

I GENERALITES SUR L'EXPRESSION GENIQUE

⇒ Aujourd'hui, grâce aux techniques d'analyses à haut débit, nanotechnologies, on a la possibilité d'étudier la science des X-omes : X étant le génome, protéome, transcriptome et -ome : ensemble

RECAP

Expression d'un gène :

- 1) ADN (*génome*)
- 2) Transcription en ARN (*transcriptome*)
- 3) Épissage
- 4) Transport ARN en dehors du noyau
- 5) Traduction protéine (*protéome*)
- 6) Maturation
- 7) Adressage de la protéine
- 8) Dégradation de la protéine

Génotype : Ensemble des caractères génétiques d'un individu dont l'expression différentielle est visible dans le phénotype. Les individus d'une même espèce ont les mêmes gènes, mais des allèles différentes (=génotype)

☑ La transcription et la traduction constituent l'expression du génotype pour obtenir le phénotype.

Phénotype : expression des gènes dans un **environnement** donné

- ❖ Ce que l'on transmet à nos cellules filles :
 - ☞ La séquence d'ADN
 - ☞ Cette séquence d'ADN dans un **environnement donné**. (Épigénome). L'environnement aura un effet sur la structure de la chromatine : association particulière entre l'ADN et certaines protéines => c'est l'épigénome, cet « épi » génome est transmis de générations en générations, mais cette transmission ne sera pas aussi stable que la transmission d'ADN.
- ❖ Au niveau de chaque compartiment cellulaire, des informations régulatrices vont établir et maintenir un **programme d'expression génique**, chaque étape est très régulée.



⇒ On a **20 000 gènes** dans nos cellules, mais ils ne sont pas tous exprimés (*on dit qu'un gène s'exprime quand il est transcrit en ARN*).

Les gènes s'expriment sont dits : **ON**, les autres, la plupart, sont : **OFF**

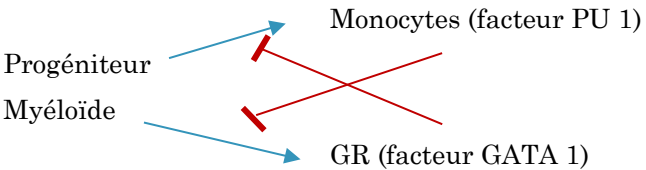
⇒ **10% seulement des gènes s'expriment** dans les 200 différents types cellules (gènes différents selon l'organe concerné).

- *Quels sont les mécanismes qui maintiennent les gènes ON et OFF ?*
- *Ces mécanismes sont-ils transmissibles avec les divisions ? Sont-ils héréditaires*

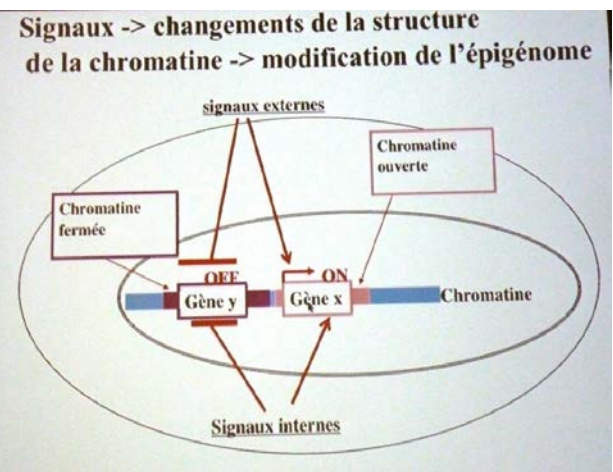
A. PROGRAMME TRANSCRIPTIONNEL

Un programme est établi en réponse à une signalisation spécifique (exogène/endogène), grâce à des facteurs spécifiques

- 1) Ceci entraîne **une activation combinée de plusieurs gènes** (les gènes ON)
- 2) Permettant la diversité des programmes cellulaires
- 3) C'est la différenciation cellulaire
- 4) Puis il faut le maintien de ce programme de transcription ;

