

Cours de Biologie Moléculaire



F. GESBERT

franck.gesbert@universite-paris-saclay.fr

Expression des gènes





Que savons nous?

- L'ADN est le support de l'information génétique
- L'ADN est structuré en une double hélice
- Chaque brin est composé de nucléotides puriques et pyrimidiques
- Les deux brins sont complémentaires et antiparallèles

- L'ADN se réplique de manière semi-conservative
- La réplication fait intervenir des complexes macromoléculaires
- Les ADN polymérase synthétisent l'ADN de 5'-3'



Le messager manquant

L'intuition de l'existence d'un "messager manquant"

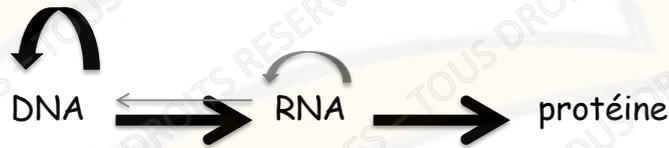
Observations:

- L'ADN porte l'information
- les ribosomes, qui contiennent de l'ARN (ARNr) participent à la synthèse protéique
- chez les eucaryotes ADN est nucléaire, ribosomes sont cytoplasmiques
- sans ARNr il n'y a pas de protéines

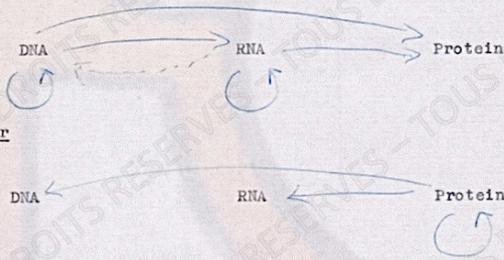
Question: comment une protéine est elle produite à partir d'un gène ?

Hypothèse: en 1960, Jacob et Monod émettent l'hypothèse et démontrent l'existence d'une nouvelle population hétérogène d'ARN, les ARN messagers

F.C. Crick, 1956: "the central dogma* in molecular biology"



The Central Dogma: "Once information has got into a protein it can't get out again". Information here means the sequence of the amino acid residues, or other sequences related to it. That is, we may be able to have



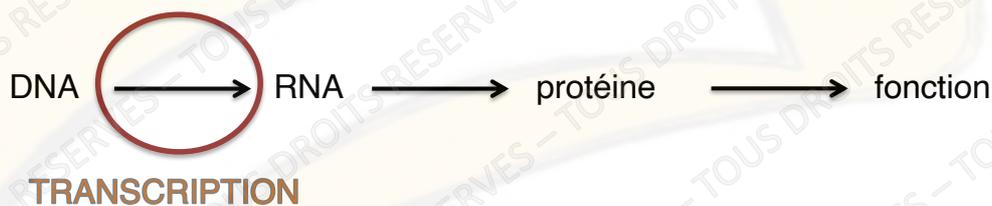
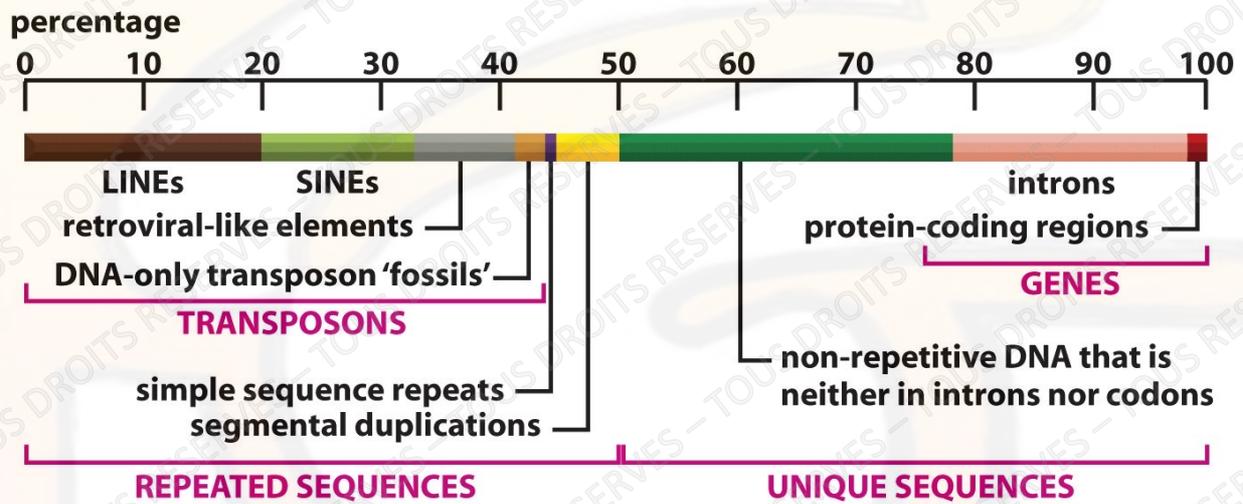
Crick's first outline of the central dogma, from an unpublished note made in 1956. Credit: Wellcome Library, London.



Credit: Cold Spring Harbor Laboratory

Définition contemporaine du gène

Un gène est une séquence d'acide nucléique ayant un locus précis sur un chromosome et qui code pour au moins un ARN ou une protéine ayant une fonction définie



Transcription:

- synthèse d'ARN à partir d'une matrice ADN

- Nécessite:

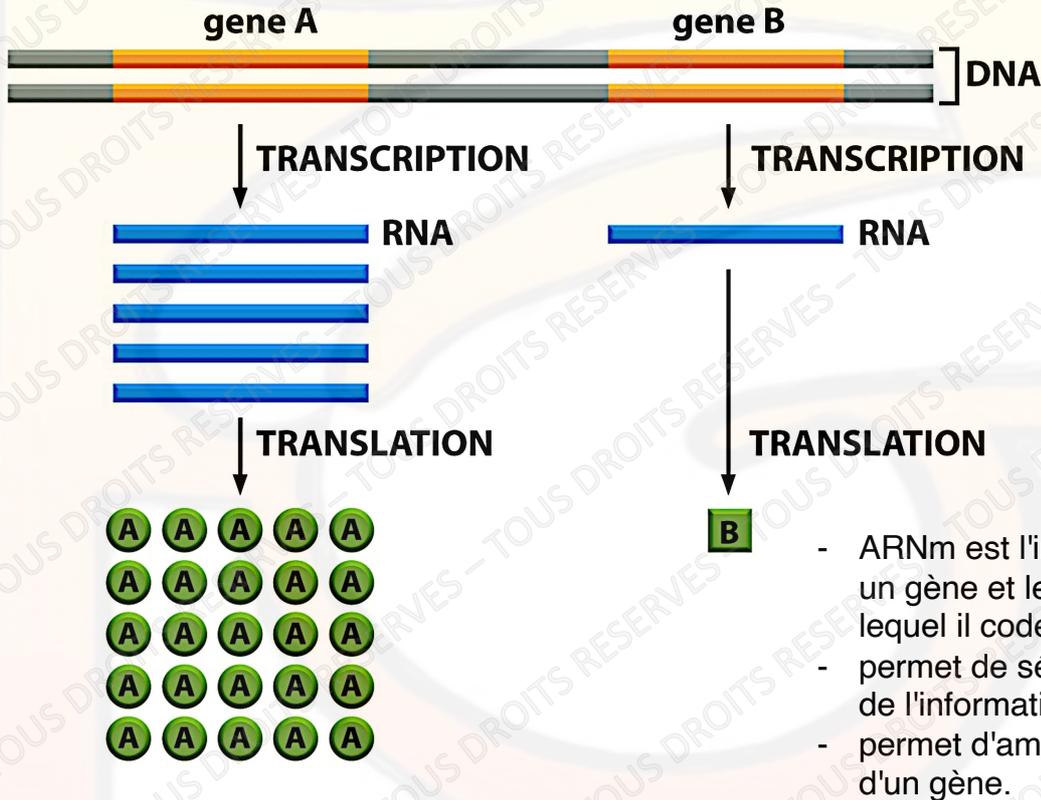
- Une matrice ADN

- une ARN polymérase

- des NTP (ATP, GTP, CTP et UTP **sous forme Ribo...**)



La Transcription



La Transcription



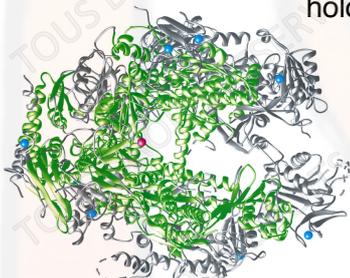
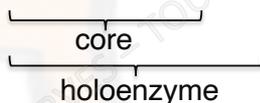
Les ARN polymérases

Les ARN polymérases connues sont formées de plusieurs sous unités

Procaryotes

Une seule ARN polymérase pour tous les gènes.

5 sous unités (α , α' , β , β' , ω) + σ



Eucaryotes

ARN pol I : ARNr 18s; 28s; 5,8s (nucléole)

ARN pol II: ARNm (noyau), miARN

ARN pol III: ARNt, ARNr (5s), μ ARN (noyau)

ARN pol IV: siARN (plantes)

ARN pol V: petits ARN (plantes)

de 12 à 16 sous unités

ARN pol mitochondriale est un cas particulier

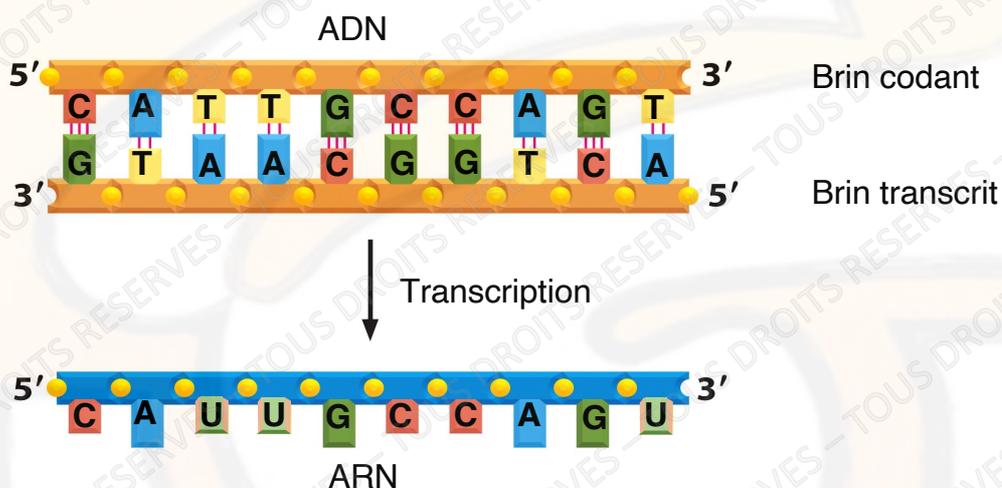


La transcription: mécanismes de base

- Les ARN polymérases sont des **ARN synthétases ADN dépendantes**
 - Ne nécessitent pas d'amorce
 - Synthétisent un ARN de 5'-3' en utilisant un simple brin d'ADN matrice
 - La séquence synthétisée est complémentaire et antiparallèle du brin ADN matrice
- Pour initier la transcription, l'ARN polymérase s'associe à un site sur la double hélice d'ADN
 - Ce site est le promoteur
 - le promoteur est toujours en 5' du gène à transcrire
 - le promoteur est donc orienté et il détermine le brin qui sera transcrit



Quelques conventions à connaître et à respecter



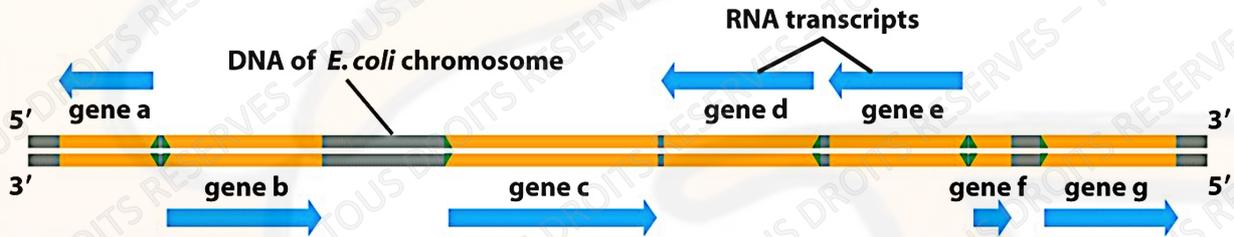
Le brin transcrit, est le brin lu par la polymérase

Le brin codant est le brin qui possède la même séquence que l'ARNm, au U près.

Par convention, lorsque l'on donne la séquence d'un gène, on donne la séquence du brin codant.



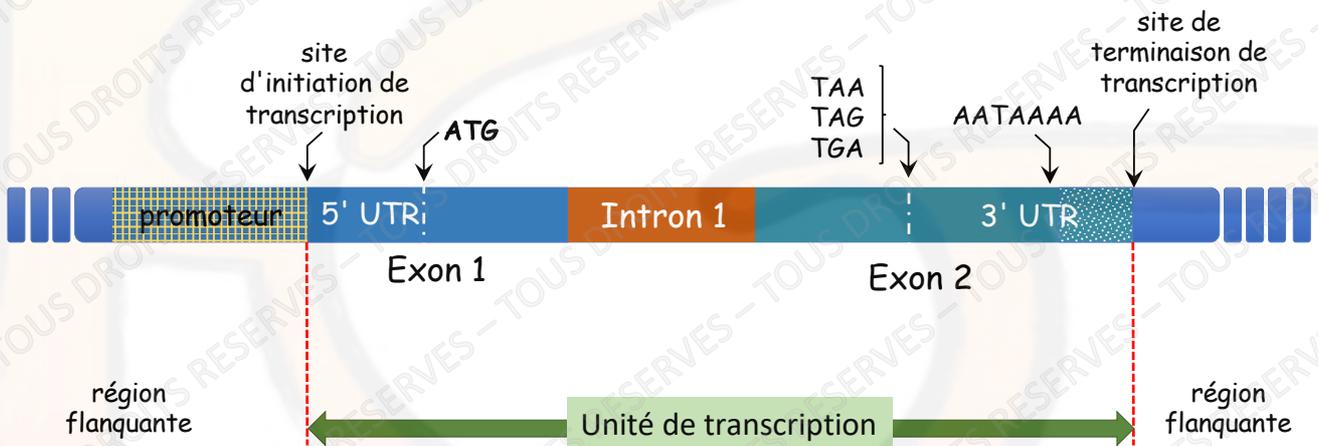
Les gènes peuvent résider sur les deux brins de la double hélice



Transcription eucaryote

définition de l'unité de transcription eucaryote

Rappel: les gènes eucaryotes sont morcellés



UTR: "UnTranslated Region": Région transcrite et non traduite

Transcription eucaryote-Initiation

Promoteur minimum: séquence d'environ 40 à 60 nucléotides, encadrant le site d'initiation et permettant l'assemblage du **Complexe de Pré-Initiation**.

Complexe de Pré-Initiation: Complexe constitué par l'ARN pol II et des Facteurs Généraux de la Transcription (l'ensemble des FGT remplit les fonction de σ des procaryotes).

Expression des gènes et antibiotiques

Transcription et antibiotiques



Rifamycin B

Rifamycine: naturellement produite par *Amiclatopsis mediterranei*.

Rifampicine: produit de synthèse

Mode d'action: **inhibition de l'initiation** de la **transcription bactérienne** par occlusion stérique.



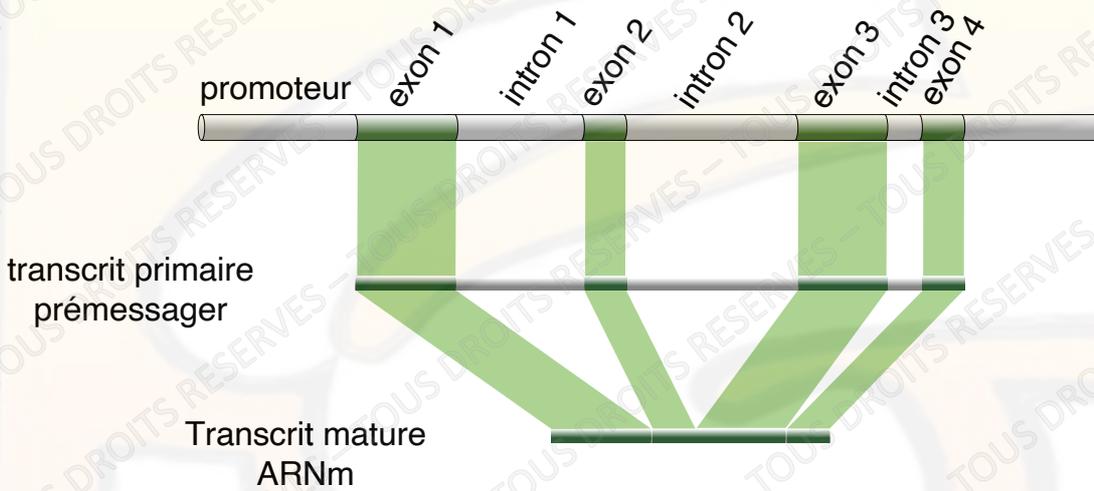
Actinomycine D: produit par *Streptomyces*

Mode d'action: **bloque l'élongation** de la transcription en s'intercalant dans la double hélice d'ADN au moment de sa déstabilisation par l'ARN polymérase en progression.



La Transcription - Maturation du transcrit

A+

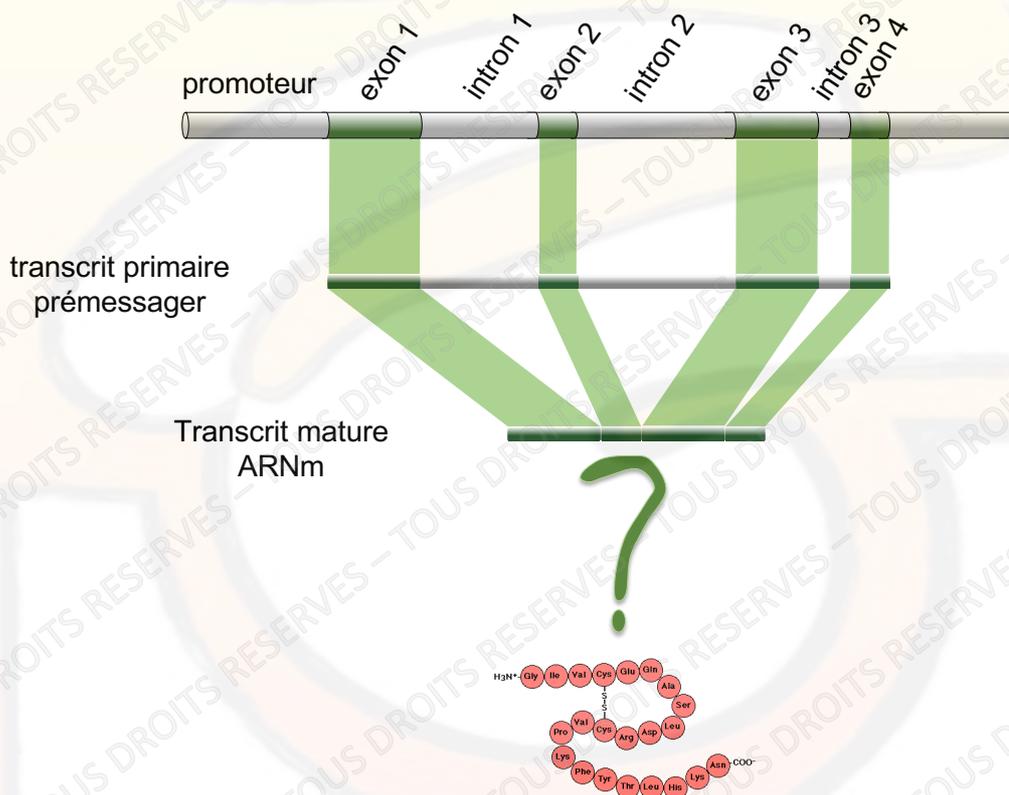


Compartimentation cellulaire et morcellement des gènes impliquent une maturation et un transports des transcrits



La Traduction

B





Question: comment est réalisé le transfert d'information pour passer d'une molécule d'acide nucléique (ARNm) à une molécule d'acides aminés?

