

Structure du chromosome

Pr Dali-Sahi Majda

Faculté SNV/.STU :Département de Biologie

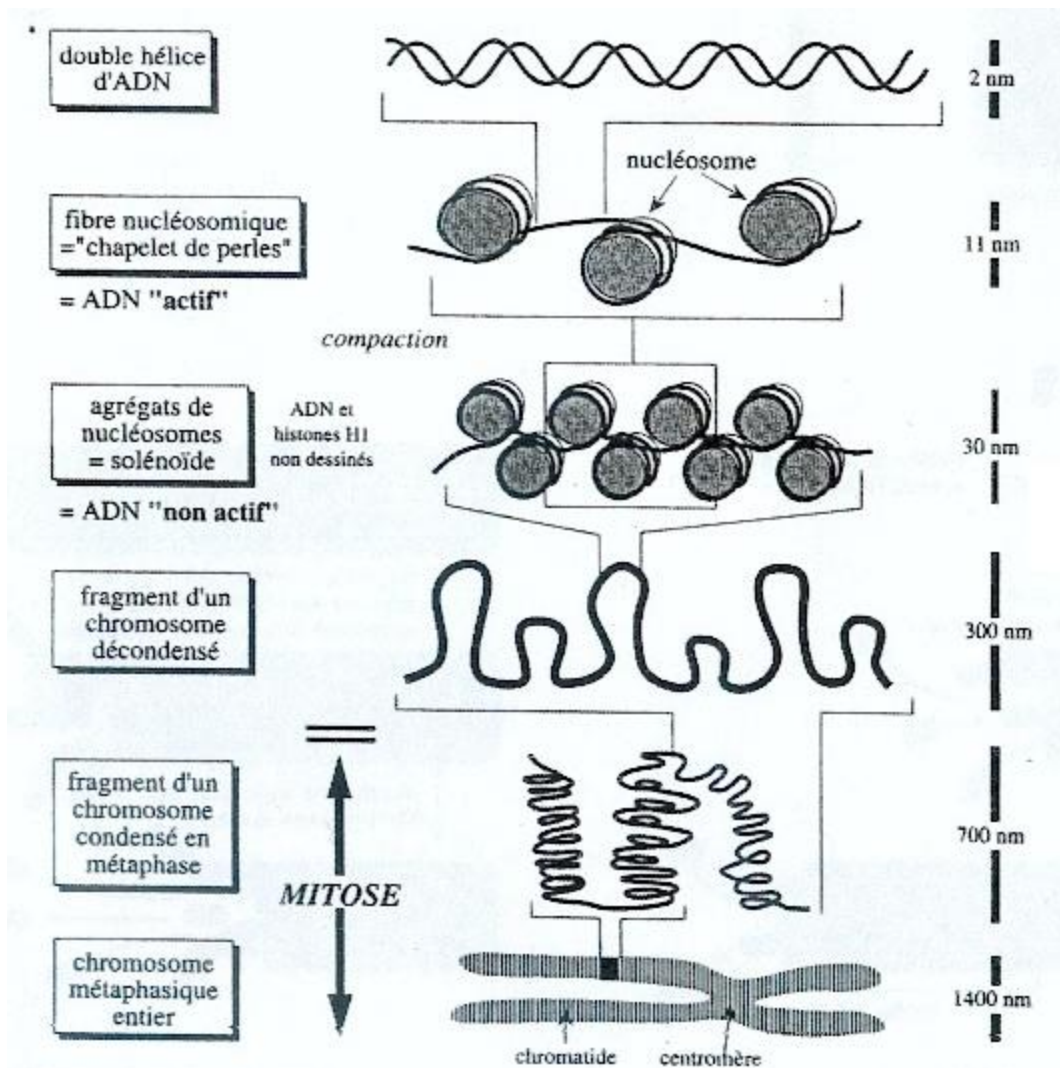
Chromatine : ADN associé à des protéines de liaison (histones)

□ Histones : petites protéines liée à l'ADN formant nucléosomes ; 4 types : H1, H2B, H3 & H4

□ Nucléosomes : enroulement de 146pb (paires de bases de l'ADN autour des histones

□ Eurochromatine : ADN codant (gènes)

La chromatine se présente sous différentes formes durant le cycle cellulaire : dense et fibreuse ou hyper condensée (chromosomes)



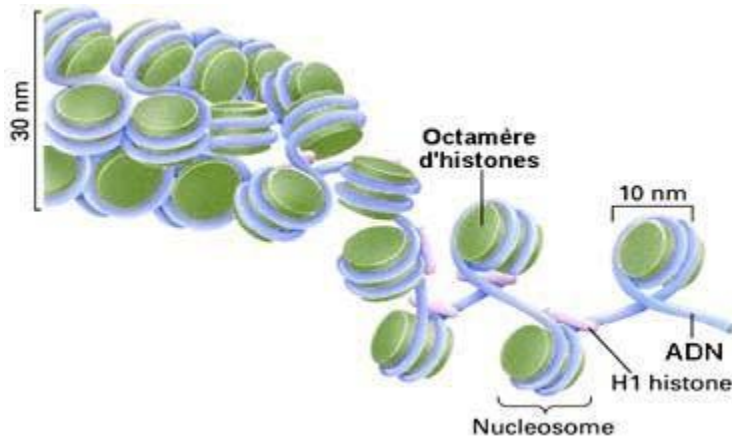
les différents degrés de compaction de l'ADN nucléaire

Figure 3 : Etat du chromosome au cours du cycle cellulaire

Au delà des nucléosomes, la chromatine présente des niveaux d'organisation supérieurs encore peu caractérisés. Le nucléofilament se compacte pour former la fibre de 30 nm qui s'organise en boucles de 150 à 200 kpb (250 nm pendant l'interphase) pour atteindre un niveau de compaction maximum dans le chromosome métaphasique (850 nm).