<p><u>Exemple 1</u> : Comment passer du fibroblaste au myoblaste ?</p>	<ol style="list-style-type: none"> 1) On prend un fibroblaste 2) On « active un gène » grâce à un facteur de transcription myogénique (signal spécifique) 3) Cela provoque un programme transcriptionnel transformant notre fibroblaste en myoblaste (réponse au signal) <p>Le fibroblaste n'était pas destiné à former une cellule musculaire, mais la présence du facteur de transcription myogénique déclenche un programme entraînant une telle différentiation.</p>
<p><u>Exemple 2</u> : Progéniteur myéloïde</p> 	<ul style="list-style-type: none"> ➤ A partir d'un progéniteur myéloïde, on a deux grands types de différenciation possibles : la voie monocyttaire et la voie érythrocytaire. ➤ La cellule progénitrice peut aller dans les deux directions : elle décide selon les facteurs de transcription spécifique qu'elle reçoit. → PU1 entraîne la lignée vers la voie monocyttaire. → GATA-1 entraîne la lignée vers la différenciation en globule rouge. ➤ Pour être sûr qu'il n'y ait pas de retour. On va avoir un rétrocontrôle négatif par le facteur de la lignée activée.

⇒ Une fois que le programme transcriptionnel est établi, comment faire pour qu'il soit maintenu (4) ?

<p>Signaux -> changements de la structure de la chromatine -> modification de l'épigénome</p> 	<p>Les signaux (exogène/endogène) induisent un changement de structure de la chromatine ⇒ Ils modifient l'épigénome</p> <p>NB : l'épigénome porte les marques d'activation et d'inactivation des gènes.</p> <table border="1" data-bbox="708 1599 1522 1688"> <tr> <td>Chromatine ouverte : expression</td> </tr> <tr> <td>Chromatine fermée : non expression</td> </tr> </table> <p>L'épigénome est héritable, même en l'absence de signaux inducteurs (hormones ...) : ceci permet l'expression des gènes et le maintien du programme transcriptionnel (différenciation) entre les différentes générations cellulaires.</p>	Chromatine ouverte : expression	Chromatine fermée : non expression
Chromatine ouverte : expression			
Chromatine fermée : non expression			



Cette mémoire épigénétique est assurée par les changements chromatinien.

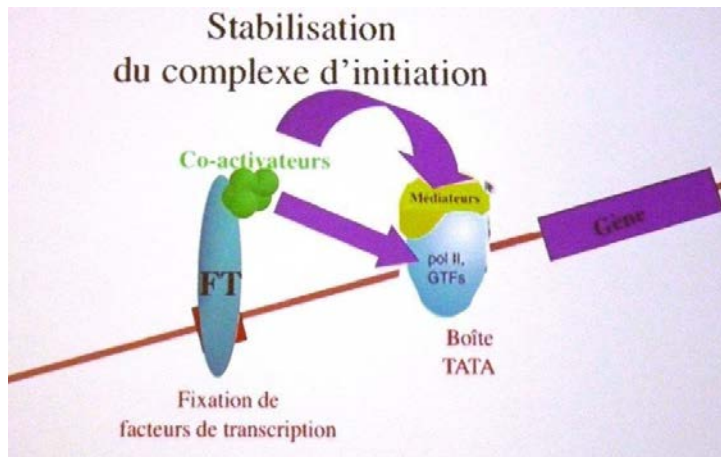
B. REGULATION DE L'EXPRESSION DES GENES

⇒ On a différents niveaux de régulation de l'expression des gènes.