- Les acides aminés n'ont pas d'affinité particulière pour les purines ou pyrimidines
- 1955: F. Crick a l'intuition qu'il faut que les acides aminés soient associés à une molécule d'ARN qui puisse interagir (par complémentarité) avec la molécule d'ARNm
- 1957: Zamenick et Hoagland démontrent la présence d'une fraction de petits ARN qui se lient aux acides aminés avant leur incorporation dans les peptides.
- Ces ARN sont appelés ARN de transfert: **ARNt**
- Les enzymes responsables de la liaison AA- ARNt sont les **Amino-Acyl-ARNt synthétases**



Le code génétique

preuve génétique apportée par Crick et Brenner :

- 1- le code est lu par blocs de trois nucléotides: un bloc=un codon
- 2- Il existe 64 codons possibles (4^3)
- 3- 64 codons codent pour 20 acides aminés
- 4- le code n'est pas strict: il est dégénéré

Crick F H C, Barnett L, Brenner S, Watts-Tobin R J. General nature of the genetic code for proteins. Nature 192 : 1227-1232. 1961



Décryptage du code génétique

UUU	UCU	UAU	UGU
UUC	UCC	UAC	UGC
UUA	UCA	UAA	UGA
UUG	UCG	UAG	UGG
CUU	CCU	CAU	CGU
CUC	CCC	CAC	CGC
CUA	CCA	CAA	CGA
CUG	CCG	CAG	CGG
AUU	ACU	AAU	AGU
AUC	ACC	AAC	AGC
AUA	ACA	AAA	AGA
AUG	ACG	AAG	AGG
GUU	GCU	GAU	GGU
GUC	GCC	GAC	GGC
GUA	GCA	GAA	GGA
GUG	GCG	GAG	GGG



<chem>C(C)C(=O)O</chem> Valine	<chem>CC(C)C(=O)O</chem> Isoleucine	<chem>CC(C)CC(=O)O</chem> Leucine	<chem>CSCC(=O)O</chem> Méthionine	<chem>c1ccc(cc1)CC(=O)O</chem> Phénylalanine
<chem>C(=O)O</chem> Glycine	<chem>CC(=O)O</chem> Alanine	<chem>C1CCNC1</chem> Proline	<chem>C1=CC=C(C=C1)C(=O)O</chem> Tryptophane	<chem>Oc1ccc(cc1)CC(=O)O</chem> Tyrosine
<chem>C(O)C(=O)O</chem> Sérine	<chem>CC(O)C(=O)O</chem> Théonine	<chem>CSCC(=O)O</chem> Cystéine	<chem>NC(=O)CC(=O)O</chem> Asparagine	<chem>CC(C)C(=O)O</chem> Glutamine
<chem>C(CCN)C(=O)O</chem> Arginine	<chem>CCCCN</chem> Lysine	<chem>C1=CN=C(N1)C(=O)O</chem> Histidine	<chem>C(C(=O)[O-])C(=O)O</chem> Acide aspartique	<chem>C(CC(=O)[O-])C(=O)O</chem> Acide glutamique

64 (4³) codons possibles

20 acides aminés



Décryptage du code génétique

	Seconde base					
	U	C	A	G		
P r e m i è r e b a s e	U	Phe Phe Leu Leu	Ser Ser Ser Ser	Tyr Tyr STOP STOP	Cys Cys STOP Trp	U C A G
	C	Leu Leu Leu Leu	Pro Pro Pro Pro	His His Gln Gln	Arg Arg Arg Arg	U C A G
	A	Ile Ile Ile Met	Thr Thr Thr Thr	Asn Asn Lys Lys	Ser Ser Arg Arg	U C A G
	G	Val Val Val Val	Ala Ala Ala Ala	Asp Asp Glu Glu	Gly Gly Gly Gly	U C A G
						T r o i s i è m e b a s e

- un codon spécifie toujours le même acide-aminé*.
- 3 codons ne codent pas pour un acide aminé. Spécifient un STOP.
- Met n'est codé que par AUG



Le code génétique- caractéristiques

- 1- Le code est lu sur l'ARNm de 5'-3'
- 2- Le code est lu par triplet
- 3- Le code est non chevauchant
- 4- Le code est non-ambigüe
- 5- Le code est dégénéré
- 6- Le code est universel
- 7- Le code comporte les signaux de début et de fin de lecture.



Toute molécule d'ARNm contient donc **trois cadres** possibles de lecture
C'est le **codon d'initiation (AUG)** qui spécifiera le **cadre ouvert de lecture (Open Reading Frame:ORF)**



Définition d'une séquence codante (CDS pour CoDing Sequence)

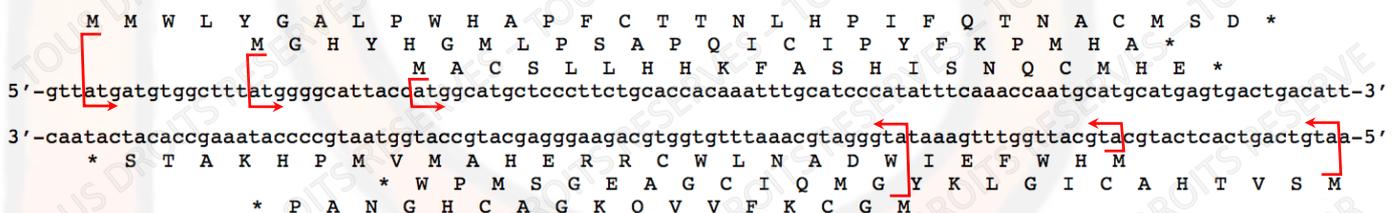
- Un codon initiateur (ATG)
- Une suite de codons sens
- Un codon STOP (Non sens)

Définition d'un cadre ouvert de lecture (ORF)

Région d'ADN qui sépare deux codons STOP

Sur chaque brin d'ADN 3 CDS possibles

Sur une séquence d'ADN double brin 6 CDS possibles



L'adaptateur: l'ARNt

- 70 à 90 nucléotides
- bases modifiées par édition post-transcriptionnelle
- structure tige-boucle (forme de trèfle)
- l'anticodon se situe dans une boucle (région simple brin, pouvant donc s'apparier...)
- séquence CCA en 3'-OH, impliquée dans la liaison avec l'acide aminé

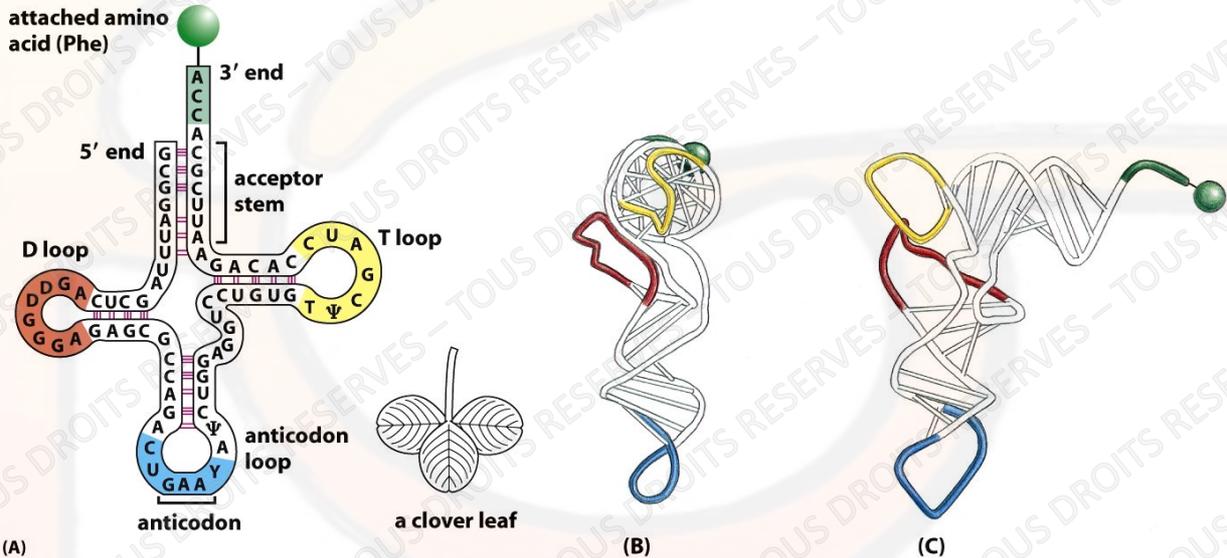


Figure 6-52 *Molecular Biology of the Cell* (© Garland Science 2008)

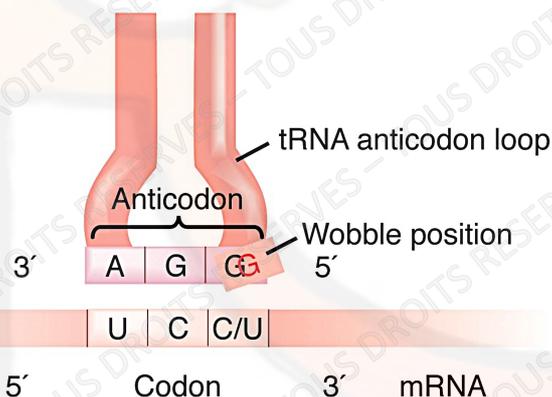
La Traduction – le WOBBLE

chez l'Homme: **WOBBLE** concept, F. Crick 1966

- 64 codons possibles
- 20 acides aminés
- 48 ARNt différents

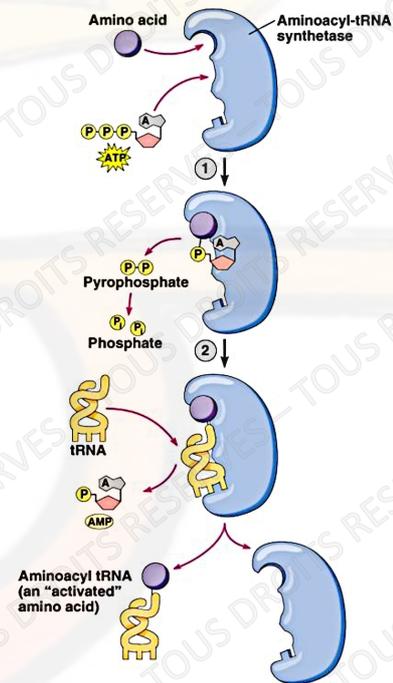
2 conclusions:

- un même acide aminé peut être codé par plusieurs codons: code dégénéré
- un même ARNt peut reconnaître plusieurs codons: WOBBLE



dans cet exemple UCC et UCU codent pour Ser à partir d'un même ARNt

- Activation des acides-aminés par les aminoacyl ARNt synthétases
- en règle générale 20 aminoacyl ARNt synthétases différentes
- il y a donc une aa-synthétase dédiée par acide aminé.
- réaction de couplage en 2 temps (intermédiaire aminoacyl-AMP)
- Etape de très grande fidélité avec un taux d'erreur $\leq 0,01\%$
- 1 reconnaissance de l'acide aminé
- 2 reconnaissance de caractéristiques structurales des ARNt
- 3 édition de l' aminoacyl-ARNt



Le ribosome n'aura aucun moyen de vérifier si l'ARNt est correctement chargé

- L'un des plus gros complexes moléculaires qu'il soit donné d'observer (2500 kDa).
- synthétise 2 à 20 acides aminés par seconde.

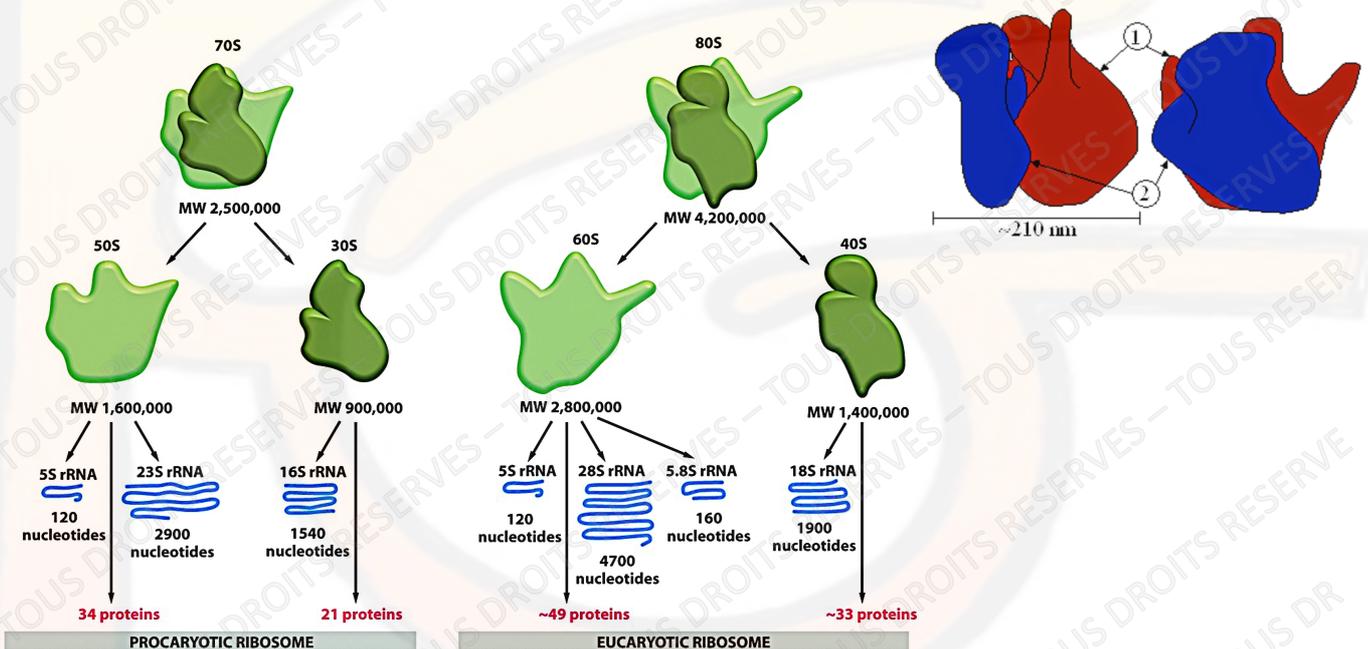


Figure 6-63 Molecular Biology of the Cell (© Garland Science 2008)

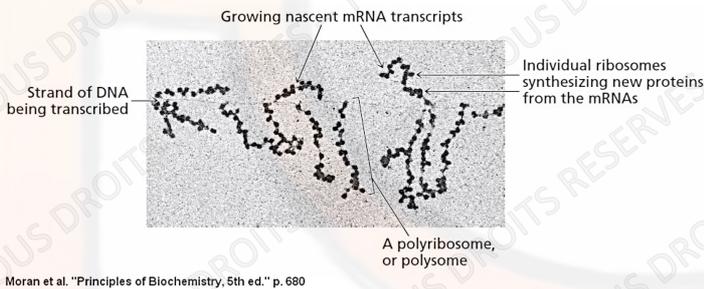
quelques rappels:

procaryotes:

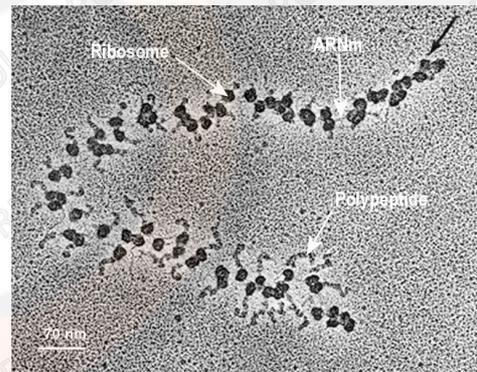
possibilité de traduction concomitante à la transcription
 le transcrit en cours de synthèse commence à être traduit
 messagers polycistroniques

eucaryotes:

la traduction n'a lieu que dans le cytoplasme sur un ARNm mature (monocistronique)
 les ARNr sont synthétisés (transcrits) dans le noyau
 les protéines ribosomales sont synthétisées dans le cytoplasme
 l'assemblage protéines ribosomales-ARNr est réalisé dans le noyau



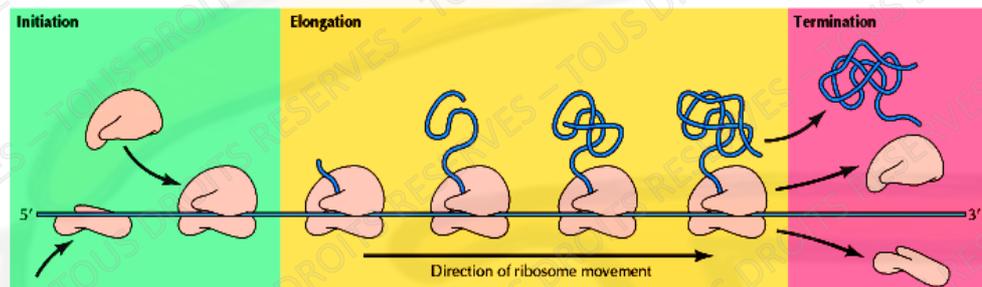
Moran et al. "Principles of Biochemistry, 5th ed." p. 680



