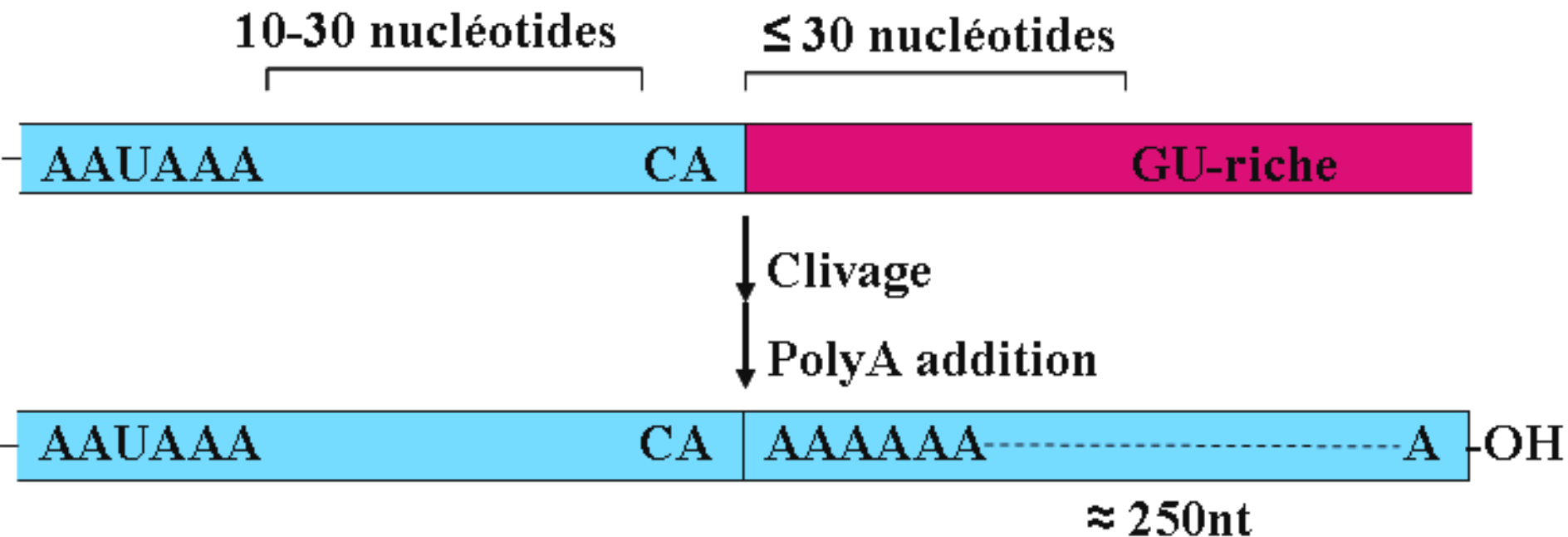
L'épissage de l'ARNm

- Des exons peuvent être exclus de l'ARNm
 - L'épissage alternatif augmente la diversité des protéines produites à partir d'un pré-ARNm
 - 200 000 protéines sont produites par 30 000 gènes
 - Protéines de fonctions analogues ou totalement différentes
 - Des facteurs d'épissage spécifiques d'un tissu → Protéine spécifique de ce tissu à partir d'un gène présent dans toutes les cellules



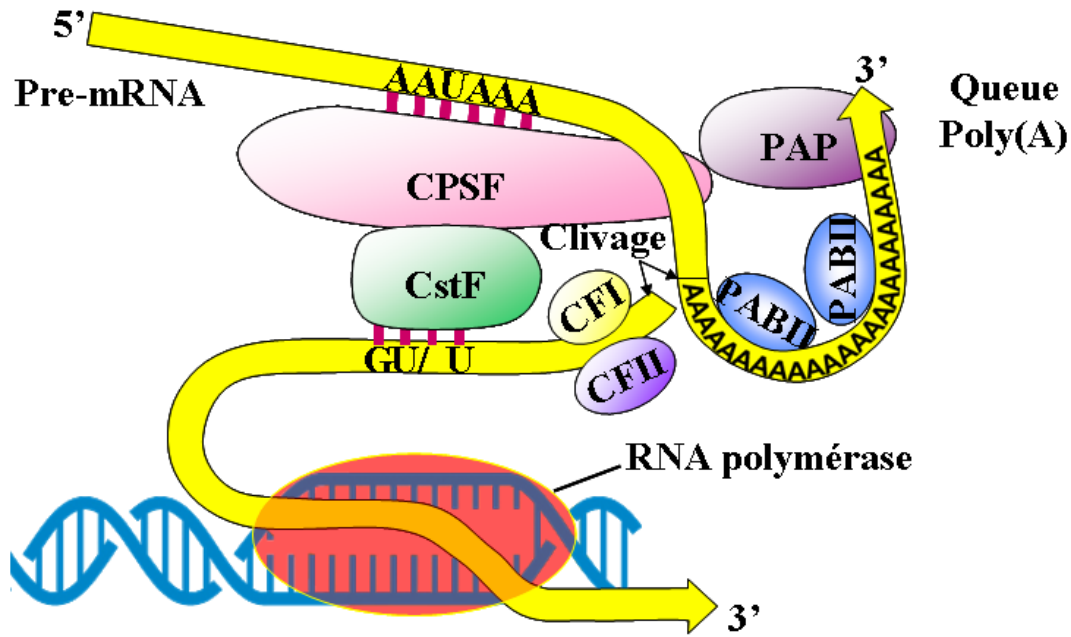
La polyadénylation de l'ARNm

- **Nécessite plusieurs séquences**
 - **Séquence de polyadénylation AAUAAA**
 - **Séquence de clivage du messenger CA**
 - La queue Poly(A) (suite ~250 nucléotides à adénine) lui sera ajoutée
 - **Séquence GU-riche**



La polyadénylation de l'ARNm

- **Séquence GU-riche fixe CstF (*Cleavage Stimulating Factor*)**
 - Active les protéines CFI et CFII (*Cleavage Factor*)
- **Site de clivage reconnu par CFI et CFII qui le coupent**
- **Séquence de polyadénylation AAUAAA fixe CPSF (*Cleavage and Polyadenylation Specificity Factor*)**
 - Active la poly(A) polymérase (PAP) qui synthétise la queue poly(A)
 - Cette synthèse est accélérée par PABPII (PolyA Binding Protein II)



Mentions légales

L'ensemble de ce document relève des législations française et internationale sur le droit d'auteur et la propriété intellectuelle.

Tous les droits de reproduction de tout ou partie sont réservés pour les textes ainsi que pour l'ensemble des documents iconographiques, photographiques, vidéos et sonores.

Ce document est interdit à la vente ou à la location par un tiers autre que l'Université de Nice-Sophia Antipolis.

La diffusion, la duplication, la mise à disposition du public (sous quelque forme ou support que ce soit), la mise en réseau, de tout ou partie de ce document, sont strictement réservées à l'Université de Nice-Sophia Antipolis.

L'utilisation de ce document est strictement réservée à l'usage privé des étudiants inscrits aux cours et au tutorat organisés par l'UFR de Médecine de l'Université de Nice-Sophia Antipolis, et non destinée à toute autre utilisation privée ou collective, gratuite ou payante.