- ❖ On a une séquence juste en amont du gène = le **promoteur**, où se fait l'**initiation de la transcription et un certain nombre de régulation**

COMPLEXE D'INITIATION

- Les chercheurs ont trouvé, à travers plusieurs expériences (cf. cours), la **structure nécessaire** à avoir pour une **transcription active** :



- Le **complexe d'initiation** (ARN polymérase II + facteurs généraux de transcription) se fixe sur la **boîte TATA**, sur laquelle va se fixer le **médiateur**.
- Ce complexe a besoin d'être stabilisé ; c'est le rôle des **facteurs de transcription** qui se fixent sur des régions spécifiques du promoteur, -eux-mêmes couplés éventuellement avec des **protéines co-activatrices**.
- Les **interactions protéines/protéines** entre les facteurs de transcription, les médiateurs et le complexe d'initiation vont venir stabiliser le complexe → le **gène s'exprime**

ORGANISATION MODULAIRE DES FT

- Jacques Monod (prix Nobel années 60) qui a découvert les **mécanismes d'expression des gènes** avec l'opéron lactose chez la bactérie (procaryote) :
- C'est un modèle de référence pour comprendre comment un complexe d'initiation agit chez la cellule eucaryote.

-En présence de galactose :

GAL4 (**facteur de transcription**) est activé ainsi que des gènes cibles du catabolisme du galactose : on veut **dégrader** pour faire de l'énergie.

Gal4 se fixe sur une **séquence** spécifique nommé **UAS** chez la levure et **stabilise** le complexe via des médiateurs.

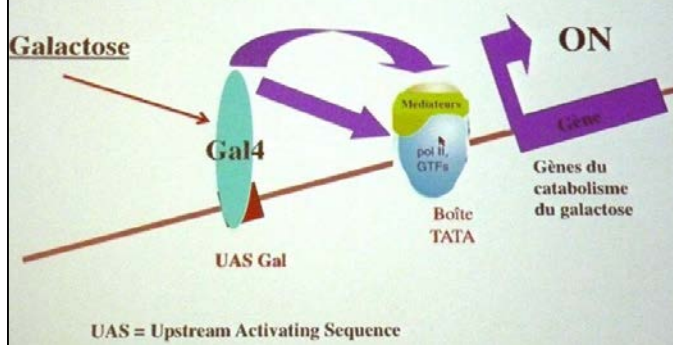
-Gal4 (comme beaucoup de facteur de transcription) est formé de **deux domaines** fonctionnels :

DFA = Domaine de fixation spécifique de l'ADN de Gal4 qui va permettre le recrutement spécifique de facteur de transcription sur le site de réponse.

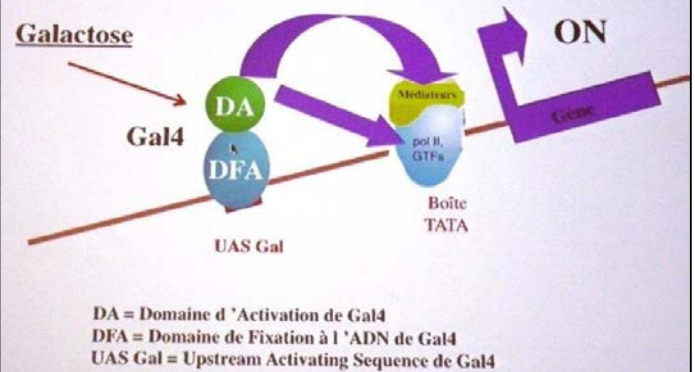
Da = domaine d'activation qui interagit avec les protéines du médiateur pour stabiliser le complexe et favoriser la transcription.

-Donc ces **deux domaines** sont **portés par la même protéine mais peuvent être séparés**. Différentes expériences vont montrer que les deux parties du facteur de transcription sont indépendantes et peuvent agir l'une indépendamment de l'autre si on remplace l'une ou l'autre des parties.