La synthèse peptidique par le ribosome

3 étapes

- initiation
- élongation
- terminaison



Mécanismes procaryotes assez bien décrits

Spécificités eucaryotes avec des étapes moins bien différenciées

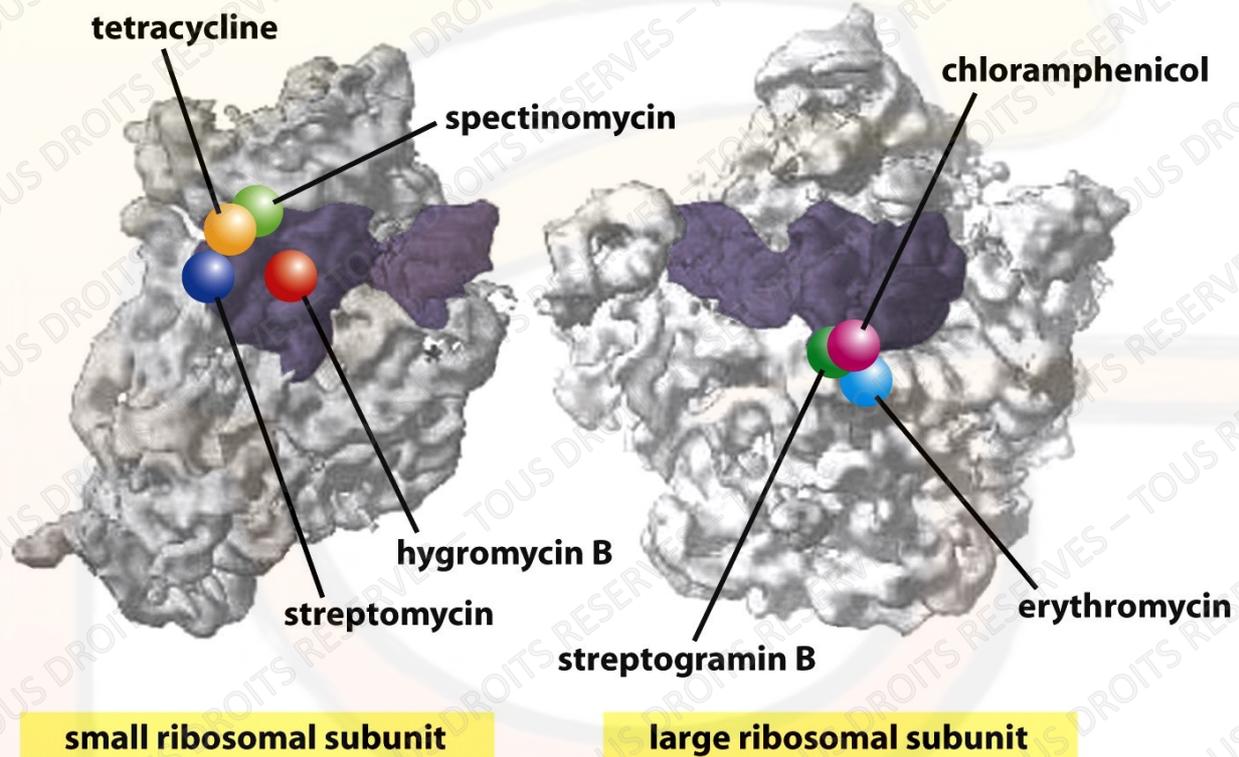


Table 6–4 Inhibitors of Protein or RNA Synthesis

INHIBITOR	SPECIFIC EFFECT
<i>Acting only on bacteria</i>	
Tetracycline	blocks binding of aminoacyl-tRNA to A-site of ribosome
Streptomycin	prevents the transition from translation initiation to chain elongation and also causes miscoding
Chloramphenicol	blocks the peptidyl transferase reaction on ribosomes (step 2 in Figure 6–66)
Erythromycin	binds in the exit channel of the ribosome and thereby inhibits elongation of the peptide chain
Rifamycin	blocks initiation of RNA chains by binding to RNA polymerase (prevents RNA synthesis)
<i>Acting on bacteria and eucaryotes</i>	
Puromycin	causes the premature release of nascent polypeptide chains by its addition to the growing chain end
Actinomycin D	binds to DNA and blocks the movement of RNA polymerase (prevents RNA synthesis)
<i>Acting on eucaryotes but not bacteria</i>	
Cycloheximide	blocks the translocation reaction on ribosomes (step 3 in Figure 6–66)
Anisomycin	blocks the peptidyl transferase reaction on ribosomes (step 2 in Figure 6–66)
α-Amanitin	blocks mRNA synthesis by binding preferentially to RNA polymerase II

The ribosomes of eucaryotic mitochondria (and chloroplasts) often resemble those of bacteria in their sensitivity to inhibitors. Therefore, some of these antibiotics can have a deleterious effect on human mitochondria.

La régulation de l'expression des gènes



Régulation de l'expression des gènes

- Nous avons vu la **structure** de l'ADN et comment les gènes sont organisés dans le **génom**.
- Nous verrons plus tard comment des mutations peuvent survenir au sein du génome et comment ce génome peut être réparé.
- Nous avons vu comment un génome est **répliqué**.
- Nous savons que certaines parties d'un génome sont **transcrites** et parfois **traduites**.
- Cependant, l'ensemble d'un génome n'est **pas exprimé en continu**.

Il existe des mécanismes de régulation de l'expression des gènes

Procaryotes

- Opéron lactose: système inductible par levée de répression.
- Opéron tryptophane: système de répression traductionnelle de la transcription.

Eucaryotes

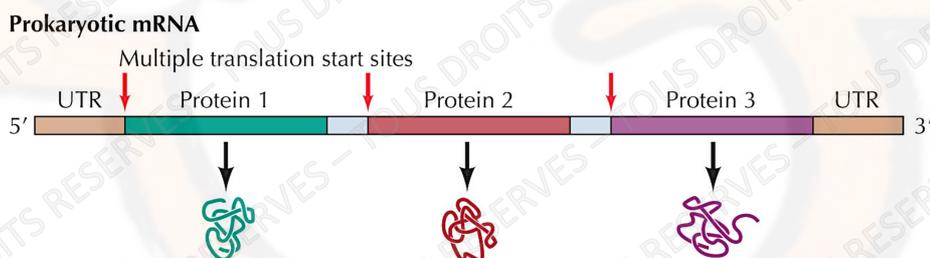
- au niveau chromatinien
- régulation transcriptionnelle
- post-transcriptionnelle
- traductionnelle

Partie à aborder en ayant connaissance des mécanismes transcriptionnels et traductionnels

Rappel:

- 1960: mise en évidence des ARNm

existence de messagers polycistroniques (comportent plusieurs RBS)



- 1961: Jacob, Monod et Lwoff montrent la régulation de l'opéron lactose

Mise en évidence des éléments CIS et éléments TRANS

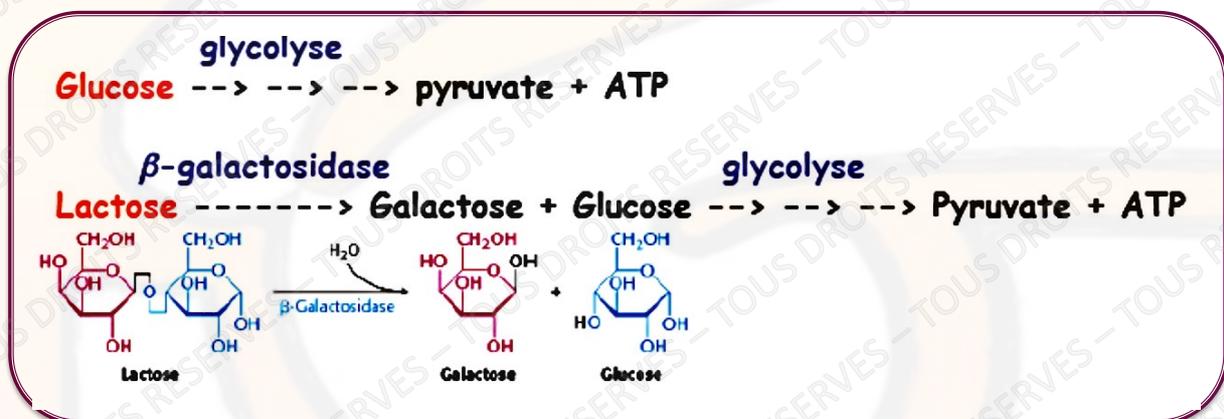
- **éléments CIS**: séquences d'ADN permettant une interaction avec des facteurs protéiques (TRANS) et participant à la régulation de gènes situés à proximité ou à distance. Eléments également nommés "opérateur".
- **éléments TRANS**: facteurs protéiques se fixant à l'ADN et participant à la régulation de l'expression de gènes, par activation ou inhibition de l'activité de l'ARN polymérase.
Ces facteurs peuvent activer ou inhiber l'activité transcriptionnelle du complexe ARN pol.

Régulation de l'expression des gènes - opéron lactose

A

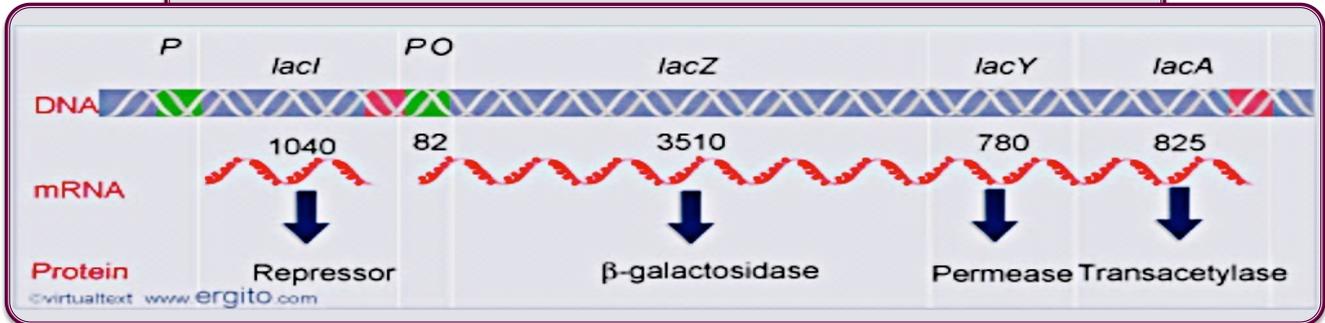
L'Opéron Lactose

- Etudes génétiques et biochimiques d'E. coli
- adaptation à la source de carbone disponible dans le milieu de culture



induction de l'expression de la galactosidase en fonction de l'absence de glucose **ET** de la présence de lactose

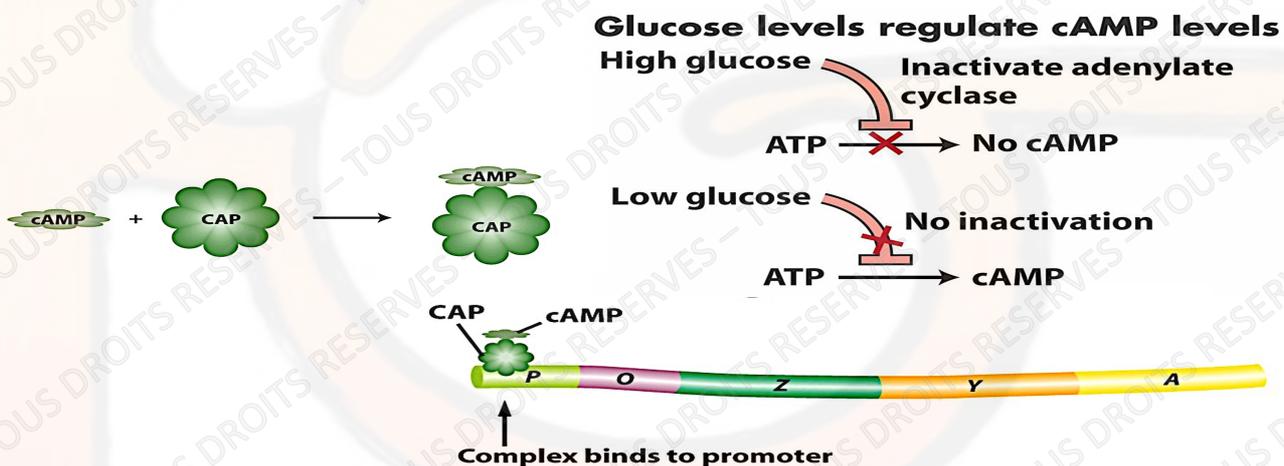
Opéron lactose contient des éléments régulateurs agissant en *cis* et des gènes de structure codant pour des protéines



- P: promoteur
- lacI*: répresseur
- O: Opérateur, séquence sur laquelle se lie le répresseur.
- lacZ*: β-galactosidase
- lacY*: β-galactoside perméase
- lacA*: thiogalactoside transacétylase

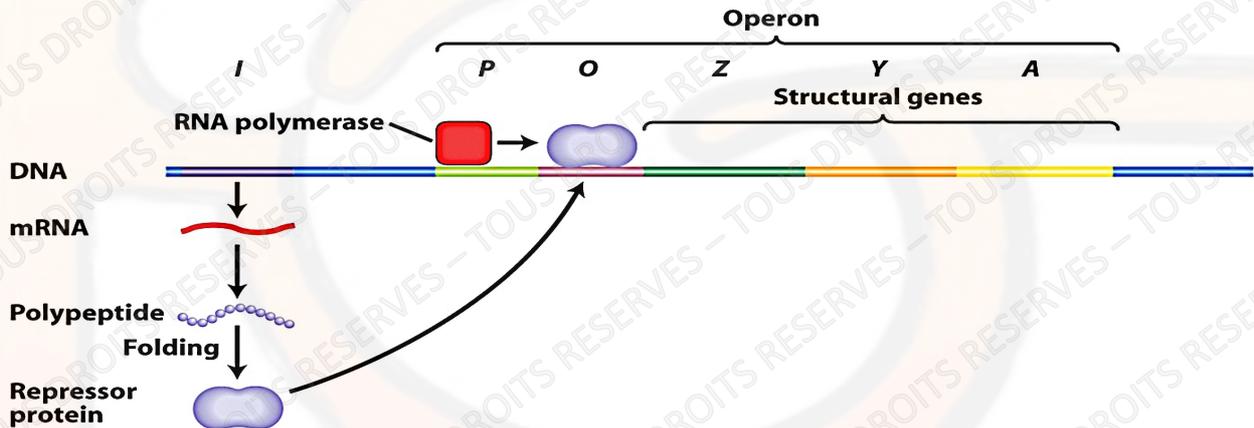
Répression catabolique par le système CAP-AMPc

Le gène *crp* encode pour la protéine CAP: **Catabolite Activator Protein**, dont l'effecteur allostérique est l'AMPc. En présence d'AMPc, CAP se lie à une séquence sur P, en 5' du site de fixation de l'ARN polymérase et active la transcription: **phénomène d'activation par levée de la répression catabolique**.

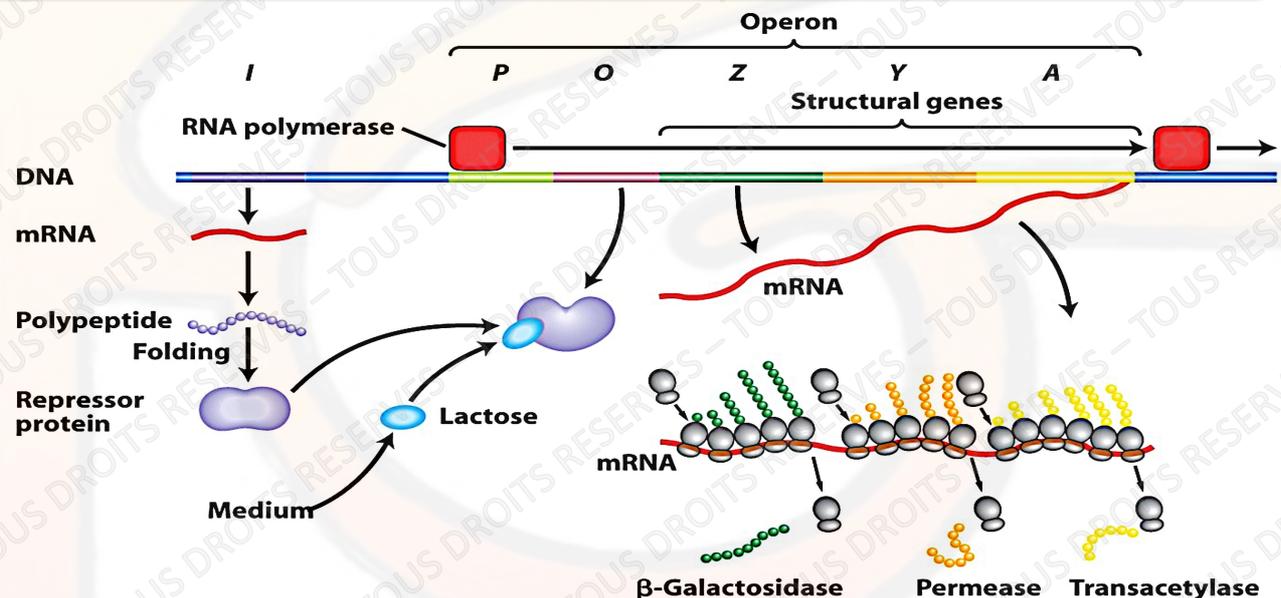


L'opéron lactose est un système inductible basé sur un système de **répression/dérépression**.

En absence de lactose, le répresseur, produit du gène *lacI*, se lie à O et réprime la transcription : **Régulation Négative**



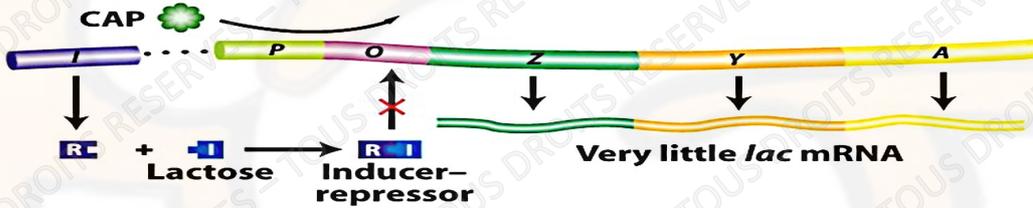
En présence de lactose, le lactose se lie au répresseur qui subit une modification conformationnelle (transition allostérique), et **ne peut plus se lier à O**. Il y a donc une induction de la transcription par le lactose.