L'ADN possède plusieurs degrés de compaction dans le noyau interphasique :

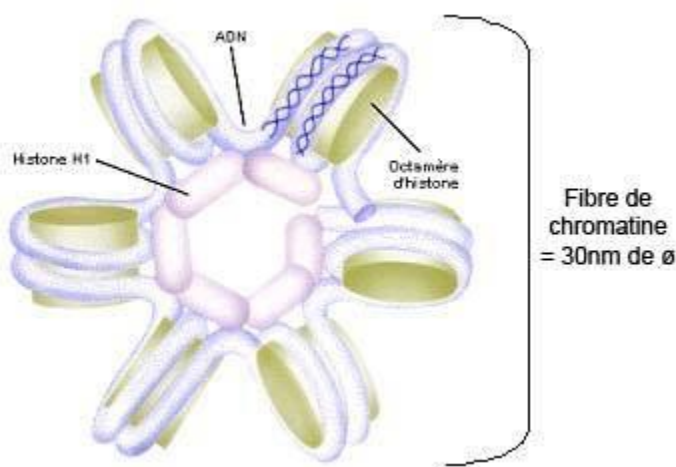
a-Le premier niveau de compaction : les nucléosomes.



Ce sont des structures ayant la forme d'un cylindre de 10 nm de diamètre, formées de petites protéines appelées histones nucléosomiques qui sont chargées positivement, ce qui facilite leur fixation à l'ADN (qui est chargé négativement). Il s'agit d'un octamère (H2A, H2B, H3 et H4)_{x2}.

L'ADN fait 2 tours (=146 paires de bases) autour de chaque cylindre. Les nucléosomes sont séparés par un court segment d'ADN de taille variable 0 à 60-80 pb. La chaîne d'ADN ressemble alors à un collier de perles.

b-Le deuxième niveau de compaction : le solénoïde



Les nucléosomes sont associés par 6 par une autre histone, l'histone H1, pour former des "Solénoïdes". L'histone H1 se lie à l'ADN à sa sortie du nucléosomes. Les molécules d'histone

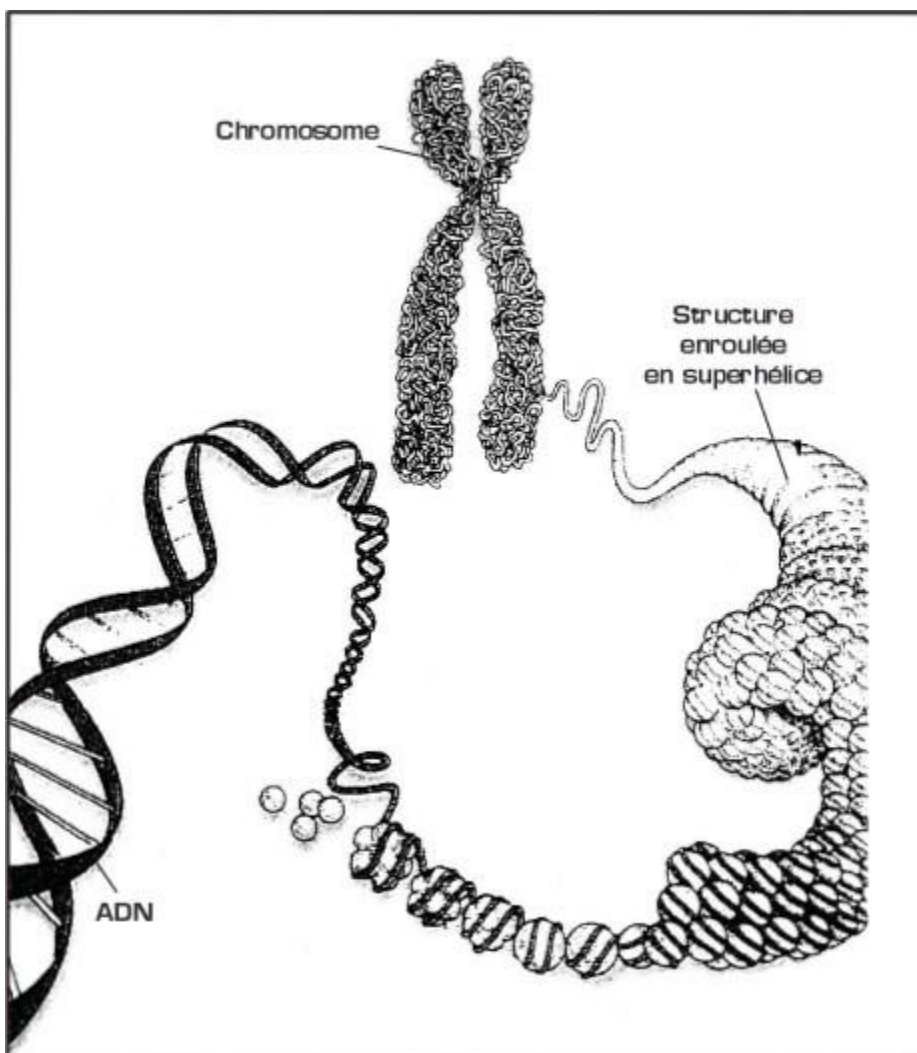
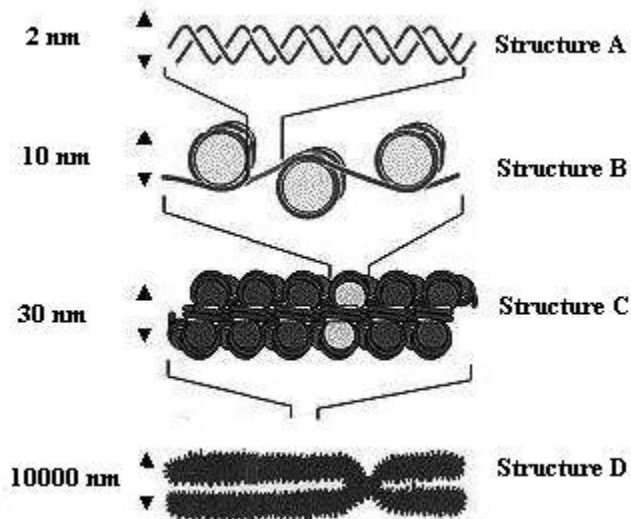
H1 sont reliées entre elles par des liaisons peptidiques. Elles sont responsables de la constitution des fibres de chromatine de 30nm de diamètre.