Activation simple chez la levure en réponse au galactose

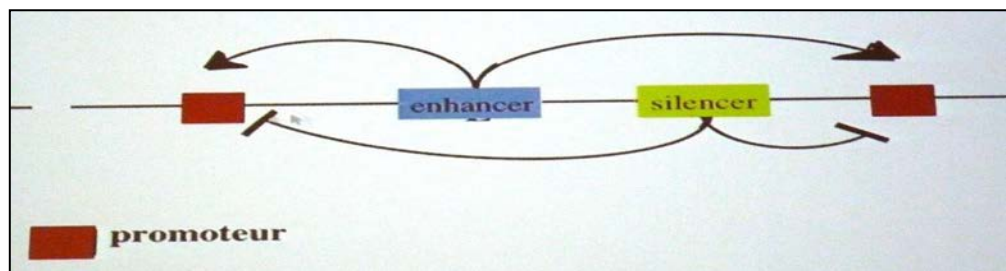


Gal4 est formé de deux domaines fonctionnels



- Ce sont des mécanismes très conservés avec des phénomènes universels (un virus de l'herpès fonctionne chez la levure).
- Tout ceci, sera transposé chez les eucaryotes, cependant on parlera d'Elément réponse (ER) pour UAS.

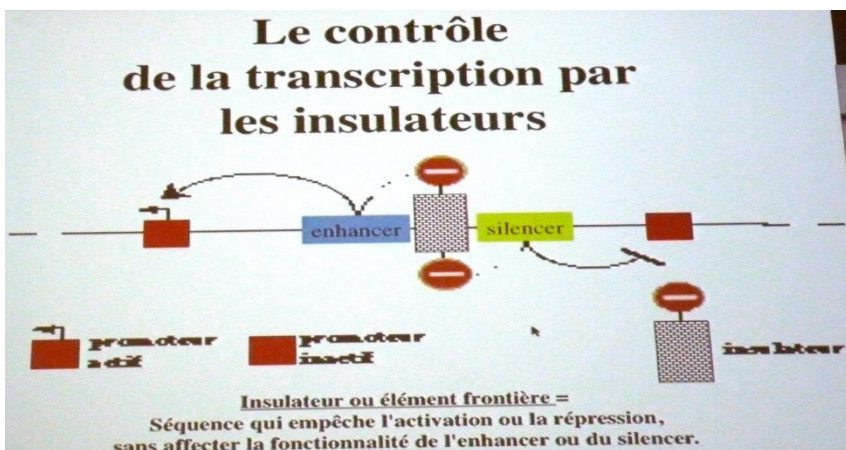
- Chez les mammifères (et les métazoaires en général), on a des contrôles distaux de la transcription pour déterminer si le gène est ON ou OFF par deux catégories de séquence : les enhanceurs (activateurs) et silencers (inhibiteurs)



- Ils sont très différents des promoteurs ; en effet :
 - Ils ont une **position variable** par rapport au promoteur (en 5' ou 3'),
 - Ils peuvent être **très éloignés** (à des milliers voire des millions de paires de base(pb))
 - Ils agissent **indifféremment de l'orientation**.
 - Ils agissent le plus **souvent en cis** (même chromosome) mais quelques exceptions en **trans**

- Ils donnent la direction de l'action des enhanceurs et des silencers.