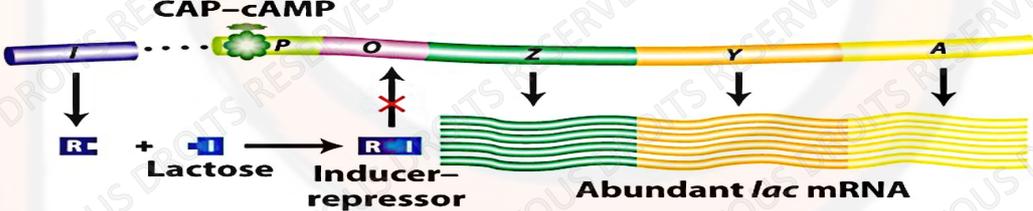
(a) Glucose present (cAMP low); no lactose; no *lac* mRNA



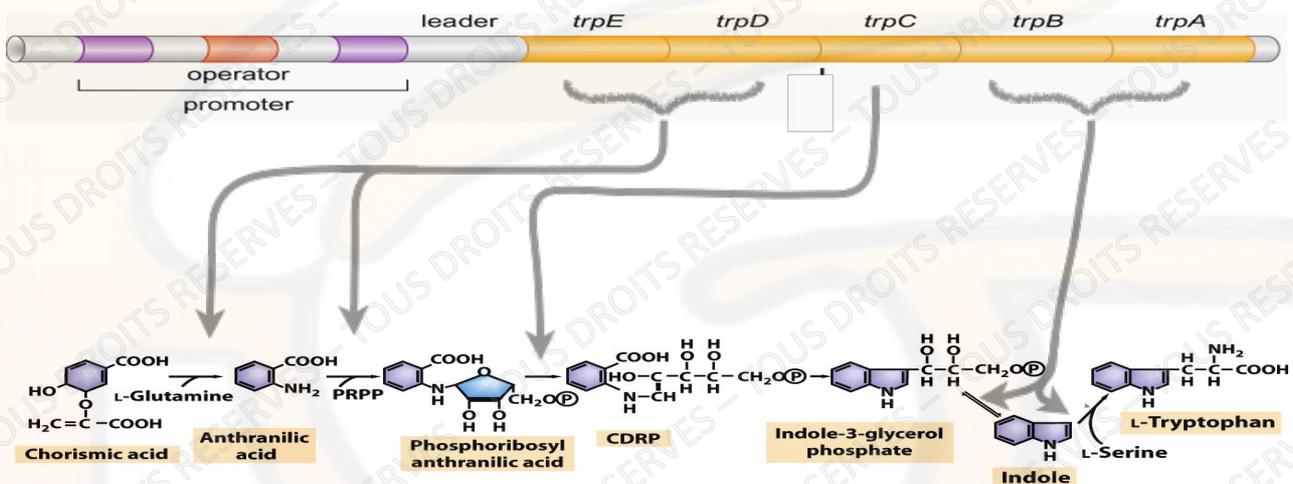
(b) Glucose present (cAMP low); lactose present



(c) No glucose present (cAMP high); lactose present

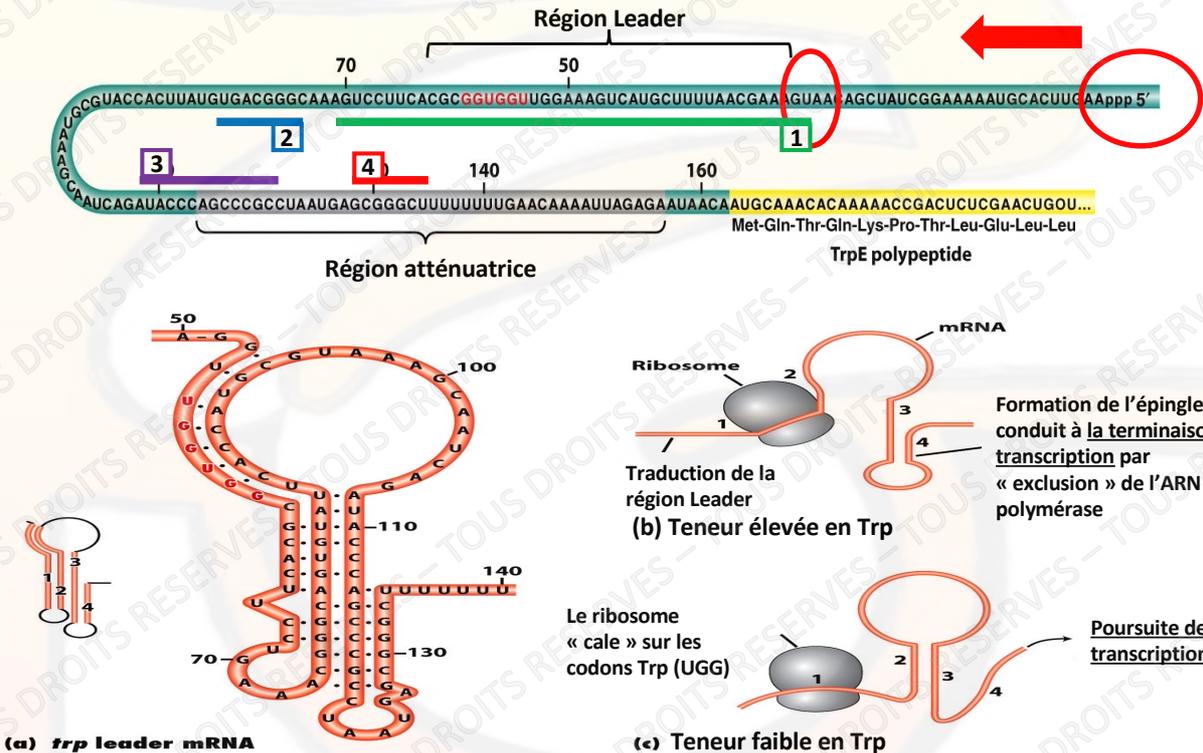


Atténuation génétique



Régulation de l'expression des gènes - opéron Tryptophane

A



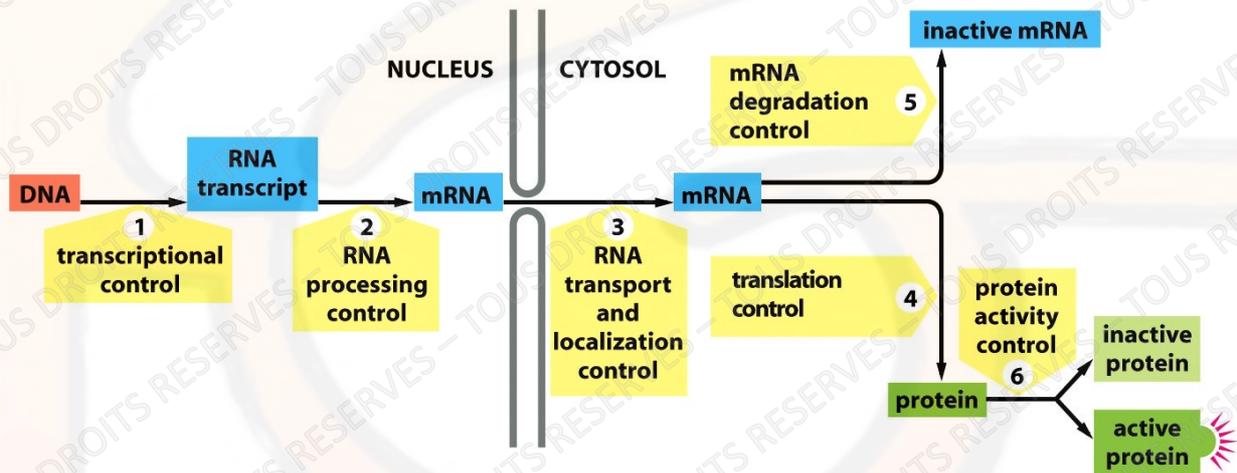
(a) *trp* leader mRNA
 Figure 10-23
 Introduction to Genetic Analysis, Ninth Edition
 © 2008 W. H. Freeman and Company

Régulation de l'expression des gènes - opéron Tryptophane

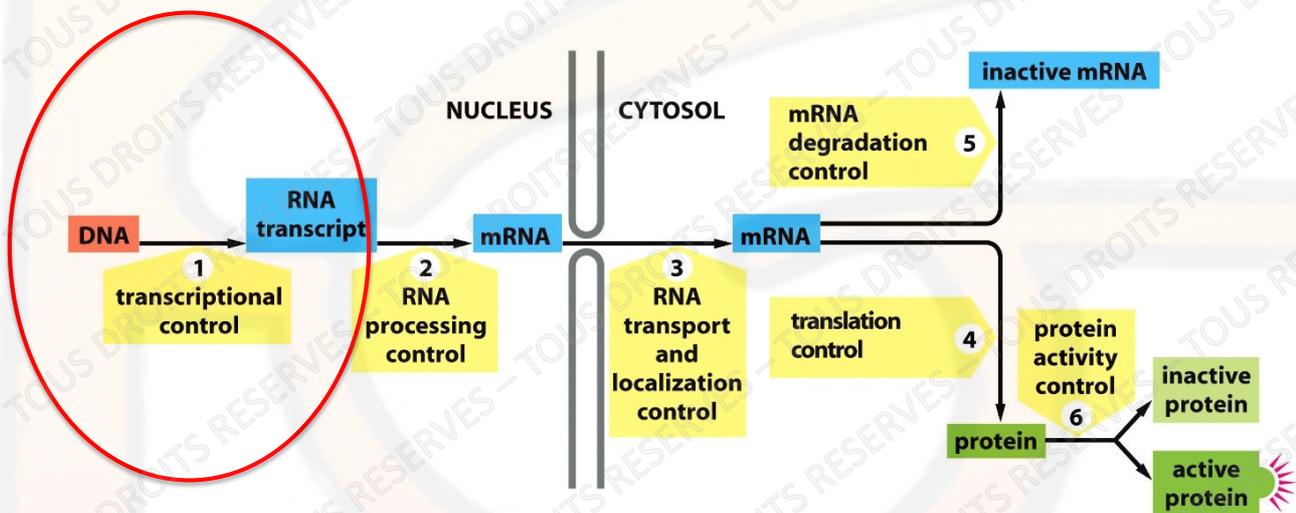
B

- (a) *trp* operon Met - Lys - Ala - Ile - Phe - Val - Leu - Lys - Gly - Trp - Trp - Arg - Thr - Ser - Stop
 5' AUG - AAA - GCA - AUU - UUC - GUA - CUG - AAA - GGU - UGG - UGG - CGC - ACU - UCC - UGA 3'
- (b) *phe* operon Met - Lys - His - Ile - Pro - Phe - Phe - Phe - Ala - Phe - Phe - Phe - Thr - Phe - Pro - Stop
 5' AUG - AAA - CAC - AUA - CCG - UUU - UUU - UUC - GCA - UUC - UUU - UUU - ACC - UUC - CCC - UGA 3'
- (c) *his* operon Met - Thr - Arg - Val - Gln - Phe - Lys - His - His - His - His - His - His - Pro - Asp
 5' AUG - ACA - CGC - GUU - CAA - UUU - AAA - CAC - CAC - CAU - CAU - CAC - CAU - CAU - CCU - GAC 3'

La régulation de l'expression d'un gène eucaryote est réalisée à des niveaux variés...



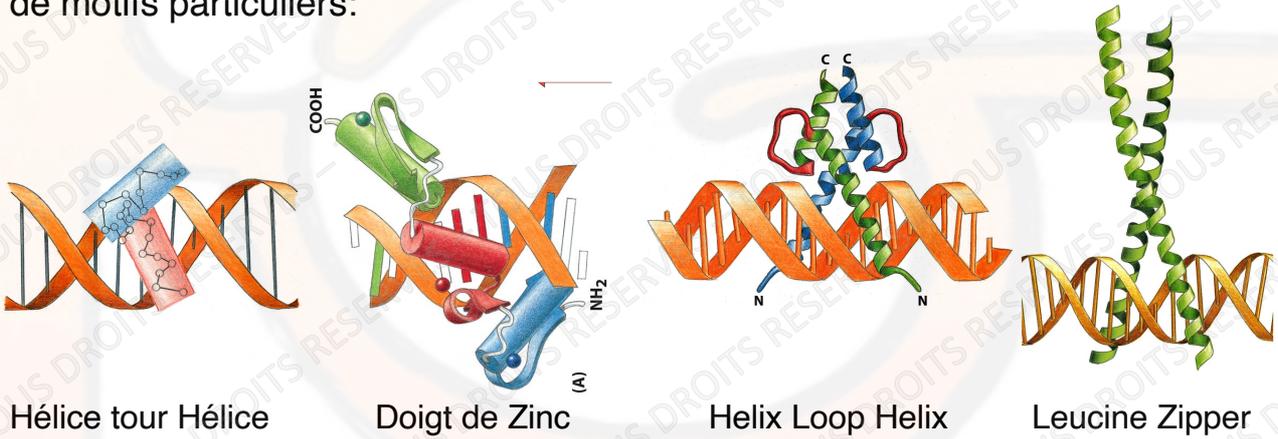
Régulation transcriptionnelle



Notion de commutateurs

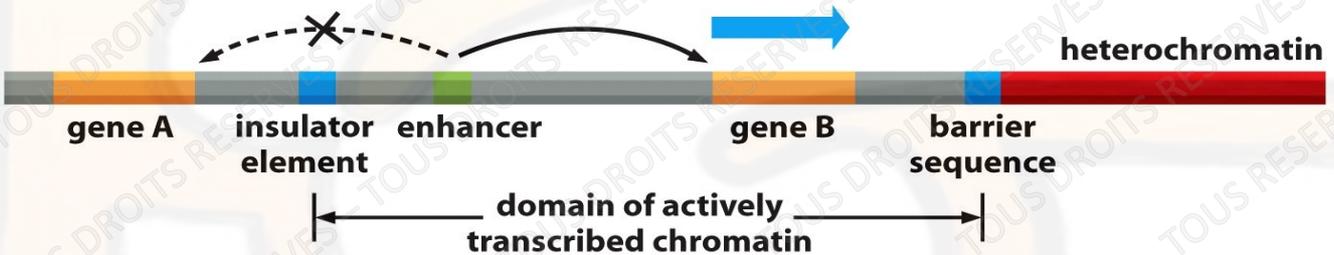
- simples ou complexes
- constitués de séquences situées sur le génome et de protéines pouvant lier ces séquences

protéines liant l'ADN, agissent souvent en dimères et portant 4 grands types de motifs particuliers:



La régulation transcriptionnelle

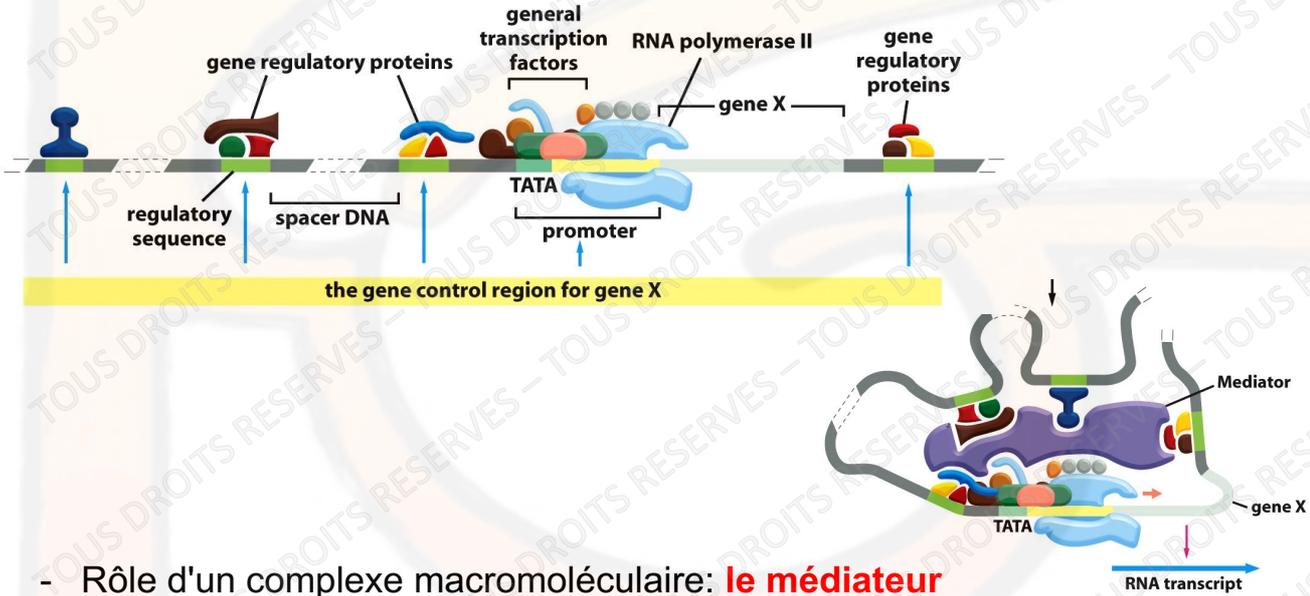
Différents types de séquences régulatrices participent à la régulation de l'expression des gènes eucaryotes



Régulation transcriptionnelle Eucaryote

A B

- l'ARN polymérase II est "incapable" d'initier la transcription à elle seule:
- Complexe d'Initiation de la Transcription (Facteurs Généraux de la transcription)
- Un unique promoteur peut être contrôlé par une infinité de séquences



- Rôle d'un complexe macromoléculaire: **le médiateur**

Régulation de la transcription par méthylation de l'ADN

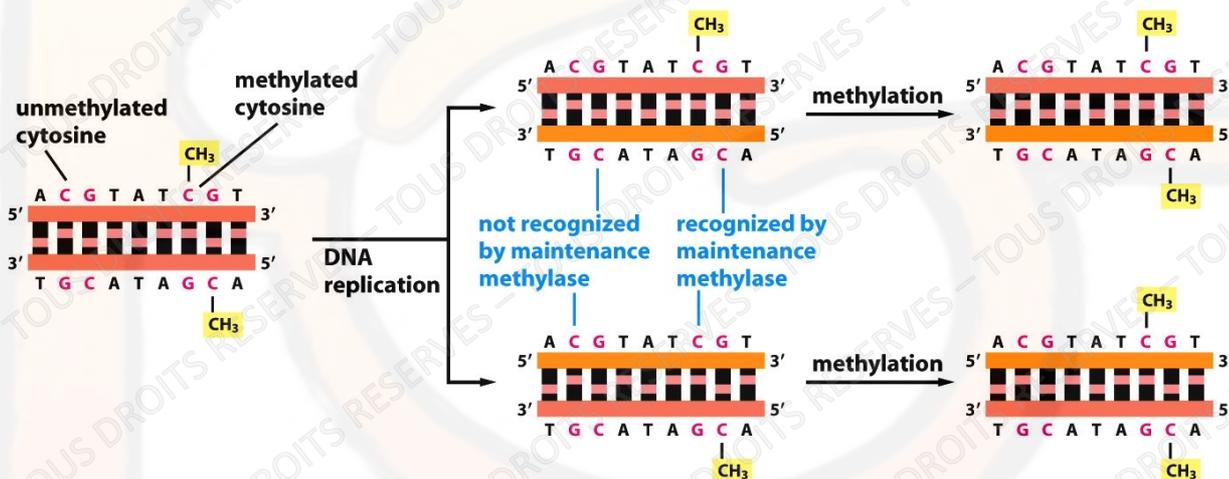
A

méthylation de l'ADN

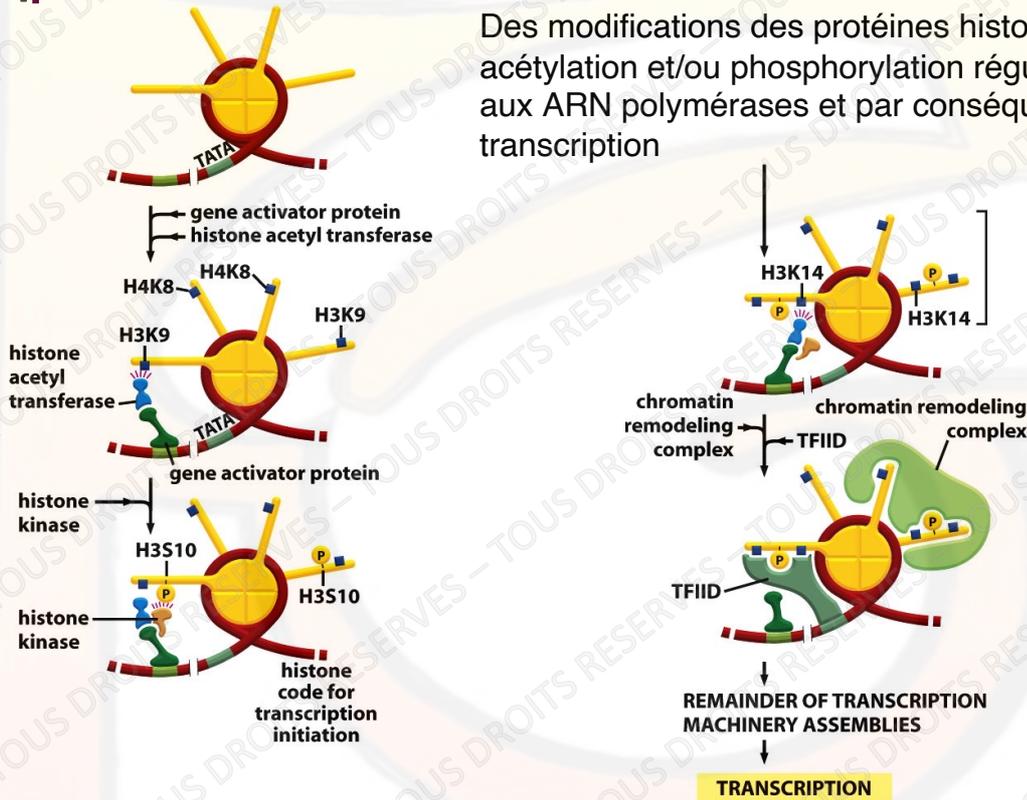
Réalisée par des ADN méthylases (déjà rencontrées précédemment)

Méthylation des îlots CG (ou CpG, répétition de CG de 300 à 3000 pb)

Notion d'empreinte.