c-Le troisième niveau de compaction : les boucles d'ADN

Les fibres de 30 nm forment des boucles des quelques mégabases qui s'attachent à une armature protéique flexible

d-Le dernier niveau de compaction : le chromosome métaphasique

La forme de compaction ultime est observée au niveau du chromosome métaphasique



Dans le chromosome métaphasique, la compaction est de l'ordre de 10^4 .

Il y a 2 protéines (ou enzymes) importantes:

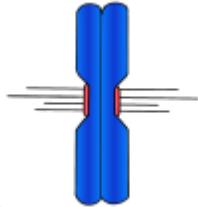
La topoisomérase II : Elle intervient pour ordonner les boucles qui étaient enchevêtrées.

La SMC ATPase : Elle est indispensable pour condenser la chromatine, c'est une protéine motrice.

LE CENTROMERE ET LE TELOMERE.

A. Le centromère.

C'est la structure centrale du chromosome. De part et d'autre de ce centromère, il existe un ensemble de protéines que l'on nomme le «kinétochore»..



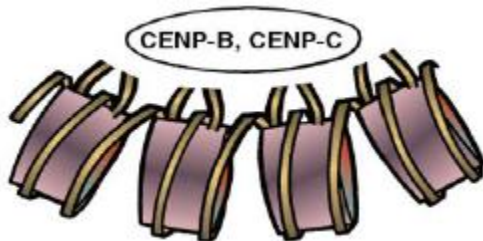
Chromosome

kinétochore»

La taille du centromère est de **250 à 5000** kilobases. Au niveau du centromère, il y a **toujours** une protéine qui se nomme CENPA. Elle prend la place des histones H3 au niveau de la région centromériques.

Cette protéine possède une extrémité Nter de grande taille qui va accrocher 2 autres protéines: CENPB et CENPC.

Elles seront responsables de l'interaction avec les microtubules au moment de la métaphase



B. Télomères.

Ce sont les structures extrêmes du chromosome.

Deux protéines sont importantes: TRF1 et TRF2 (elles s'associent de façon coopérative) vont stabiliser la chromatine et inhiber le complexe télomérique (télomérase).

L'HETEROCHROMATINE ET L'EUCHROMATINE.

Ce sont 2 parties structurellement et fonctionnellement différentes.

A. Hétérochromatine.

C'est de la chromatine fortement condensée qui résiste à l'activité des nucléases.

(Les nucléases sont des enzymes qui attaquent les molécules du noyau).

Caractéristiques:

- Un % de 5 méthyl-cytosine très important **N**
- Un grand nombre d'histones H1
- Un taux d'acétylation plus faible

La présence de 5 méthyl-cytosine est reconnue par des protéines qui vont contribuer à former de la chromatine condensée.

B. Euchromatine.

On retrouve les protéines HMG 14 et 17 qui prennent la place de H1 et stabilisent la structure en collier de perles.

Caractéristiques:

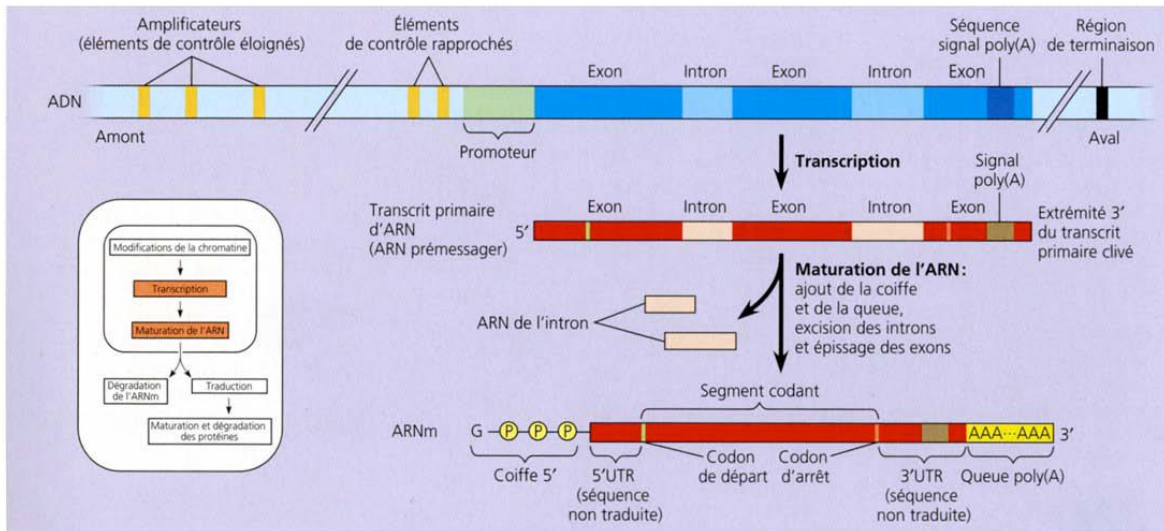
- Il y a moins de 5 méthyl-cytosine N
- Il y a moins d'histones H1 N
- Le degré d'acétylation est plus élevé N
- Elle correspond à de la chromatine transcrite N