Le contrôle de la transcription par les insulateurs

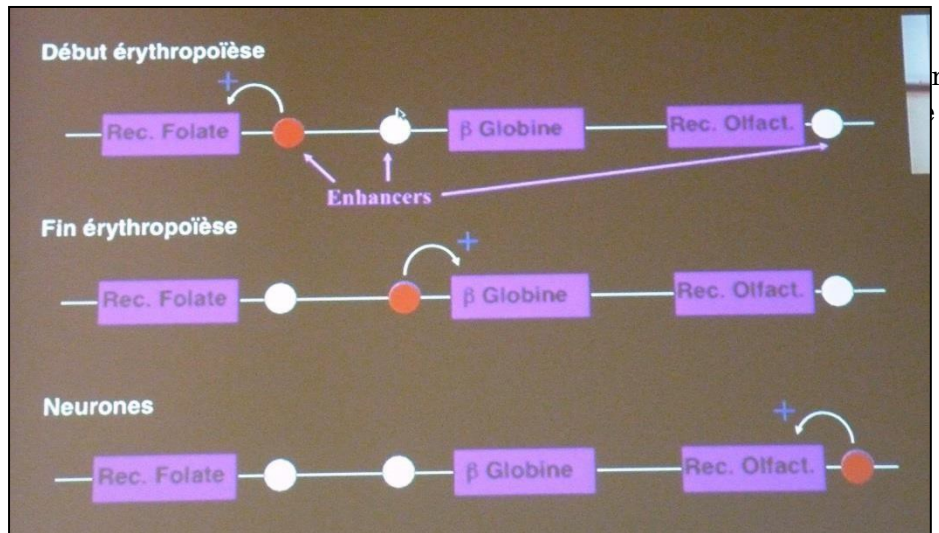


Ces éléments **frontières** n'empêchent pas l'action des enhanceurs et des silencers dans notre génome, mais vont les empêcher d'agir que dans **une direction** (exemple de la signalisation routière avec des STOP).

Ces insulateurs définissent des domaines de régulation spécifique dans notre génome

Exemple = l'Érythropoïèse

- On a trois gènes (le récepteur folate, le gène β -globuline (les deux impliqués dans l'érythropoïèse) et celui des récepteurs olfactifs qui sont côte à côte dans notre génome et on va voir ce qu'il se passe dans différentes cellules.
- On a des enhancers tissu/spécifique qui détermine l'expression de ces gènes.



En début d'érythropoïèse :

Le gène des récepteurs folate est ON, le gène de la β globine n'est pas encore utile, et celui des récepteurs olfactifs est totalement inutile : ils sont OFF.

En fin d'érythropoïèse : on a plus besoin des récepteurs folate le gène est OFF, on a besoin de la β -globine il est ON, et le gène des récepteurs olfactifs est toujours OFF.

Dans les neurones : on ne fait pas d'érythropoïèse, les 2 premiers sont OFF, le gène des Rc olfactifs est ON

Contrôle Proximal	Contrôle distaux	Régulation différentielle/ distal
Promoteur	Enhancers & Silencers	Insulateurs
En amont du gène	Position variable ++ à distance	A distance
Agit de manière unidirectionnelle	Agissent dans les deux directions	Réduisent les enhancers et les silencers à une seule direction

II STRUCTURE DE LA CHROMATINE

- La chromatine est le support des modifications liés aux facteurs de transcriptions.
- Les modifications de la chromatine sont héréditaires : l'épigénome se transmet.

A. COMPACTION DE L'ADN NUCLAIRE

Réalité/Problème

2m d'ADN doivent rentrer dans un noyau de 10 μ m

⇒ La compaction de l'ADN est indispensable

2 nucléotides A et B séparés de 80 paires de bases dans la séquence linéaire peuvent être côte à côte dans le noyau

⇒ Qu'est-ce qui provoque cela ?

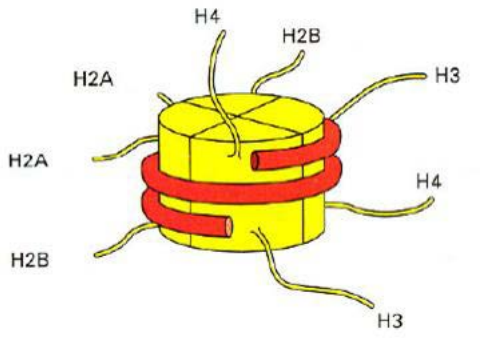
➤ La structure en chromatine bloque l'accès aux facteurs de transcription ou d'élongation, aux protéines qui veulent s'associer à la TATA BOX.