Des modifications des protéines histones, par acétylation et/ou phosphorylation régulent l'accessibilité aux ARN polymérase et par conséquent la transcription



Existence d'un "code" de modification des Histones

Euchromatin (active/open)	Heterochromatin (inactive/condensed)
<p>H3 ARTKQTARKSTGGKAPRKQL Ac P Ac 9 10 14</p> <p>H3 ARTKQTARKSTGGKAPRKQL Ac Me 4 14</p> <p>H4 SGRGKGGKGLGKGGAKRHRK Ac Me 3 5</p>	<p>H3 ARTKQTARKSTGGKAPRKQL P P 10</p> <p>CENP-A MGPRRRSRKPEAPRRRSPSP 7</p> <p>H3 ARTKQTARKSTGGKAPRKQL Me Ac 9 12</p> <p>H4 SGRGKGGKGLGKGGAKRHRK 12</p>

Exemple d'un mécanisme global de régulation de l'expression de gènes:

L'Inactivation du Chromosome X



L'Inactivation du Chromosome X

Le chromosome Y est beaucoup plus petit que le chromosome X, il contient beaucoup moins de gènes et est principalement responsable de la "détermination mâle".

X Chromosome



Duchenne muscular dystrophy

Melanoma

X-inactivation center

X-linked severe combined immunodeficiency (SCID)

Colorblindness

Hemophilia

Y Chromosome



Testis-determining factor

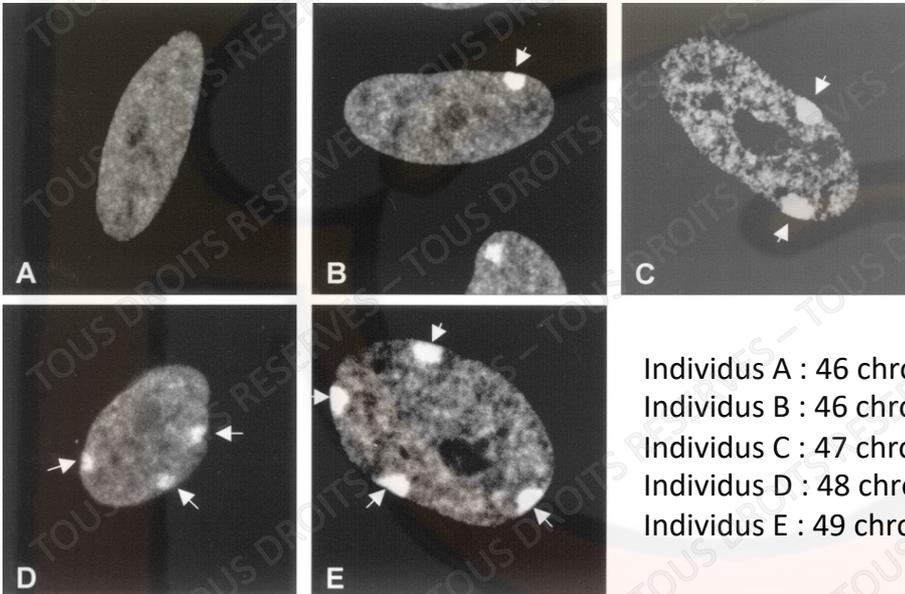


Le corpuscule de Barr

A B

Observation d'une hétérochromatine sexuelle dans les cellules de mammifère de sexe femelle.
Dr. Murray Barr, 1948.

Association à un mécanisme de "comptage" des chromosomes X.



Individus A : 46 chromosomes dont XY, 0 cb;
Individus B : 46 chromosomes dont XX, 1 cb;
Individus C : 47 chromosomes dont XXX, 2 cb;
Individus D : 48 chromosomes dont XXXX, 3 cb;
Individus E : 49 chromosomes dont XXXXX, 4 cb



Le phénomène de compensation de dosage génique

A

L'inactivation du chromosome X chez les mammifères:

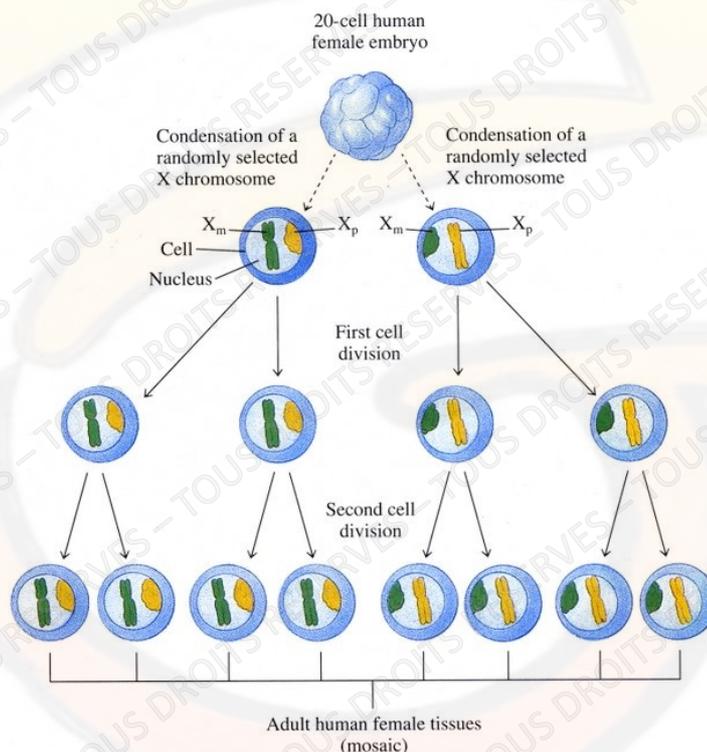
1. Barr observe la chromatine condensée dans les cellules nerveuses de chats femelles. Exclusivement dans les cellules somatiques.
2. Ohno propose que le corpuscule de Barr est un chromosome X condensé
3. Le nombre de corpuscules est toujours égal à :

"Nbre de chromosomes X -1"

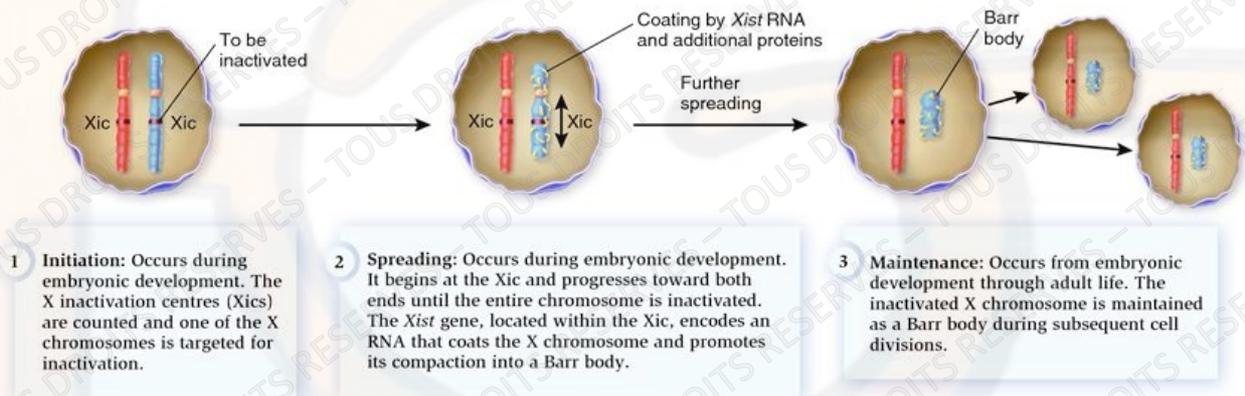
1962, Mary Lyon (Généticienne de la Souris), explique le phénomène de compensation de dosage, qui porte aujourd'hui le nom de Lyonisation):

1. Le corpuscule de Barr est un chromosome X Inactivé. Ses gènes sont inactivés, en grande partie.
2. Le Chromosome inactif peut être d'origine maternelle ou paternelle.
3. Le chromosome inactif tarde à se répliquer.
4. L'inactivation du chromosome X a lieu aux alentours du 16^{ème} jour de développement chez l'Homme. Il s'agit d'un mécanisme aléatoire.
5. Une fois qu'un chromosome est inactivé dans une cellule ce sera toujours le même chromosome qui sera inactivé dans toutes les cellules dérivées de la cellule d'origine.

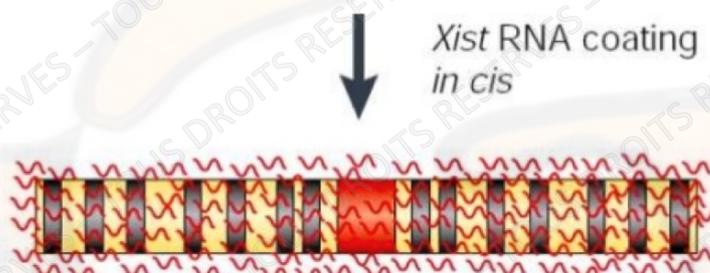
L'inactivation du Chromosome X



- Le chromosome X porte le XIC (X-Inactivation Center) sur son bras q
- XIC contient un gène nommé XIST (Xi-Specific Transcripts)
- XIST exprime un Long ARN Non Codant (lncRNA) de 17kb.
- XIST n'est exprimé que lorsqu'une cellule exprime plus d'un chromosome X



Les molécules d'ARN XIST vont recouvrir le chromosome à inactiver



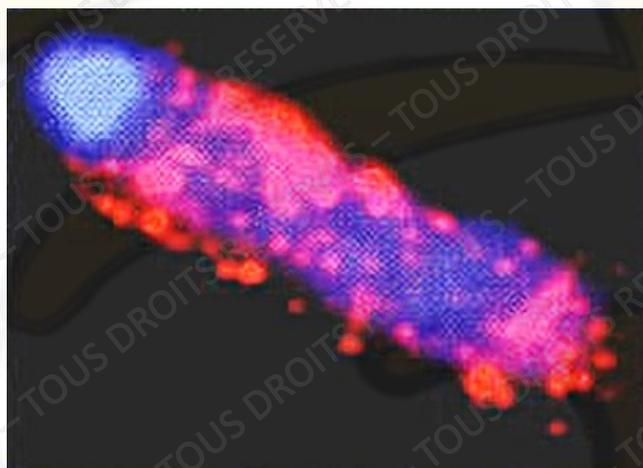
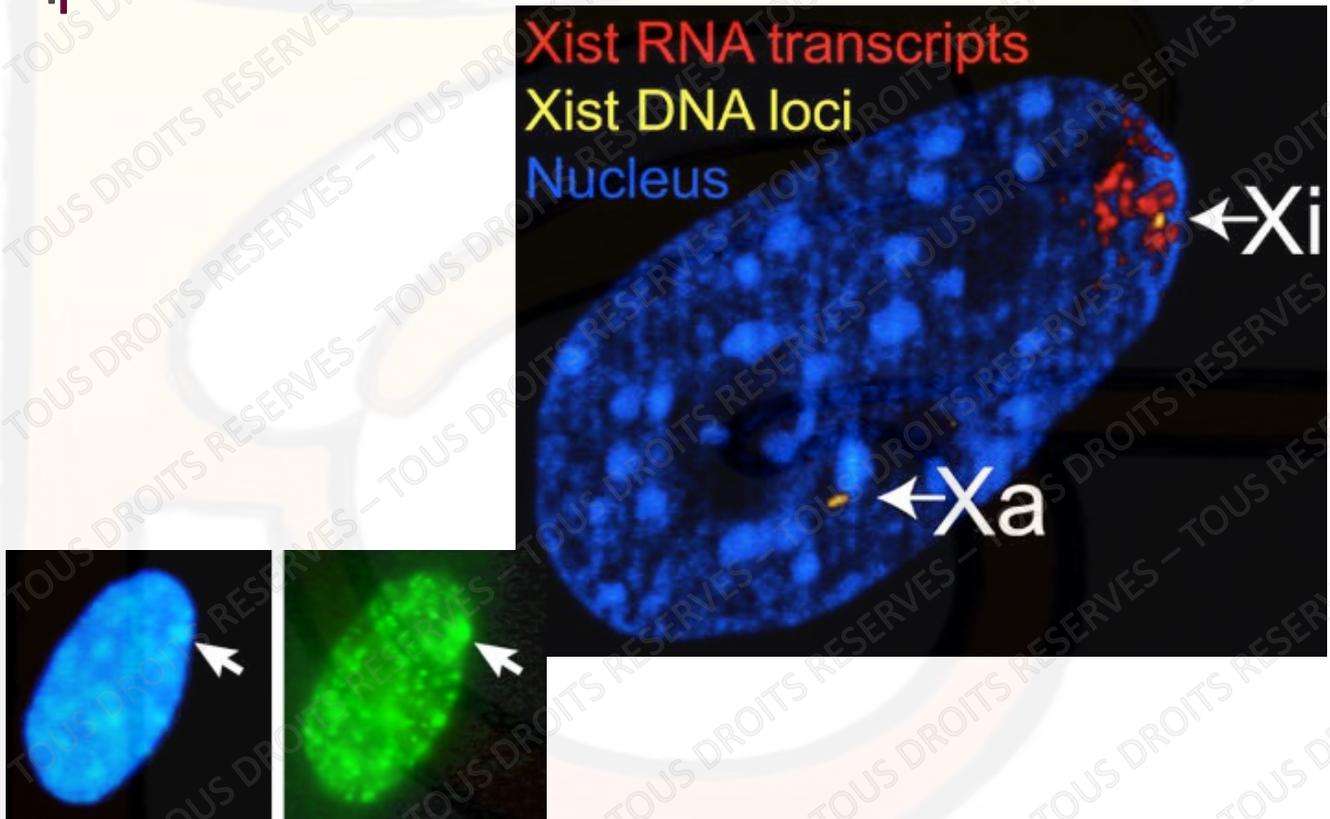


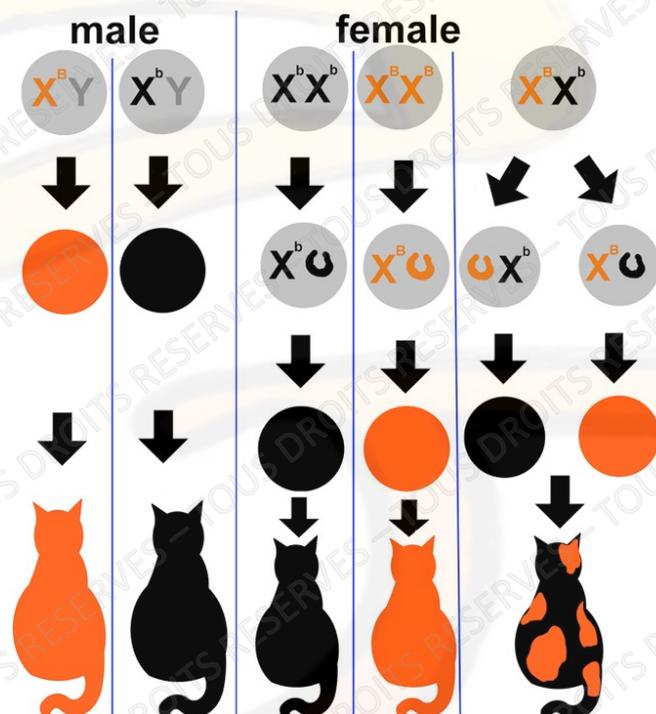
Figure 4 : Association de l'ARN Xist sur le chromosome X inactivé. On observe en fluorescence un chromosome murin métaphasique (bleu) recouvert par l'ARN Xist (rouge).

Source : [Developmental Biology](#).

- L'inactivation débute en un point précis du chromosome X : Xq13.2
- L'inactivation n'a pas lieu sur un chromosome dont la région XIC est absente. Des embryons portant plusieurs X actifs (dû à la perte d'un XIC) ne survivent pas.
- XIC est aussi impliqué dans le mécanisme de comptage
- Le X inactif (Xi) exprime le gène XIST. Le transcrit XIST n'est jamais traduit mais il est maturé comme un ARNm.
- Les histones sur Xi sont sous-acétylées
- La région 5' du gène XIST sur Xa est très méthylée alors que cette région 5' est peu méthylée sur Xi.

L'Inactivation du Chromosome X

Robe "calico" ou "écaille de tortue" (inactivation aléatoire d'un des deux chromosomes X chez la femelle au cours du développement embryonnaire)



Un autre exemple de régulation épigénétique : Empreinte parentale



L'empreinte génomique parentale

A

Exemples de mise en évidence d'expression monoalléliques : phénotypes différents selon l'origine des allèles malgré un génome identique

Bardot: croisement étalon X ânesse

Mulet: croisement âne X jument



Bardot



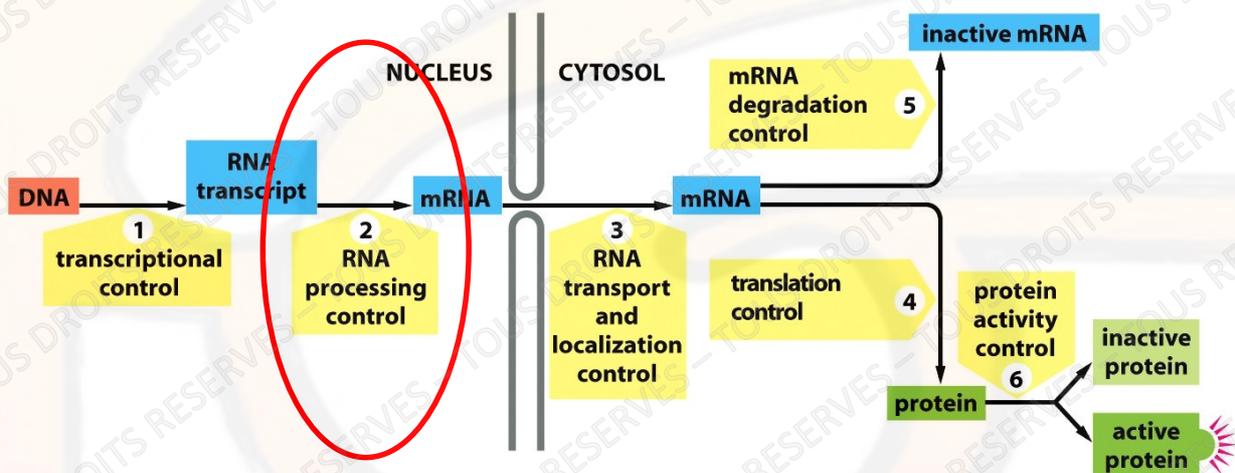
Mulet

Conclusions:

- Certaines pathologies ont pour origine une différence fonctionnelle entre allèle paternel et allèle maternel.
- Différents mécanismes moléculaires peuvent être à l'origine d'une perte d'empreinte (LOI):
 - microdélétion
 - Disomie uniparentale
 - Anomalie ou mutation affectant un centre d'empreinte
- Un même mécanisme, selon qu'il affecte l'un ou l'autre allèle peut conduire à des pathologies différentes.