La régulation de l'expression des gènes : régulation de la structure de la chromatine

La structure de la chromatine permet de condenser l'ADN mais elle permet aussi de réguler l'accessibilité de l'ADN pour la transcription : les gènes situés dans l'hétérochromatine ne sont pas transcrits. En plus, des modifications chimiques des histones régulent aussi la transcription : les queues des histones peuvent être **acétylées** par des enzymes spécifiques, sur des acides aminés lysines. Lorsque les histones sont acétylées, elles perdent leurs charges positives et interagissent moins avec l'ADN et les nucléosomes voisins --> la chromatine a une structure plus lâche et l'ADN est plus accessible aux facteurs de transcription.

La régulation de l'expression des gènes : régulation de l'initiation de la transcription

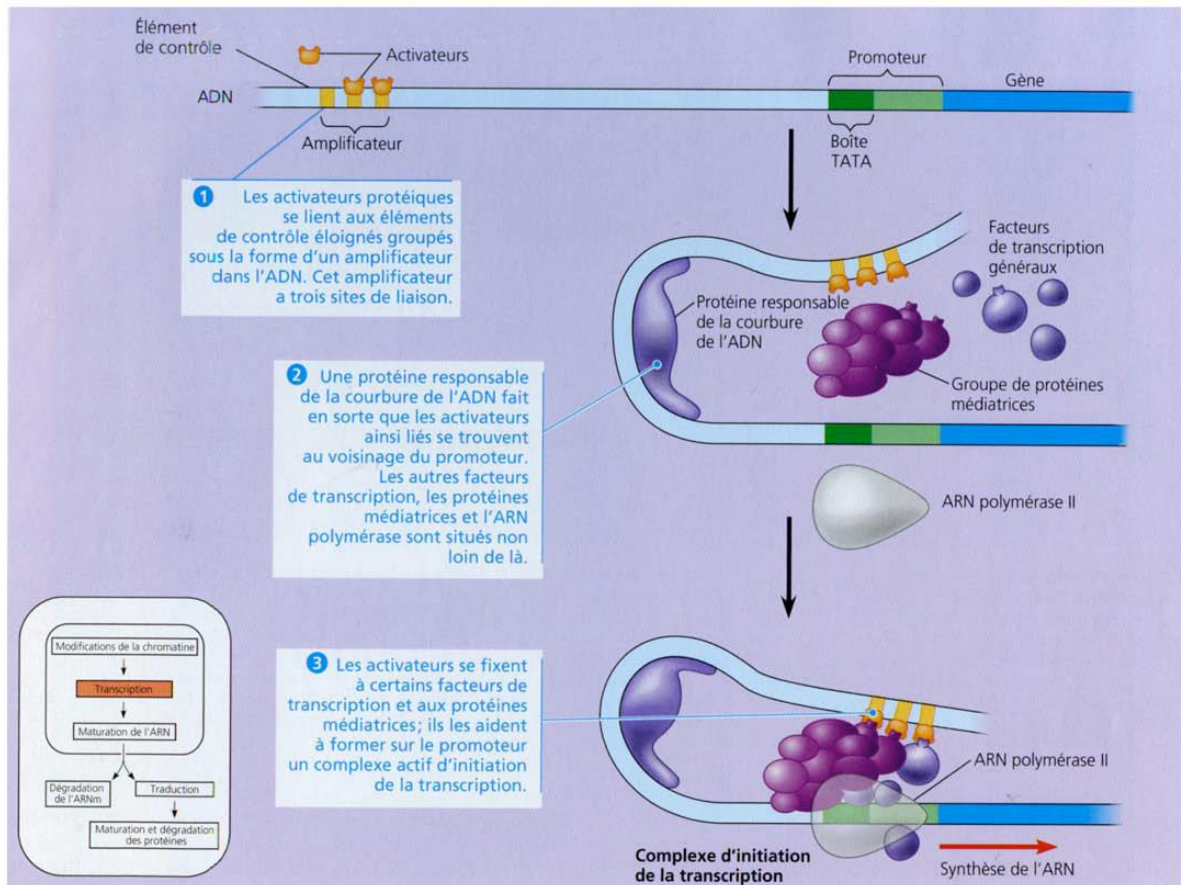
La transcription est initiée quand un complexe d'initiation de la transcription s'assemble au niveau du promoteur du gène. Ce complexe comprend l'ARN polymérase qui va synthétiser l'ARN pré-messager mais aussi différents *facteurs de transcription généraux* nécessaires à la transcription de tous les gènes. Pour assurer un niveau de transcription élevé de certains gènes (par ex à un moment donné ou dans une cellule précise), *d'autres facteurs de transcription appelés alors spécifiques* sont nécessaires.



▲ Figure 19.5 Gène d'eucaryote et son transcrit.