➤ La chromatine assure une organisation du matériel génétique ce qui :

- ☞ Facilite la ségrégation lors de la mitose
- ☞ Défavorise la transcription

- ❖ Tout cela est dû au nucléosome qui est à la base de la structure de la chromatine, c'est le premier niveau de compaction de l'ADN.

B. STRUCTURE DU NUCLEOSOME

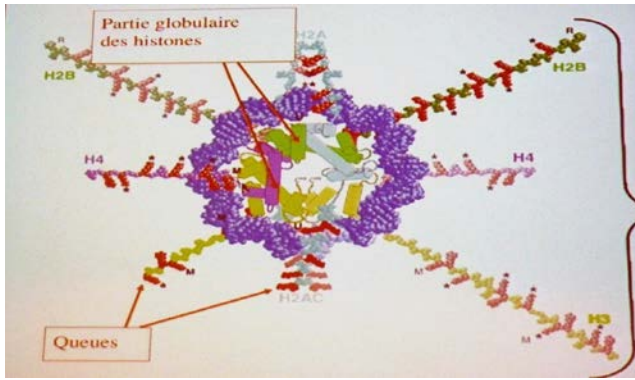
<p>LE NUCLEOSOME</p>	<ul style="list-style-type: none"> Le nucléosome est le constituant de base de la chromatine. Cylindre de 6nm d'hauteur et 11 de diamètre. Il est constitué de 146 paires de base d'ADN enroulées autour de 8 molécules d'histones (ce qui forme un octamère: $2[H2A + H2B] + 2[H3 + H4]$). Le maintien de l'ADN autour de l'octamère d'histones est assuré par des interactions électrostatiques entre l'ADN et les protéines. S'assemble de manière spontanée mais aider par les protéines chaperon 	
<p>LES HISTONES</p>	<ul style="list-style-type: none"> Ces histones sont très conservées au cours de l'évolution (dans tous les organismes eucaryotes), ce qui montre leur importance fonctionnelle. Il n'y en a pas chez les bactéries, elles ont des systèmes de compaction différents. Elles font entre 20 et 30Kd chacune. On a 60 millions d'histones dans un de nos noyaux 	
<p>LES PROTÉINES CHAPERON</p>	<ul style="list-style-type: none"> Stimulent l'assemblage du nucléosome de manière conformationnelle sans réaction enzymatique, interagissent avec les dimères d'histones en plusieurs étapes : <ul style="list-style-type: none"> 1) ADN nu 2) Chaperons guident 2 x l'hétérodimère H3 H4 vers l'ADN => formant un « demi nucléosome » 3) Chaperons guident 2 x l'hétérodimère H2A H2B vers l'ADN => formant un nucléosome complet 	

Expérience : Pour dégrader les protéines et récupérer les histones

- Utilisation de la nucléase micrococcale Elle dégrade l'ADN linker : coupe entre les nucléosomes
 - ☞ Digestion partielle : 146 pdb + ADN linker = 200 pdb = mono,di-nucléosome
 - ☞ Digestion totale : isole le cœur protéique = 146 pdb = nucléosome
- Ajout d'une solution NaCl => destruction de l'interaction ADN-protéine : on obtient l'ADN d'un côté et les protéines histones de l'autre
- On fait descendre le Ph à 3 : cela va dénaturer nos protéines, on obtient nos histones (2 H2A, 2H2B, 2H3, 2H4)

C. DIVERSITE DES NUCLEOSOMES

- La base du nucléosome est conservée, mais des petites différences peuvent être présentes : différentes fonctions de la chromatine et la diversité cellulaires.
- Comment est organisé un nucléosome (méthode de diffraction par rayon X)



Des séquences d'ADN favorisées se courbent naturellement pour s'enrouler autour des histones

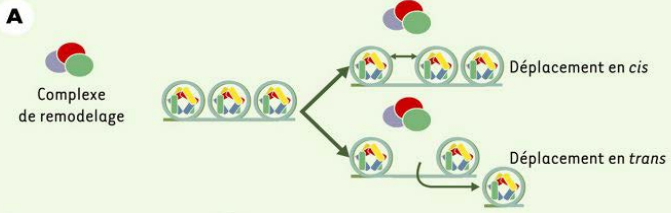
Des séquences défavorisées sont au contraire rigide

Le **nucléosome** possède :

=>Une partie globulaire centrale (« histone fold ») composés d'hélices + résidus basiques des histones.