Régulation de l'expression des gènes Eucaryotes

Régulation post-transcriptionnelle



Epissage alternatif

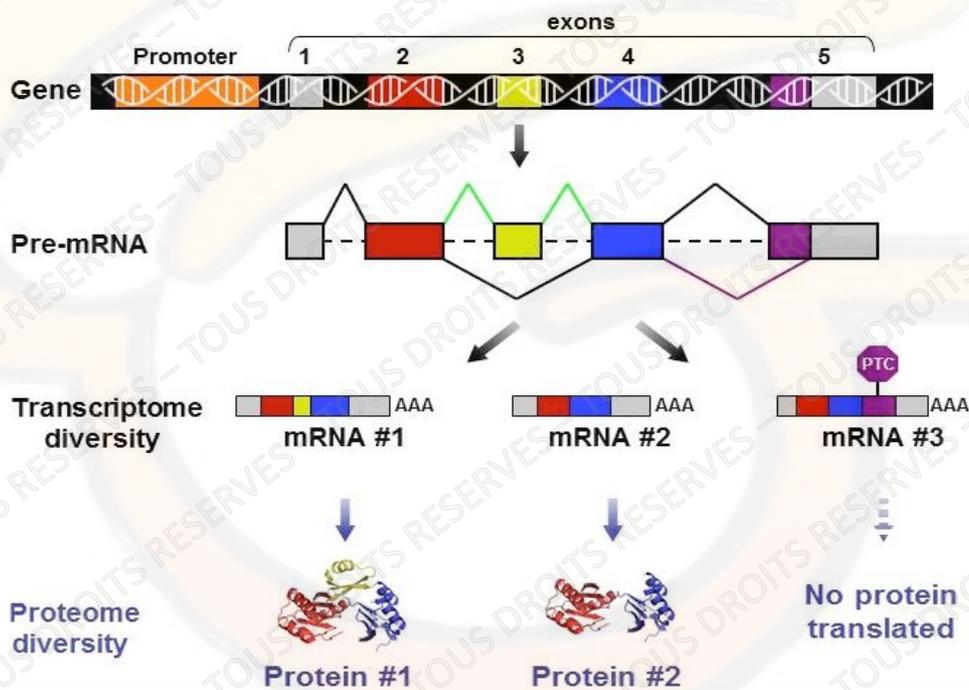


Régulation post-transcriptionnelle - épissage

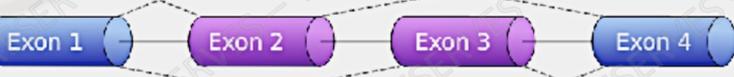
A+

Possibilités d'épissages alternatifs

- expression de transcrits matures différents selon les tissus
- expression de protéines différentes



Conséquences moléculaires d'un épissage alternatif

Exon skipping or cassette exons	
Mutually exclusive exons	
Alternative 5' splice site	
Alternative 3' splice site	
Intron retention	

Régulation post-transcriptionnelle - NMD

Le NMD (rappel)

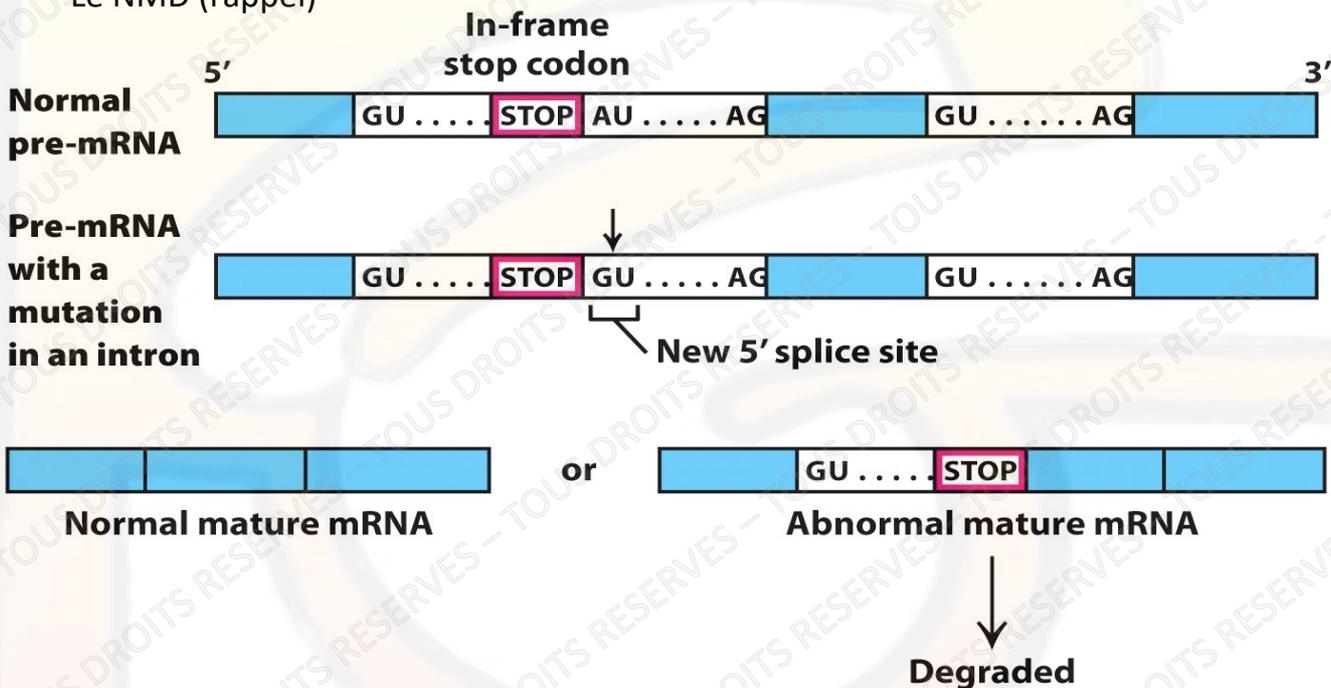
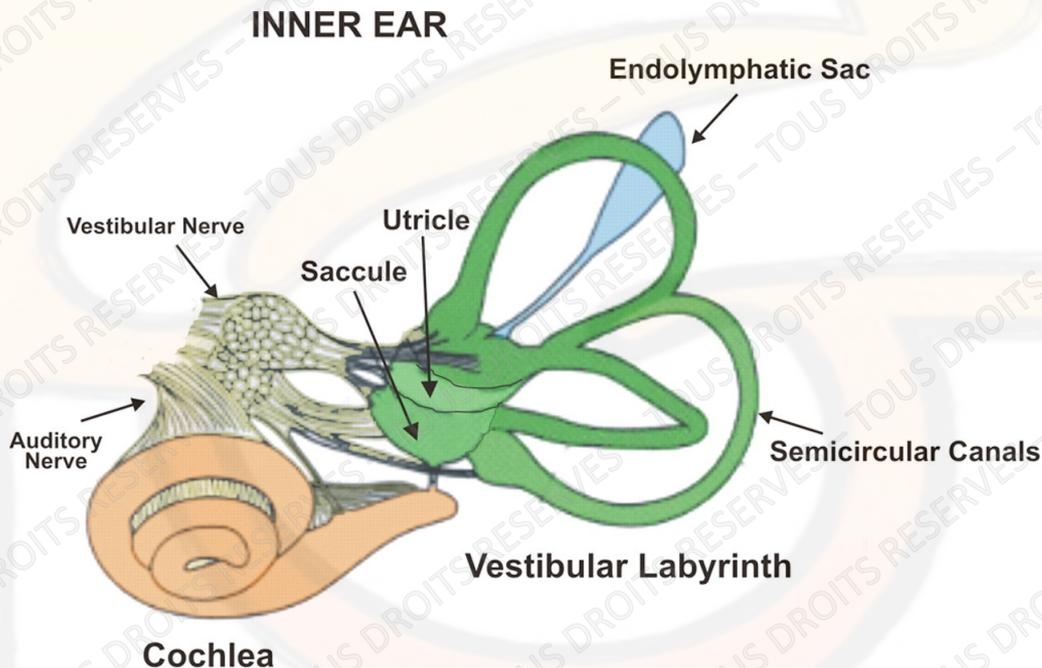


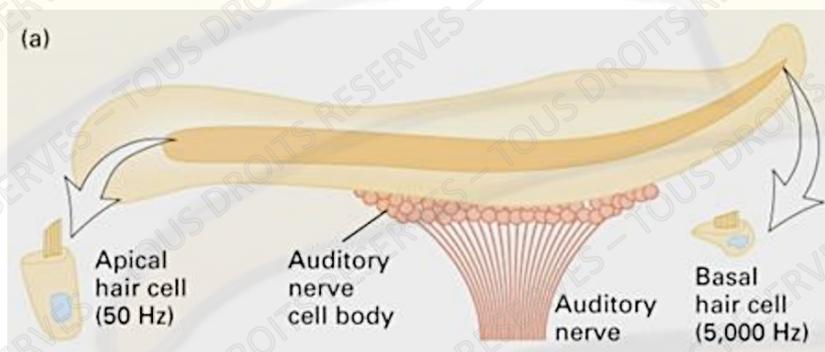
Figure 29.40
Biochemistry, Seventh Edition
© 2012 W. H. Freeman and Company

Epissage alternatif du transcrit du gène SLO



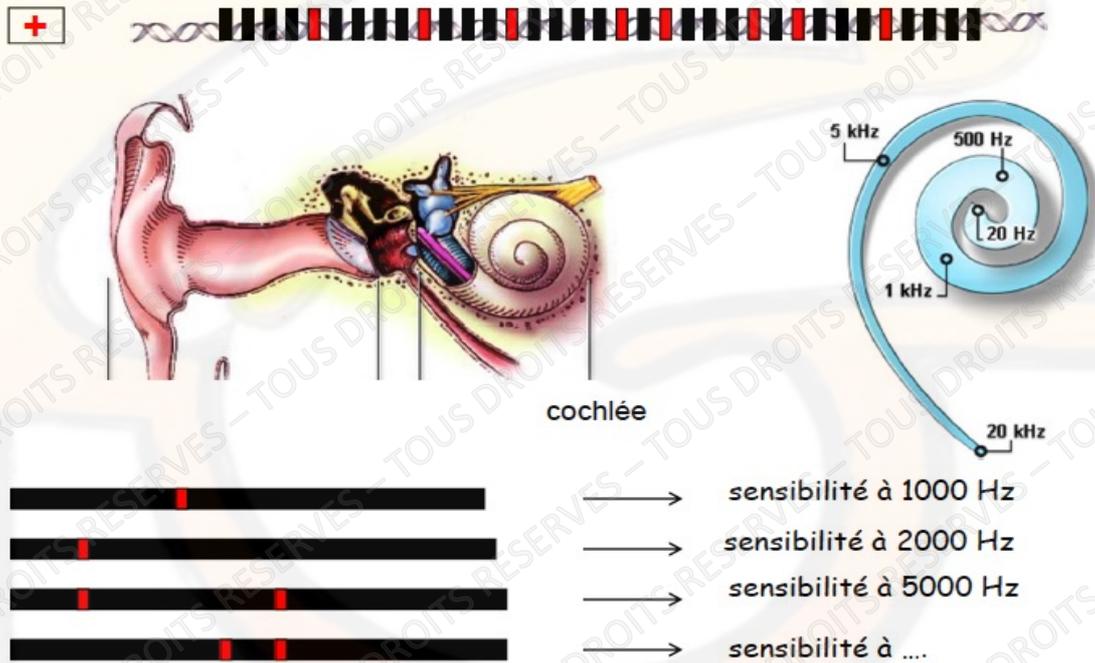
Régulation post-transcriptionnelle-épissage alternatif

Epissage alternatif du transcrit du gène SLO



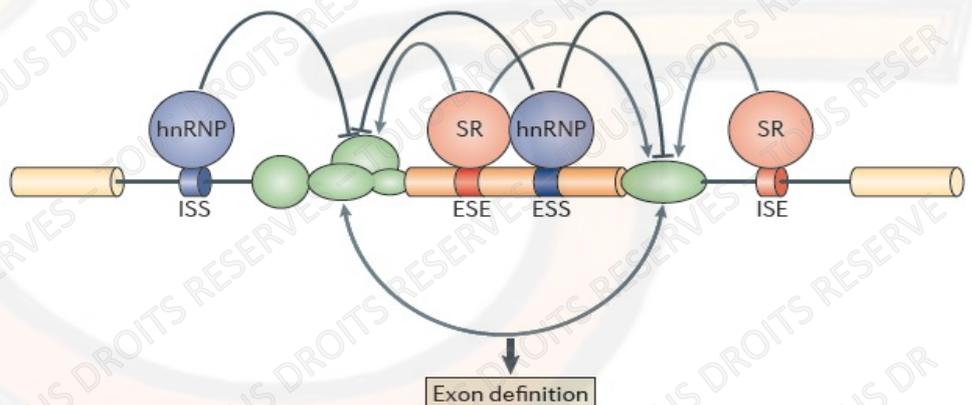
gène *slo* (35 exons dont 8 optionnels) : \approx 500 transcrits différents soit 500 protéines SLO

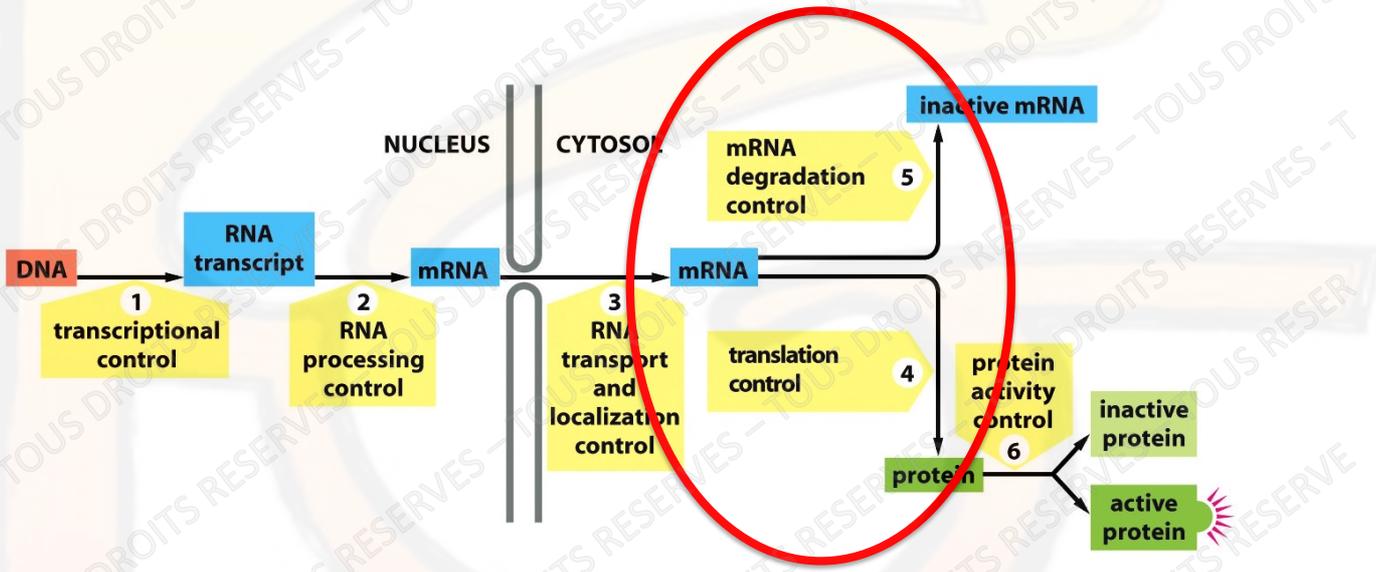
Le gène SLO code pour une protéine membranaire régulant l'entrée et la sortie de potassium dans les cellules pileuses de la cochlée. Ces cellules répondent aux différentes fréquences (Hz) en fonction de la nature de la protéine SLO. La diversité des protéines SLO obtenues grâce à l'épissage alternatif détermine l'étendue de l'audition en moulant la transmission du message nerveux.



- L'épissage d'un transcrit est un phénomène finement régulé.
Un pré-mRNA porte des éléments "CIS" régulateurs de l'épissage:
- ESE: Exonic Splicing Enhancer
 - ESS: Exonic splicing silencer
 - ISS: Intronic Splicing Silencer
 - ISE: Intronic Splicing Enhancer

Reconnus par des protéines de type "SR" et "hnRNP"
Ces protéines vont interagir avec des éléments du spliceosome au niveau des sites 5'et 3' d'épissage.





ARN interférence



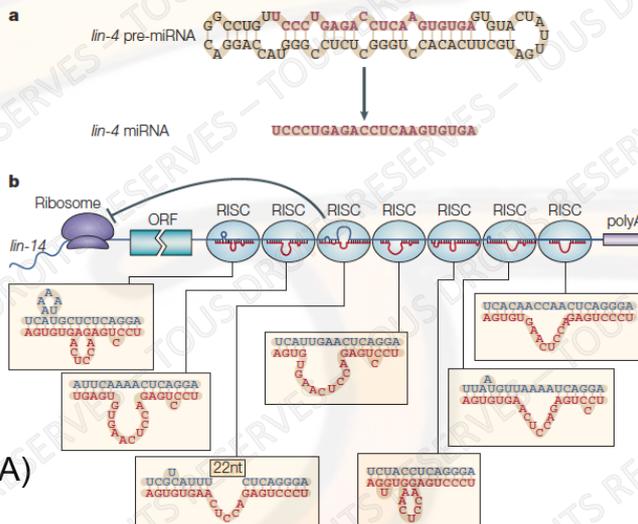


Régulation post-transcriptionnelle - miRNA

A

Les miRNA:

- à l'origine, des stRNA (≈22nt) découverts chez *C.elegans*
- régulent négativement la traduction de séquences cibles
- originaires de diverses régions du génome
- petits ARN simple brin
- La plupart sont localisés dans des régions intergéniques
- plus de 50% sont dans des clusters
- Transcription autonome ou présents dans des introns
- Transcrits par ARN polIII sous la forme de transcrits primaire polycistronique (pri-miRNA)
- plus de 200 par type d'organisme
- cibles multiples, organisation en réseaux complexes de régulation

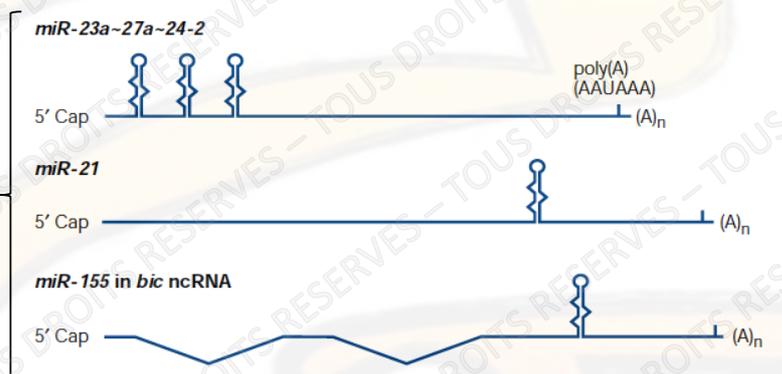


Régulation post-transcriptionnelle - miRNA

B

localisation variée des miRNA

localisation exonique dans des transcrits non-codants



localisation intronique dans des transcrits non-codants



localisation intronique dans des transcrits codants



Régulation post-transcriptionnelle - miRNA

A

Il existe une multitude de miRNA (plus de 300 chez l'Homme).
Régulation spatio-temporelle.
Chaque miRNA peut "cibler" plusieurs transcrits.