La régulation de l'expression des gènes : régulation de l'initiation de la transcription

Des éléments de contrôle proximaux (promoteur) et plus distaux (Enhancer) sont présents dans le promoteur des gènes, ils sont appelés **amplificateurs**. Ces éléments sont reconnus par des facteurs de transcription spécifiques (**activateurs** ou **répresseurs**) qui vont augmenter ou diminuer le taux de transcription du gène.



▲ Figure 19.6 Modèle d'action d'un amplificateur et des activateurs de transcription.

La régulation de l'expression des gènes : régulation de l'initiation de la transcription

On estime qu'il existe en moyenne dix éléments de régulation dans le promoteur de chaque gène et environ une douzaine de séquences de régulation différentes. C'est la combinaison particulière des éléments de régulation dans un promoteur donné qui déterminera le taux de transcription du gène en fonction du type cellulaire et des signaux que perçoit la cellule.

La régulation de l'expression des gènes au niveau post-transcriptionnel

Au niveau de la maturation de l'ARN : l'**épissage alternatif** (ou épissage différentiel) est un processus par lequel différents ARNm peuvent être produits à partir du même ARN pré-messager selon que les différents segments sont considérés comme des exons ou des introns. Chaque ARNm différent produira une protéine différente

L'épissage alternatif : un gène, combien de protéines ?

Avant la publication de la séquence complète de l'ADN du génome humain, au début des années 2000, on estimait le nombre de gènes à environ 300.000. Aujourd'hui, ce chiffre est tombé à

environ 22.000, un résultat étonnant car finalement très voisin de celui d'autres espèces parmi les préférées des généticiens : la souris, le poisson-zèbre ou même le simple ver nématode, qui possède également plus de 20.000 gènes ! En d'autres termes, le nombre de gènes d'un organisme vivant ne reflète pas sa réelle complexité biologique. Ce paradoxe résulte de la combinaison de plusieurs phénomènes, dont ce qu'on appelle l'épissage alternatif des ARN pré-messagers, une étape fondamentale de l'expression des gènes.

Les différentes étapes de l'expression d'un gène

L'expression génique est une succession d'étapes finement contrôlées à différents niveaux (**Figure 1**). En effet, bien que toutes les cellules d'un organisme contiennent le même génome, elles n'expriment pas les mêmes gènes selon l'environnement dans lequel elles se trouvent, ou selon les stimuli auxquels elles sont exposées. La première étape de l'expression d'un gène consiste en sa transcription à partir de son promoteur, une séquence précisant le début de la transcription et dont l'activité est régulée par des protéines spécifiques de chaque type cellulaire. La transcription s'effectue dans le noyau par une machinerie composée de plusieurs dizaines de facteurs protéiques qui copient la totalité de la séquence d'ADN du gène pour donner naissance à un transcrit primaire appelé ARN pré-messager (pré-ARNm).

Ce transcrit primaire subit ensuite une série de modification (« maturation ») à ses extrémités : l'ajout d'une coiffe à une extrémité et d'une queue d'environ 200 molécules d'Adénines à l'autre (« capping » et polyadénylation, respectivement). Cependant, la modification principale subie par l'ARN pré-messager consiste en l'élimination d'une grande partie de sa séquence, ce qu'on appelle l'épissage.

On sait depuis 1977 que les gènes humains sont discontinus et que les exons, c'est-à-dire les séquences contenant l'information génétique à exprimer en protéines, sont séparés par de larges séquences généralement non codantes appelées introns.

On parle ainsi de gènes multi-exoniques.

Les exons internes d'un gène mesurent en moyenne 150 nucléotides tandis que les introns mesurent 10 fois cette taille, et parfois bien plus encore. L'épissage consiste en l'élimination des introns et à la mise bout à bout des exons. Le transcrit mature ainsi formé, nommé ARN messager (ARNm), peut alors être exporté dans le cytoplasme (voir ci-dessous) où il sera traduit en protéine. Il est important également de souligner que de nombreux contrôles visant à valider la qualité des ARNm sont réalisés avant leur export, de façon pour la cellule d'éviter de poursuivre l'expression de transcrits aberrants, ce qui pourrait lui être dommageable.

Vers une nouvelle définition du gène ?

A l'heure où l'emploi de ce mot est devenu courant, il est très difficile de définir ce qu'est un gène. C'est en premier lieu un segment d'ADN qui peut être transcrit en ARN (**Figure 2**). Un gène commence par un promoteur permettant l'initiation de sa transcription et se termine par une séquence terminatrice. L'initiation et la terminaison peuvent avoir lieu à différents endroits au sein du gène et ainsi mener à la production de transcrits plus longs ou plus courts. Le pré-ARNm peut être épissé d'une manière alternative, il s'agit du processus d'épissage alternatif. En effet, il n'est pas obligatoire que tous les exons d'un gène soient inclus dans l'ARN messager mature, certains exons

sont donc considérés comme « alternatifs ». Ainsi un seul gène peut produire différents ARNm matures et par conséquent, plusieurs isoformes protéiques ayant des fonctions biologiques différentes, voire opposées. Au vu de ces données, le traditionnel dogme résumant le flux de l'information génétique à la transcription de l'ADN en un ARNm unique puis à la traduction de ce dernier en protéine, doit être revisité. Il est inconcevable de définir "un gène" comme une entité codant pour une seule protéine.