=>En périphérie les queues N-Term des histones chargés+ : ils font saillie et paraissent flexibles : ils sont la cible des modifications post traductionnelles, et sensibles aux protéases (trypsine...)

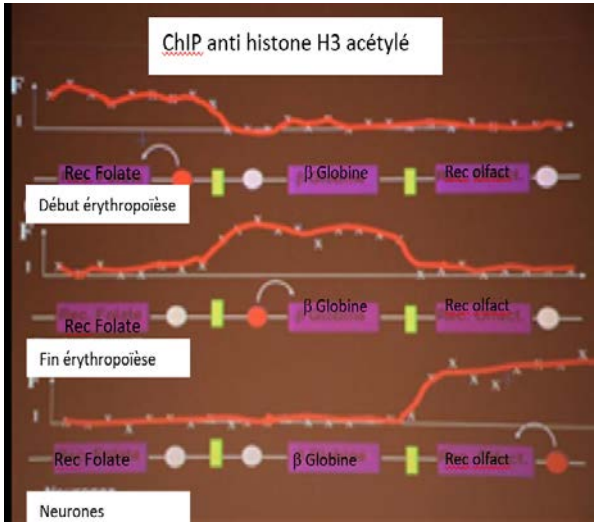
• On compte 3 voies de modifications des nucléosomes :

<div>COMPLEXE REMODELAGE</div>	<ul style="list-style-type: none">Déplacement en CIS sur la même molécule d'ADNDéplacement en TRANS sur une autre molécule d'ADN	<div><div><div><div>A</div><div><div>Complexe de remodelage</div><div>EXEMPLES</div><div>Déplacement en cis</div><div>Déplacement en trans</div></div></div></div></div>						
<div>VARIANTS D'HISTONES :</div>	<ul style="list-style-type: none">Des variants d'histones H2A, H2B, H3PAS de variant pour H4 qui n'a qu'un seul gène, un seul variant.	<div><div><div>Exemple : variant d'H2A = H2AX</div><div>H2AX ressemble à H2A en termes de séquence protéique mais les gènes sont différents. H2AX est impliqué dans la reconnaissance et la réparation des morceaux d'ADN.</div><div>Exemple : variant d'H3 = CenpA "centromeric protein A"</div><div>CenpA ressemble à l'histone H3 avec quelques petites différences. C'est une protéine des centromères.</div><div>Donc les nucléosomes des centromères ont l'histone CenpA à la place de l'histone H3. Elle permet de former les kinétochores.</div></div></div>						
<div>MODIFICATIONS POST TRADUCTIONNELLES</div>	<ul style="list-style-type: none">Ces modifications se font après la traduction grâce à des enzymes spécifiquesL'identification de ces enzymes à l'origine de ces modifications post-traductionnelles sont des cibles thérapeutiques.Les positions des lysines méthylée et le nombre de méthylation vont avoir des conséquences fonctionnelles différentes.	<div><div>⇒ Ces modifications sont contrôlées par des enzymes spécialisées :</div><table><tr><th>Acétylation sur Lysine (K)</th><th>Méthylation sur Lysine et Arginine</th></tr><tr><td><div>Ajout : -HAT (histone acétyl transférase)+ coenzyme acétyl- coA</div></td><td><div>Ajout : -HMT (histone méthyl transférase) Précisions : 1 à 3 méthylations sur Lysine au niveau du NH3</div></td></tr><tr><td><div>Suppression : HDAC (I, II, IV) -Sirtuines (utilisent NAD) font le lien métabolisme/chromatine</div></td><td><div>Suppression : -HDM (histone deméthylase)</div></td></tr></table></div>	Acétylation sur Lysine (K)	Méthylation sur Lysine et Arginine	<div>Ajout : -HAT (histone acétyl transférase)+ coenzyme acétyl- coA</div>	<div>Ajout : -HMT (histone méthyl transférase) Précisions : 1 à 3 méthylations sur Lysine au niveau du NH3</div>	<div>Suppression : HDAC (I, II, IV) -Sirtuines (utilisent NAD) font le lien métabolisme/chromatine</div>	<div>Suppression : -HDM (histone deméthylase)</div>
Acétylation sur Lysine (K)	Méthylation sur Lysine et Arginine							
<div>Ajout : -HAT (histone acétyl transférase)+ coenzyme acétyl- coA</div>	<div>Ajout : -HMT (histone méthyl transférase) Précisions : 1 à 3 méthylations sur Lysine au niveau du NH3</div>							
<div>Suppression : HDAC (I, II, IV) -Sirtuines (utilisent NAD) font le lien métabolisme/chromatine</div>	<div>Suppression : -HDM (histone deméthylase)</div>							