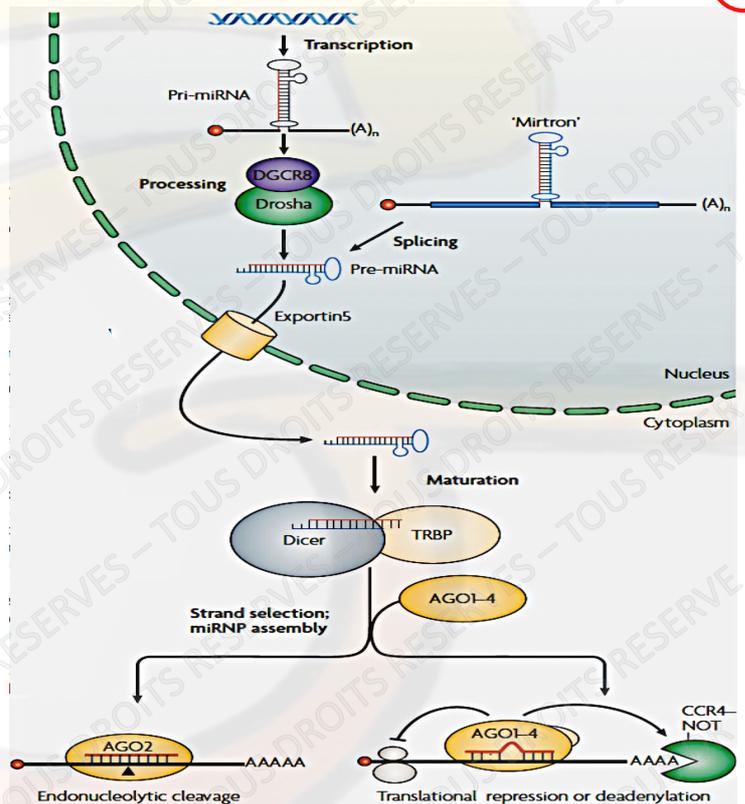
Expression pattern	microRNA
Tissue-specific expression patterns of mammalian microRNAs	
ES-cell specific	miR-296
Expressed in ES cells, but upregulated on differentiation	miR-21 and miR-22
Expressed in both ES cells and various adult tissues	miR-15a, miR-16, miR-19, b, miR-92, miR-93, miR-96, miR-130 and miR-130b
Enriched during mouse brain development	miR-128, miR-19b, miR-9, miR-125b, miR-131, miR-178, miR-124a, miR-266 and miR-103
Enriched in adult brain	miR-9*, miR-125a, miR-125b, miR-128, miR-132, miR-137, miR-139, miR-7, miR-9, miR-124a, miR-124b, miR-135, miR-153, miR-149, miR-183, miR-190 and miR-219
Enriched in lung	miR-18, miR-19a, miR-24, miR-32, miR-130, miR-213, miR-20, miR-141, miR-193 and miR-200b
Enriched in spleen	miR-99a, miR-127, miR-142-a, miR-142-s, miR-151, miR-189 and miR-212
Haemetopoietic tissues	miR-181, miR-223 and miR-142
Enriched in liver	miR-122a, miR-152, miR-194, miR-199 and miR-215
Enriched in heart	miR-1b, miR-1d, miR-133, miR-206, miR-208 and miR-143
Enriched in kidney	miR-30b, miR-30c, miR-18, miR-20, miR-24, miR-32, miR-141, miR-193 and miR-200b
Ubiquitously expressed	miR-16, miR-26a, miR-27a, miR143a, miR-21, let-7a, miR-7b, miR-30b and miR-30c
Abnormal microRNA expression during tumorigenesis	
Downregulated in chronic lymphocytic leukaemias	miR-15 and miR-16
Downregulated in lung cancer cell lines	miR-26a and miR-99a
Downregulated in colon cancers	miR143/miR-145 cluster
Upregulated in Burkitt lymphoma	miR-155

Régulation post-transcriptionnelle - miRNA

A+

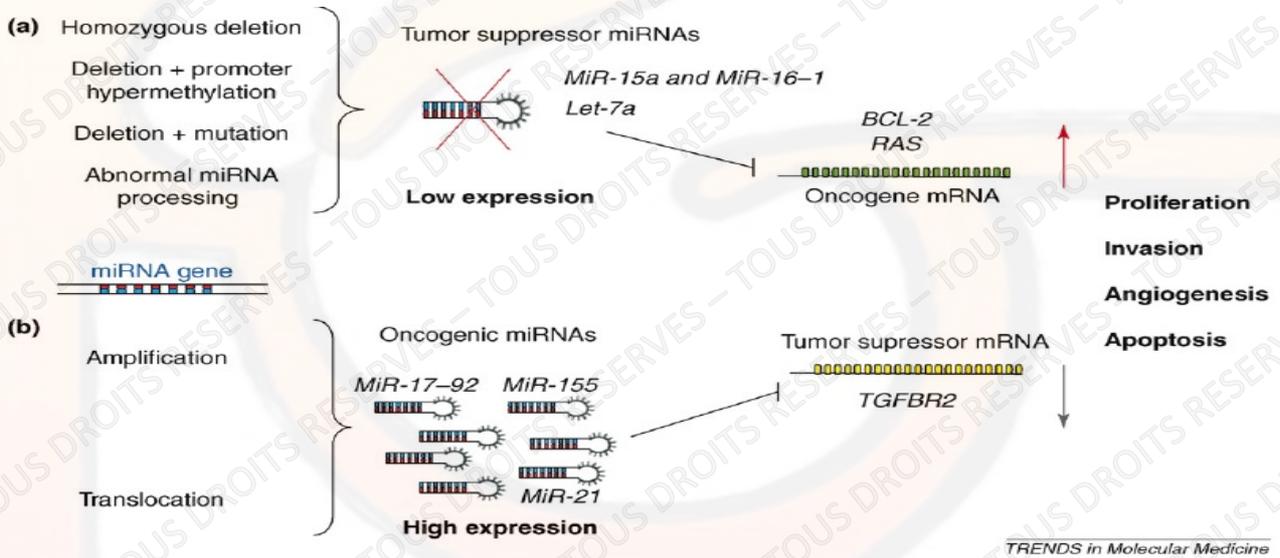
maturation des miRNA

- Production d'un Pri-miRNA
- Processing par Drosha en Pre-miRNA
- Export par Exportin5
- Maturation cytoplasmique par Dicer et sélection du brin spécifique
- Assemblage avec les protéines "argonaute" et formation du complexe RISC (RNA Induced silencing Complex).

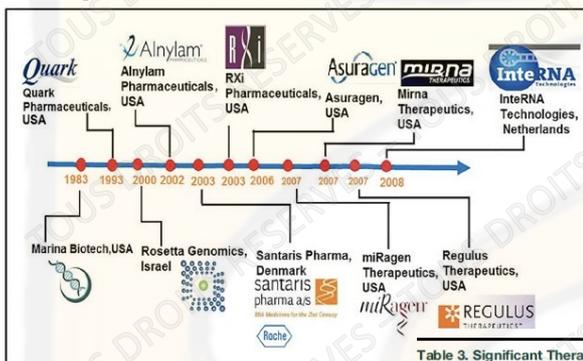


Chez l'Homme:

- Une diminution d'expression des miRNA est souvent observée dans les cellules cancéreuses. Ces miRNA pourraient donc être des suppresseurs de tumeurs
- plus de la moitié des miRNA identifiés chez l'Homme sont localisés dans des régions du génome altérées dans les cancers



Thérapeutique et miRNA



Molecular Therapy
Nucleic Acids
 Review

Therapeutic miRNA and siRNA: Moving from Bench to Clinic as Next Generation Medicine

Chiranjib Chakraborty,^{1,2,4} Ashish Ranjan Sharma,^{2,4} Garima Sharma,² C. George Priya Doss,³ and Sang-Soo Lee²

Table 3. Significant Therapeutics miRNA that Are in the Development Phase, Their Indication, and Their Biopharmaceutical Company

Therapeutic miRNAs	Indication	Biopharmaceutical Company	Remarks
Miravirsen	the indication of the hepatitis C virus (HCV) infection	Santaris Pharma	phase IIa clinical trial
MRX34	for the treatment of a variety of cancers such as colon cancer, non-small-cell lung cancer (NSCLC), hepatocellular carcinoma, cervical cancer, ovarian cancer, etc.	Mirna Therapeutics	phase 1 clinical trial halted because of immune responses
RG-101	for the treatment of HCV	Regulus Therapeutics	an owned GalNAc-conjugated anti-miR
RG-012	for the treatment of Alport syndrome	Regulus Therapeutics	in the pipeline to initiate clinical trial phase II
MGN-1374	for the treatment of post-myocardial infarction remodeling	miRagen Therapeutics	targets miR-15 and miR-195; it is in the preclinical stage
MGN-2677	for the treatment of vascular disease	miRagen Therapeutics	targets miR-143/145; it is in the pipeline
MGN-4220	for the treatment of cardiac fibrosis	miRagen Therapeutics	targets miR-29; it is in the pipeline
MGN-4893	for the treatment of disorders like abnormal red blood cell production such as polycythemia vera	miRagen Therapeutics	targets miR-451; it is in the pipeline
MGN-5804	for the treatment of cardiometabolic disease	miRagen Therapeutics	targets miR-378; it is in the pipeline
MGN-6114	for the treatment of peripheral arterial disease	miRagen Therapeutics	targets miR-92; it is in the pipeline
MGN-9103	for the treatment of chronic heart failure	miRagen Therapeutics	targets miR-208; it is in the pipeline



Objectifs:

- comprendre la génomique structurale d'un type cellulaire (organisation des gènes dans un génome)
- Avoir accès aux modification de cette organisation (déplacement, mutations, amplification)
- comprendre les informations fournies par les analyses transcriptomiques (codant ET non codant)
- Accéder au protéome et au signalosome

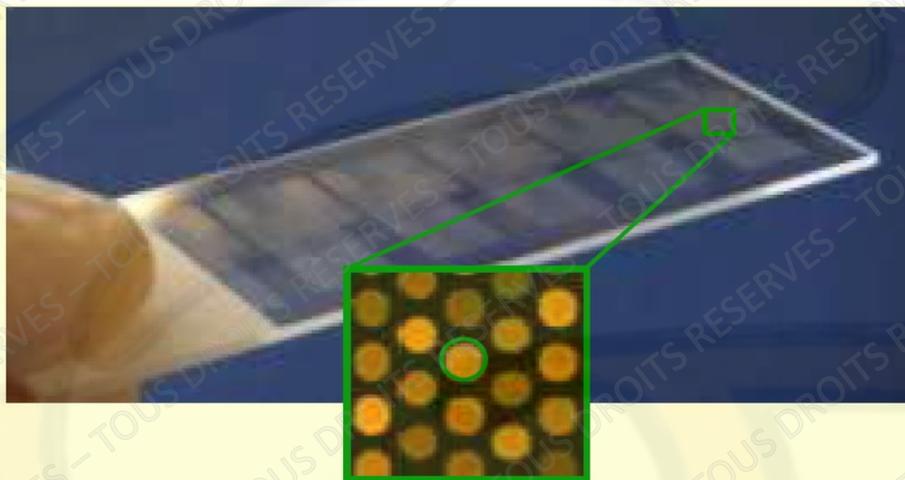
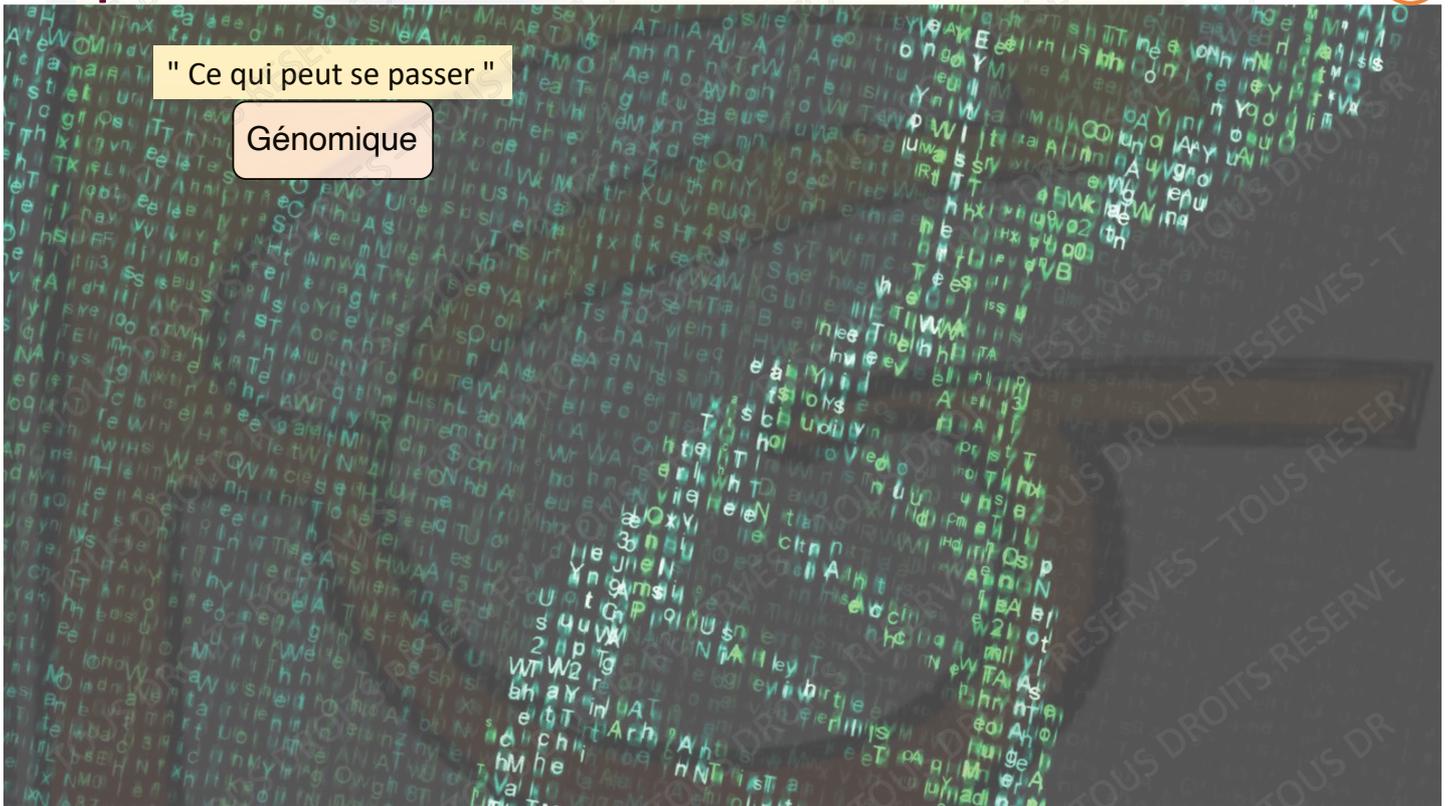
Ces objectifs vont bien au delà de l'analyse du génome obtenue en 2000. Il est rapidement apparu que l'accès au simple génome était insuffisant pour accéder à la compréhension du vivant dans son ensemble.

L'objectif à notre échelle est de pouvoir acquérir une signature moléculaire pour chaque pathologie ou pour chaque étape d'une pathologie donnée.

- Amélioration du diagnostique
- Amélioration de l'approche thérapeutique
- Identification de nouvelles cibles thérapeutiques

Les méthodes globales

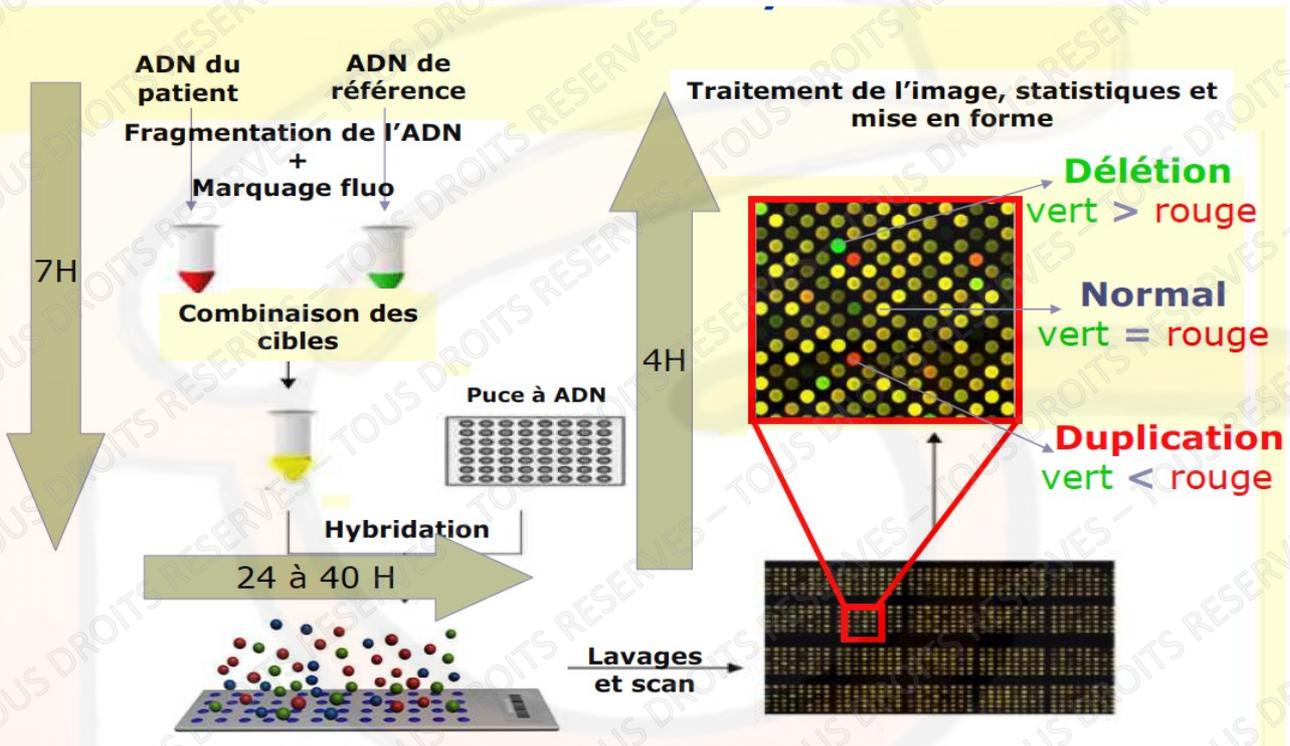


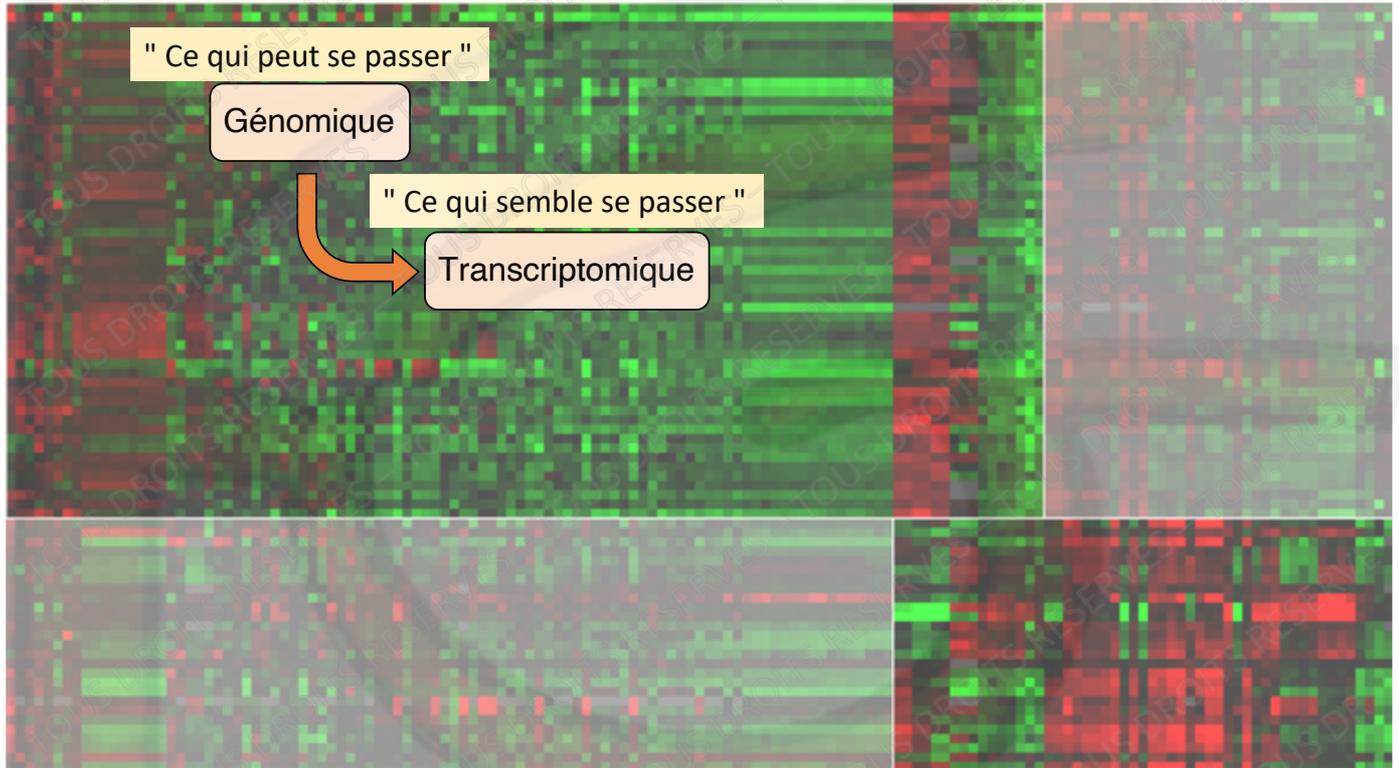


1 spot = 1 séquence d'ADN présente en multiples copies.