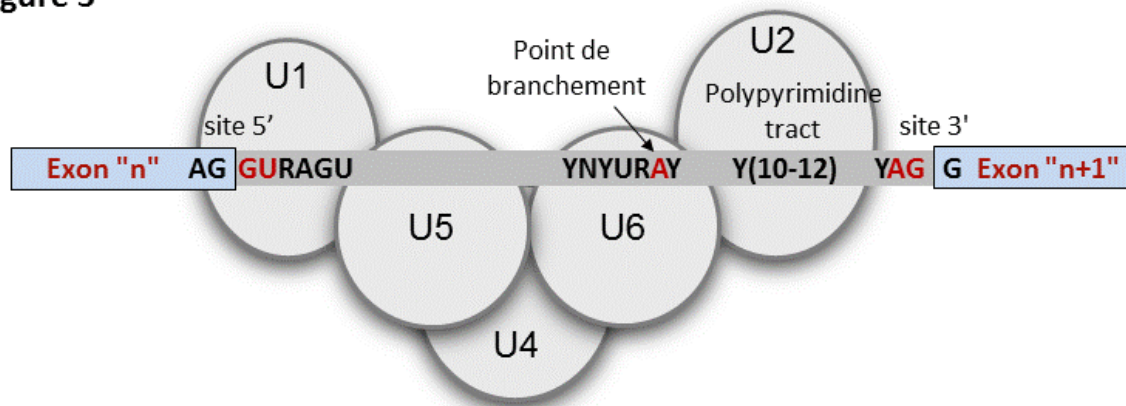
Unité fondamentale de la génétique : EXON

Aujourd'hui, l'unité de l'information génétique n'est plus le gène mais l'exon. Un gène doit plutôt être considéré comme une succession d'exons sélectionnés alternativement et permettant la production d'un ensemble de transcrits matures codant autant d'isoformes protéiques. Ainsi un gène est un message qui peut être interprété par la cellule de différentes manières pour assurer différentes fonctions biologiques selon ses conditions environnementales, son stade de développement et sa spécialisation.

La réaction d'épissage : Le spliceosome

Les exons du pré-ARNm doivent être réunis ensemble pour maintenir, d'une manière précise, le cadre de lecture (la séquence codante pour la protéine). Par conséquent, le mécanisme moléculaire de l'épissage exige une extrême précision pour reconnaître efficacement les frontières délimitant les introns et les exons dans le pré-ARNm (ce qu'on appelle les séquences 5' et 3' des introns, ou sites 5' et 3' d'épissage) (**Figure 3**). La réaction d'épissage est effectuée par le « spliceosome » (dérivé du mot anglais « splicing », épissage), une machinerie moléculaire complexe constituée de plus de 200 protéines.

Figure 3



La réaction d'épissage fait appel à 5 snRNPs (petites ribonucléoprotéines nucléaires) : U1, U2, U4/U6 et U5 qui reconnaissent des courtes séquences de nucléotides caractérisant les sites 5' et 3'.

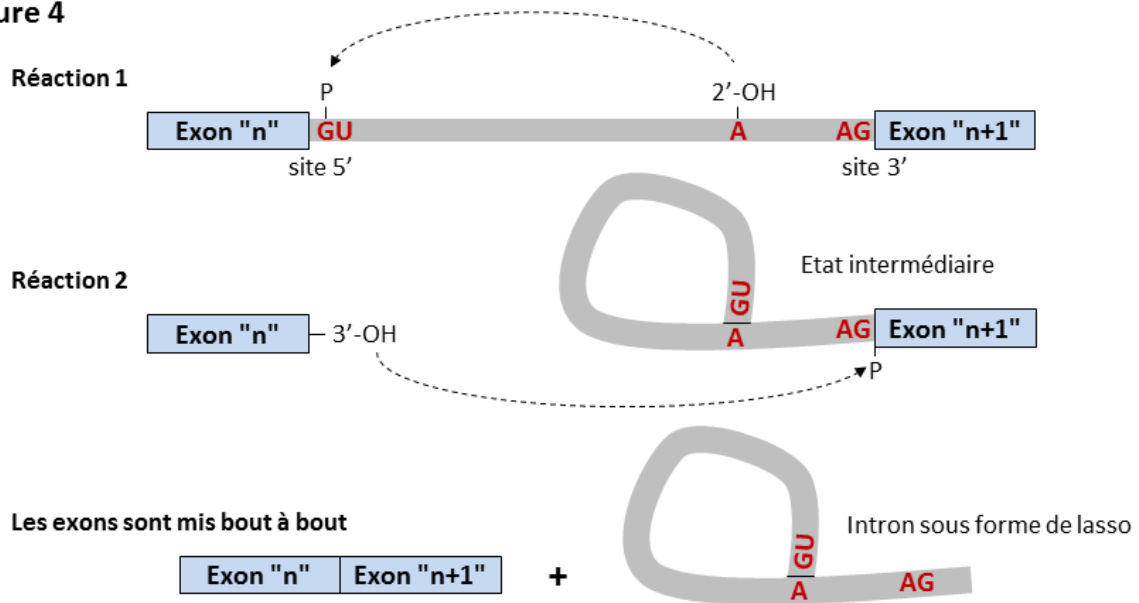
Ces séquences ont des caractéristiques (le point de branchement et la séquence riche polypyrimidique) et une localisation particulières aux extrémités de chaque intron.

La snRNP U1 s'associe au site 5' d'épissage et U2 se lie au point de branchement, le tout constitue le pré-spliceosome.

L'entrée des snRNP U4/U6 et U5 dans le complexe d'épissage s'accompagne d'une profonde réorganisation du complexe, qualifié alors de « mature », qui permet en son sein le rapprochement physique des sites d'épissage et la réaction catalytique proprement dite (**Figure 4**).