❖ Toutes ces modifications sont liées au :

LE CODE HISTONE	C'est le code de la chromatine combinant les modifications des résidus sur N ou C term + les variants d'histones. Ce code histone se superpose au code génétique. C'est ce qui nous donne le code EPIGENETIQUE.
------------------------	---

❖ Comment les étudier ?

IMMUNOPRÉCIPITATION DE CHROMATINE : CHIP	<p>Technique de référence permettant l'étude des modifications post-traductionnelles de queues N-term d'histones :</p> <ol style="list-style-type: none"> 1) On fixe nos cellules au formaldéhyde, pontage entre les protéines et l'ADN 2) Sonication => permet la fragmentation de l'intégralité de la chromatine => purification de la chromatine 3) Ajout d'anticorps (Ac) spécifiques de la modification recherchée, qui se fixe à son niveau 4) Immunoprécipitation => on garde les complexes immuns : Ac-chromatine avec modifications 5) Reverse pontage : purification de l'ADN (on n'a plus les protéines à partir de là, c'est un peu l'inverse de l'étape 1) 6) Etude de l'ADN : <p>Estimation par PCR ou hybridation (FISH) l'enrichissement relative d'une séquence par rapport à une l'input.</p> $F = \frac{RIP}{RC} = \frac{Q \text{ frag PCR} / Q \text{ ADN react}}{Q \text{ frag PCR} / Q \text{ ADN react}}$ <p style="text-align: right;"><i>Enrichissement = F</i> <i>RIP = quantité relative de fragment PCR (avec la modif) sur l'immunoprécipitat</i> <i>RC = quantité relative de fragment PCR (avec modif) sur l'input (ADN de départ)</i></p> <p>⇒ L'enrichissement F indique à quel degré la modification est présente dans la séquence étudiée par PCR</p> <p>Exemple de l'acétylation de H3 : calcul du facteur d'enrichissement tout le long sde ces gènes, 1 point = 1 amplification PCR</p>
	 <p>⇒ L'acétylation n'est pas constante selon les situations :</p> <p><u>Début d'érythropoïèse</u> : gène Rec folate acétylé => exprimé</p> <p><u>Fin érythropoïèse</u> : gène βGlobine acétylé => exprimé</p> <p>⇒ L'acétylation varie également selon le type cellulaire</p> <p><u>Neurones</u> ont le gène pour le Rc olfactif acétylé => exprimé</p> <p>⇒ On a une même séquence d'ADN avec des modifications de chromatine différents => l'expression des gènes sera différente => code épigénétique</p>