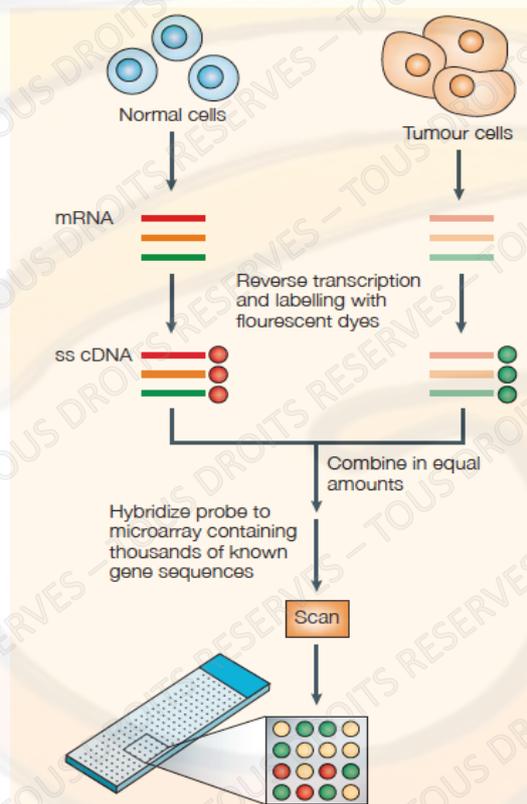
1 bonne sonde = séquence d'ADN unique présente dans le génome humain.

Plus il y a de sondes différentes sur la puce à ADN, meilleure sera la résolution de la puce (+ d'informations).





DNA arrays



- **Macroarrays** (= *spotted macroarrays*) : **membranes de nylon** de 100 à 200 cm² - 200 à 5000 cibles produits PCR – marquage, en général, radioactif (souvent, a-XTP³³)
1 condition / membrane

- **Microarrays** (= *spotted microarrays*) : **lames de verre** de quelques cm²

- produits PCR = **DNA-microarrays**

- **oligonucléotides** = **oligo-microarrays** -10 à 100 bases (le plus souvent autour de 50-70 nucléotides)

Densité de 5 000-30 000 cibles par lame

Marquage en général directement ou indirectement fluorescent (e.g. Cy3, Cy5)

1 ou 2 conditions / microarray

- **Oligochips** (= **GeneChips** = **biopuces**, Affymetrix, NimbleGen) : **support de verre** très petit (1 cm²) sur lequel sont synthétisés *in situ* des oligonucléotides [le plus souvent d'environ 25 bases (photolithographie) à 60 à 100 bases (*ink jet*)]

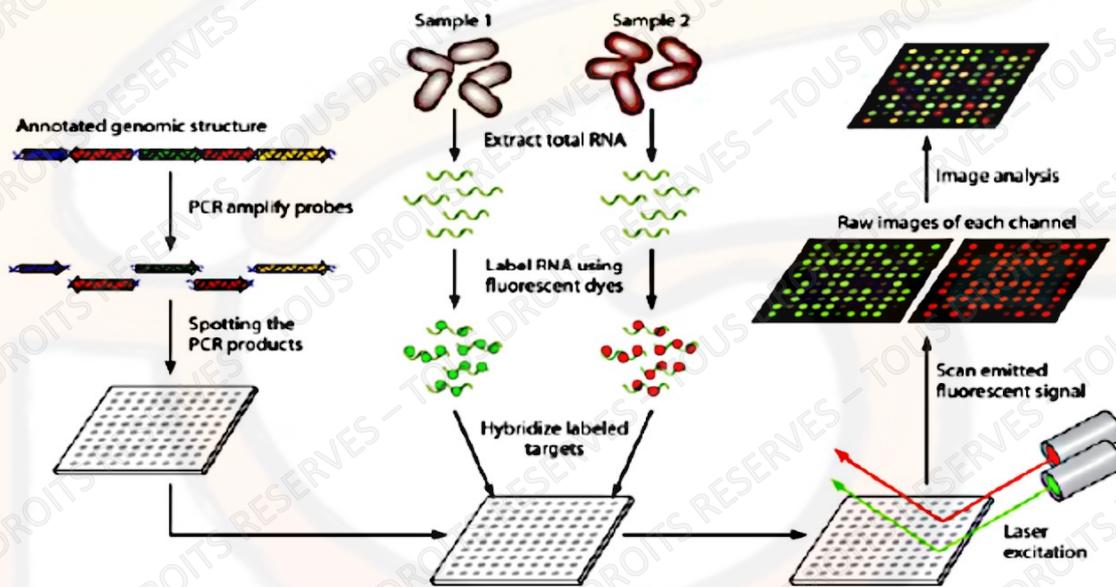
Densité de 15 000 – 10⁶ cibles par lame

Souvent 1 condition / chip



Les DNA arrays

Hybridation compétitive sur *microarrays*

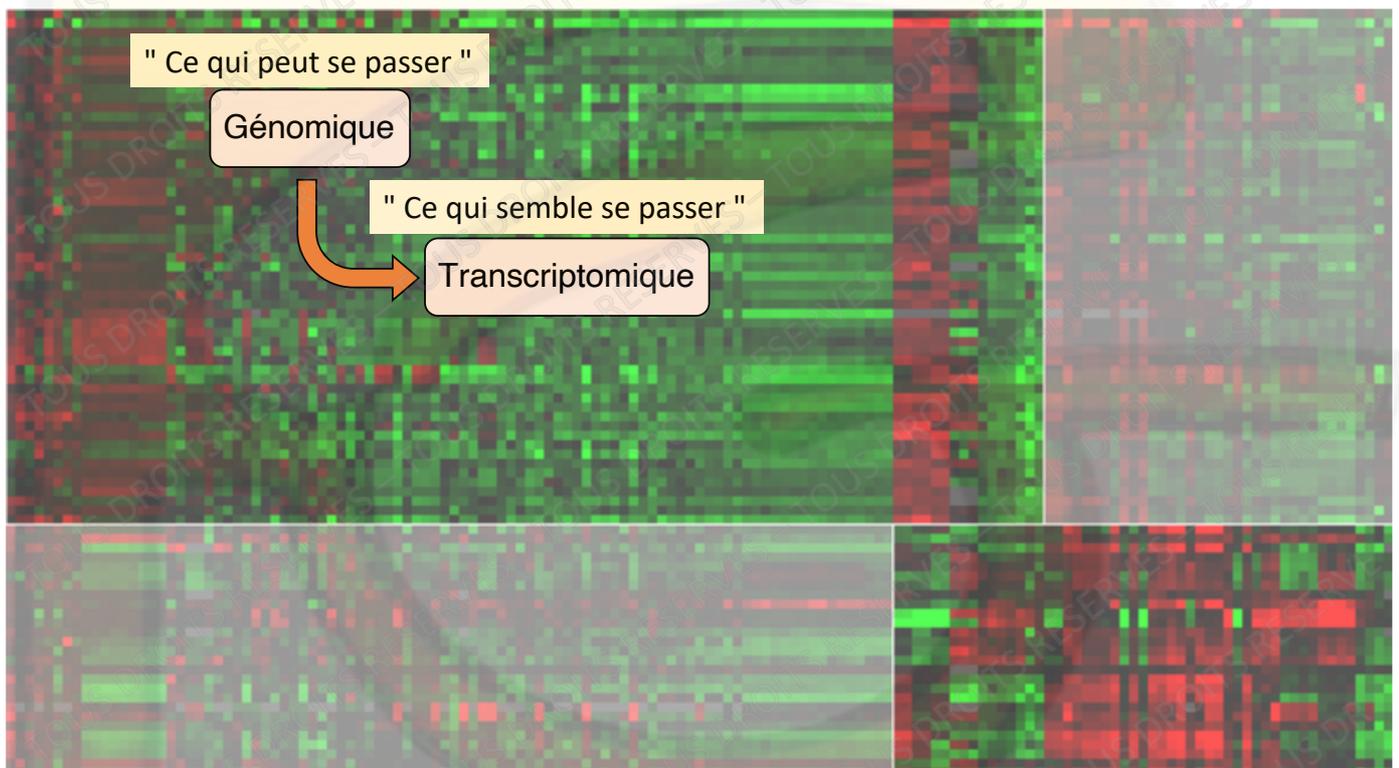
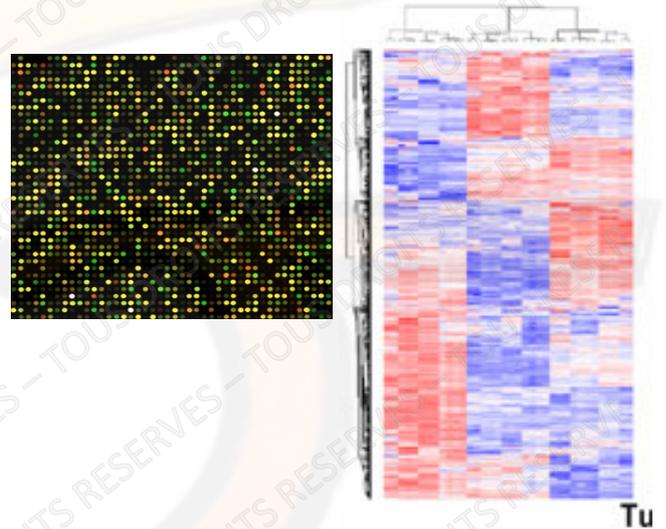


Miller & Tang, *Clin. Microbiol. Rev* (2009)



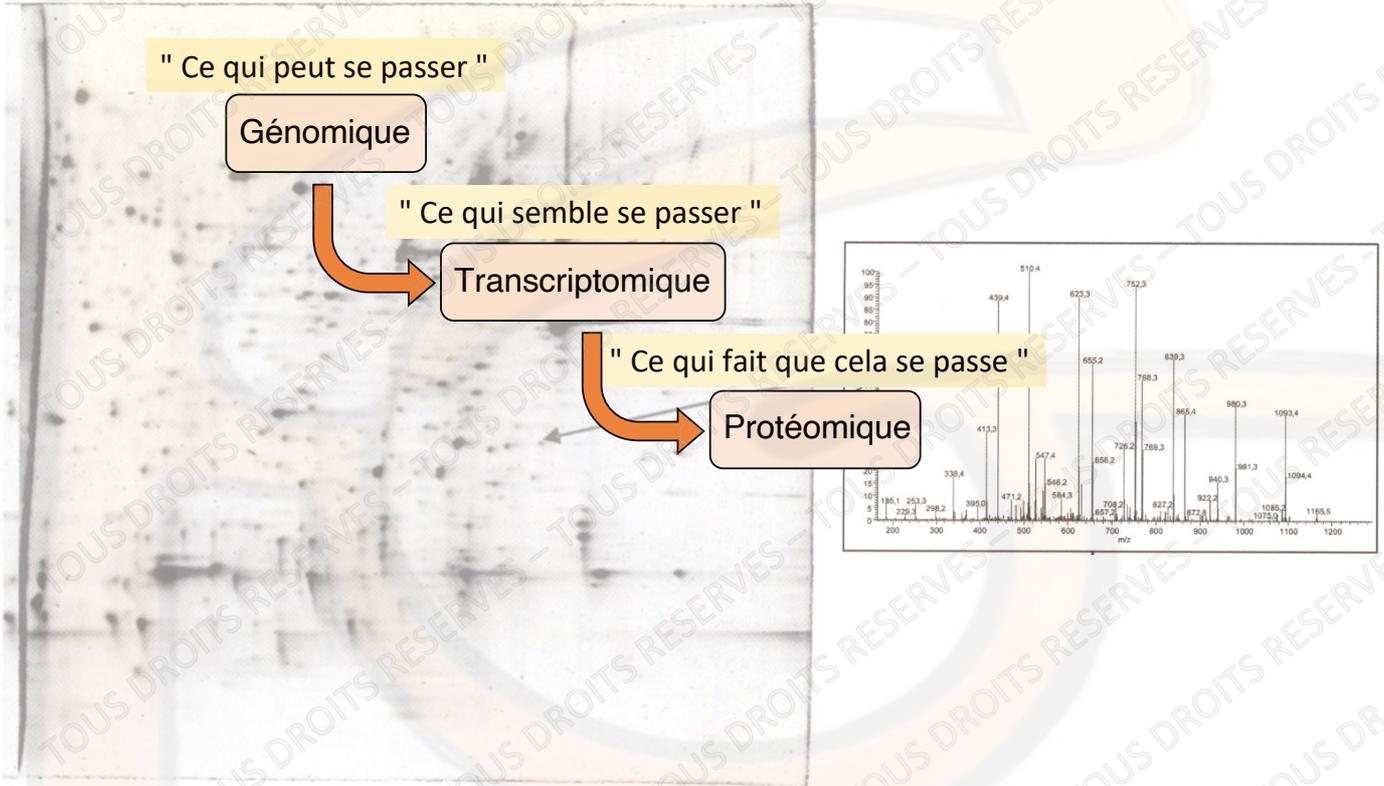
Applications des analyses en microarrays:

- 1- Identifier des gènes dont l'expression est modifiée entre échantillons expérimentaux
- 2- organiser des échantillons selon leurs profils d'expression
- 3- Etablir des profils d'expression : grouper des gènes dont l'expression change de manière coordonnées selon des traitements, des maladies ou des étapes de maladies.
- 4- Détection de mutations/polymorphismes (SNPs)
- 5- Analyse de pathogènes



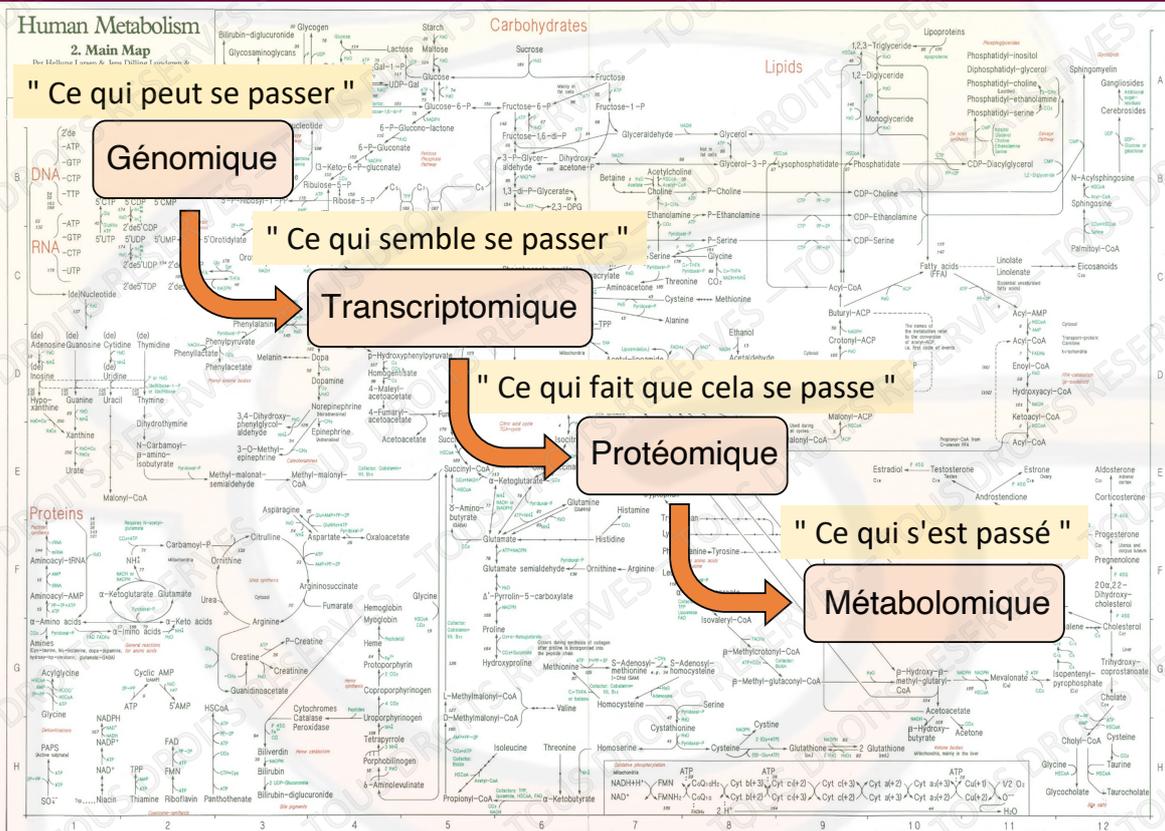
La cascade des "omiques" – la redécouverte du métabolisme

A



La cascade des "omiques" – la redécouverte du métabolisme

A





Régulation de l'expression des gènes

Ce qu'il faut retenir (1):

- régulation procaryote est principalement transcriptionnelle ou co-transcriptionnelle
 - savoir ce qu'est un opéron
 - connaître les mécanismes d'induction, de régulation négative, d'atténuation



Régulation de l'expression des gènes

Ce qu'il faut retenir (2):

- régulation eucaryote se situe à différents niveaux
 - connaître différents niveaux et mécanismes de régulation
 - connaître certains acteurs de la régulation transcriptionnelle
 - connaître certains mécanismes de régulation épigénétique, connaître les mécanismes d'acétylation/méthylation des histones et de méthylation de l'ADN. Connaître les conséquences sur l'activité transcriptionnelle.
 - Avoir une notion d'empreinte génomique parentale
 - Savoir ce qu'est un épissage alternatif et ses conséquences
 - Connaître certains mécanismes pouvant réguler la stabilité des ARNm
 - Connaître le schéma général de la régulation par les miRNA

Merci de votre attention



F. GESBERT

franck.gesbert@universite-paris-saclay.fr