Deux réactions chimiques successives assurent la ligation des deux exons et l'élimination de l'intron sous forme d'un lasso. Il est intéressant de savoir que la réaction de l'épissage s'effectue d'une manière synchrone à la transcription du gène et que certains composants protéiques des deux machineries interagissent entre eux

Figure 4



L'épissage alternatif, une source de diversité des transcrits : les exons alternatifs

Si l'épissage correspond à l'élimination des introns et à la liaison des exons bout-à-bout dans l'ARNm mature, le choix de l'inclusion des exons peut être alternatif. L'épissage alternatif fournit de larges possibilités pour enrichir le transcriptome (l'ensemble des transcrits d'une cellule) et le protéome (l'ensemble de ses protéines), sans élargir pour autant le nombre de gènes, relativement limité. Grâce à l'inclusion alternative des exons, un seul pré-ARNm peut ainsi produire plusieurs ARNm matures codant pour différentes isoformes protéiques structurellement et fonctionnellement distinctives.

.En effet, les exons alternatifs ont des sites d'épissage plus faibles (moins bien reconnus par la machinerie d'épissage) que les exons « constitutifs » (qui sont systématiquement retrouvés dans toutes les formes d'ARNm matures).

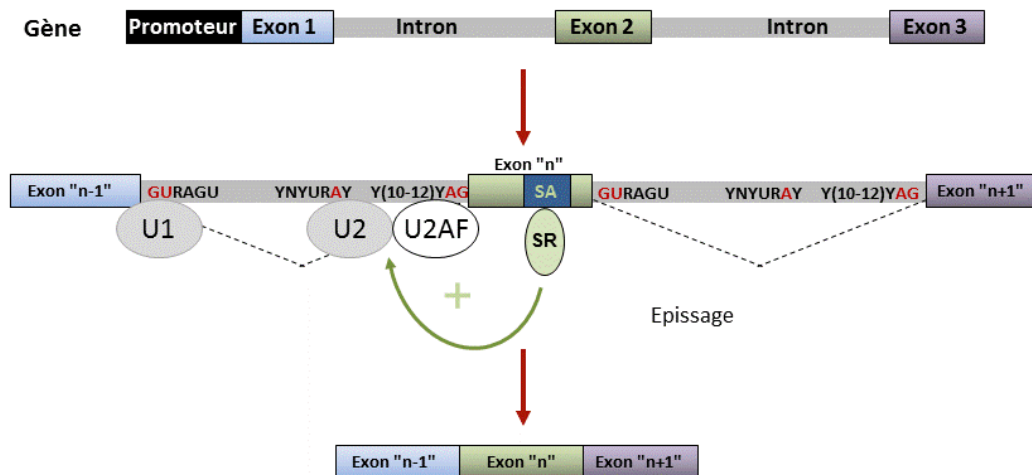
Néanmoins, il est important de noter que les sites d'épissage sont incapables, à eux seuls, d'initier l'assemblage des composants du spliceosome. En effet, d'autres séquences régulatrices, agissant en *cis* (c'est-à-dire au sein même de la molécule d'ARN, au contraire d'un facteur externe agissant en *trans*), ont été découvertes.

Ces séquences, qui peuvent être exoniques ou introniques, sont capables de favoriser ou d'inhiber la reconnaissance d'un exon à l'aide des facteurs d'épissage, dits régulateurs en *trans*.

Il existe deux grandes familles de facteurs d'épissage se liant aux séquences régulatrices en *cis* : les protéines riches en acides aminés serine et arginine, nommées protéines SR, et les protéines des particules ribonucléoprotéiques hétérogènes nucléaires (hnRNP).

Conceptuellement et de façon très simplifiée, les protéines SR, en se fixant sur les séquences activatrices (SA), favorisent la reconnaissance des sites d'épissage et facilitent le recrutement des snRNP U1 et U2 (**Figure 5**).

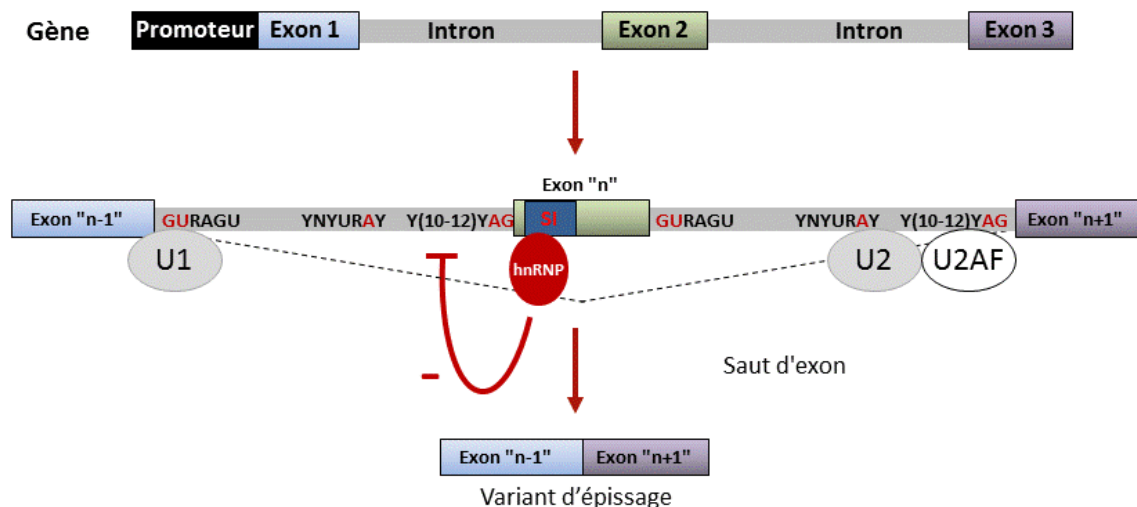
Figure 5



A l'inverse, les protéines hnRNP se lient sur les séquences inhibitrices (SI), jouant le rôle de répresseur en masquant les sites d'épissage (**Figure 6**).

L'inclusion ou l'exclusion d'un exon dépend donc de la liaison compétitive de ces différents facteurs à leurs séquences-cibles. D'autres conditions peuvent influencer le choix d'exons alternatifs telles que des modifications affectant les histones (les protéines qui soutiennent et protègent l'ADN dans le noyau), ou la vitesse à laquelle le complexe de transcription copie l'ADN d'un gène en ARN.

Figure 6



Les différents types d'épissage alternatif

L'épissage alternatif augmente significativement la diversité des transcrits grâce à l'utilisation alternative d'exons ou d'introns, mais aussi de signaux de polyadénylation situés dans des exons alternatifs. La majorité de ces événements se produit dans un maintien du cadre de lecture (la succession de trinuécléotides qui permet de décoder la séquence d'acides aminés de la protéine correspondante), résultant en l'expression de différentes isoformes protéiques.

Les exemples d'épissage alternatif peuvent être classés en cinq modèles différents (**Figure 7**) :

1. L'exon cassette (ou le saut d'exon) : il s'agit d'un événement où un exon est entièrement inclus ou retenu dans le transcrit mature. C'est le cas d'épissage alternatif le plus simple, qui prédomine chez les mammifères, donnant lieu à des transcrits plus ou moins longs.

2. Les exons mutuellement exclusifs : l'un des deux exons, qui sont généralement de taille similaire, est retenu dans l'ARNm mature, mais jamais les deux ensemble. Cet événement change peu la taille du transcrit mature mais modifie les propriétés de la protéine produite.

3. La rétention d'intron : un segment de pré-ARNm peut être épissé comme un intron ou simplement retenu. Cela se distingue du saut d'exon puisque la séquence intronique conservée n'est pas flanquée par des introns. Si l'intron ne modifie pas le cadre de lecture, il code pour des acides aminés comme les exons qui l'avoisinent, produisant ainsi une protéine d'un poids moléculaire supérieur.

Dans le cas contraire, il entraîne l'apparition d'un codon stop prématuré, ce qui provoque la dégradation du transcrit mature (voir également **Figure 13**). La rétention d'intron est l'événement d'épissage le plus rare chez les mammifères.

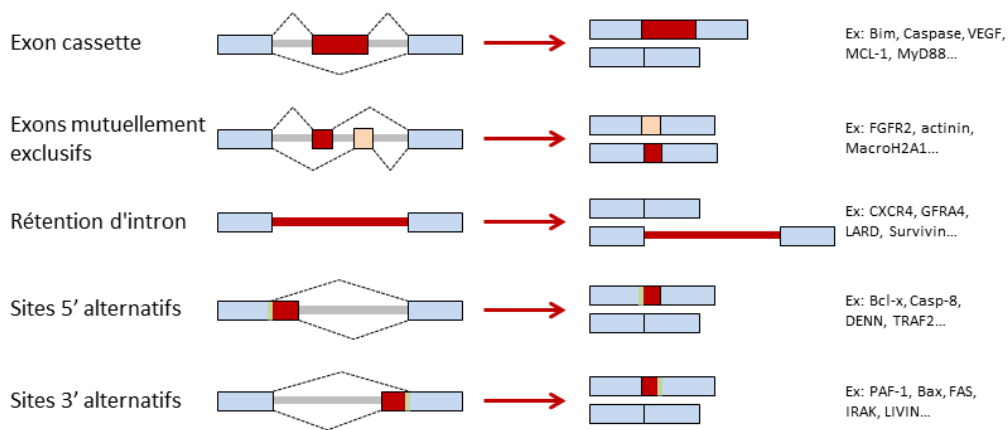
4. Site d'épissage alternatif en 5' (site donneur d'épissage) : ce variant d'épissage modifie la borne de l'exon dans sa partie aval.

5. Site d'épissage alternatif en 3' (site accepteur d'épissage) : d'une manière similaire au variant précédent, cet événement modifie la borne de l'exon dans sa partie amont, incluant ainsi un segment de l'extrémité 3' de l'intron épissé.

D'autres exemples peuvent être ajoutés aux 5 événements majeurs de l'épissage alternatif, tels que l'usage de promoteurs alternatifs ou de signaux alternatifs de polyadénylation, qui sont connus respectivement comme premier et dernier exon alternatif.

Toutefois, en soi ces variants ne sont pas des événements d'épissage puisque leur régulation dépend du choix des sites d'initiation et de la terminaison de la transcription. Néanmoins, ils augmentent également la diversité des ARNm matures générés par un seul gène.

Figure 7



L'épissage alternatif, une source de diversité fonctionnelle des produits des gènes.

Grâce à l'épissage alternatif, un gène peut générer de nombreux ARNm matures et par conséquent coder pour plusieurs isoformes protéiques.

De ce fait la définition d'un exon et d'un intron doit être actualisé on ne parle plus de séquences codantes ou non codante mais **de variants d'épissage**

En conclusion, l'ancien dogme basé sur l'hypothèse de partie codante et non codante (intron et exon n'est plus concevable.

