

Figure 1 : noyau purine et pyrimidine

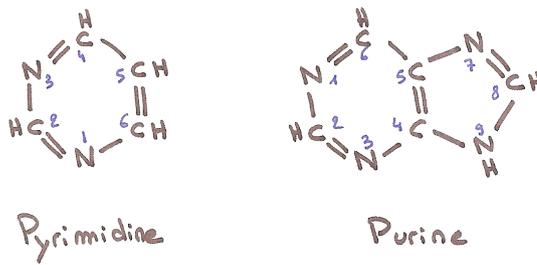


Figure 2 : les 5 bases azotées

(Luchetta et al., Dunod, 2005)

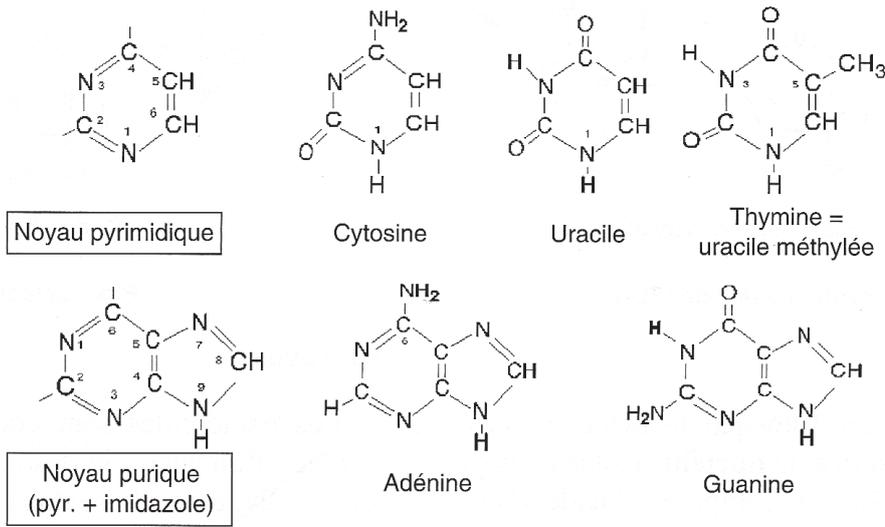


Figure 3 : propriétés des bases

Propriété	Conséquences
Bases non chargées ayant des gpmts polaires	
Riches en doubles liaisons	
Bases hydrophobes	

Figure 4 : les formes tautomères sont à l'origine d'appariements anormaux et donc de mutations (doc)

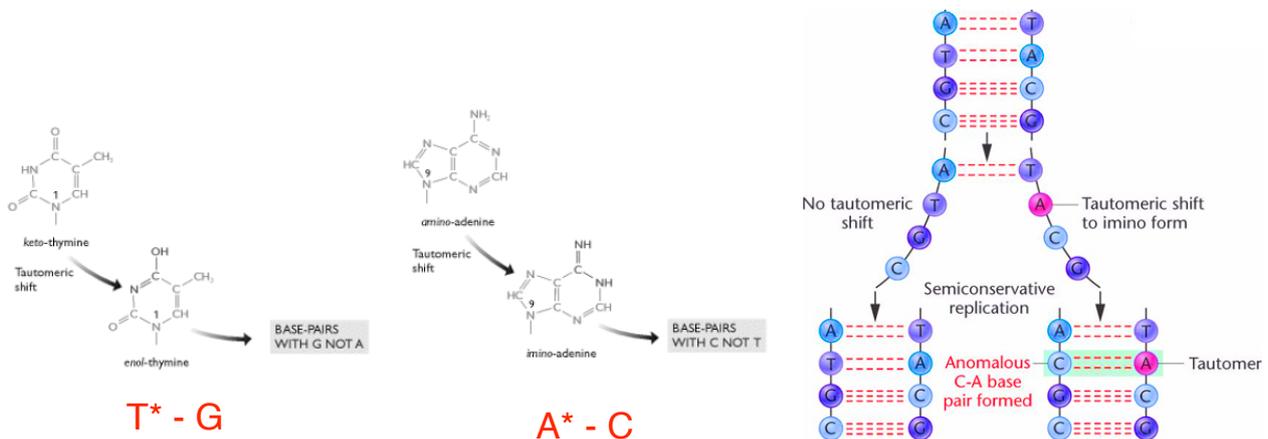
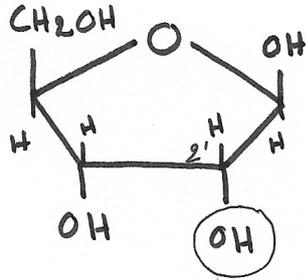


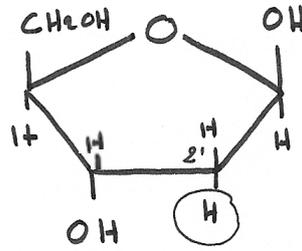
Figure 5 : formation des nucléosides (à savoir)

Deux pentoses possibles :

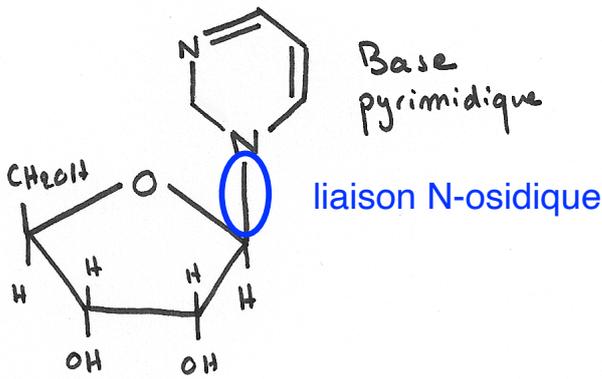
- **ribose** : on obtient un **ribonucléoside**(\rightarrow ARN) : avec A, G, C et U
- **désoxyribose** : on obtient un **désoxyribonucléoside**(\rightarrow ADN) avec A, G, C et T



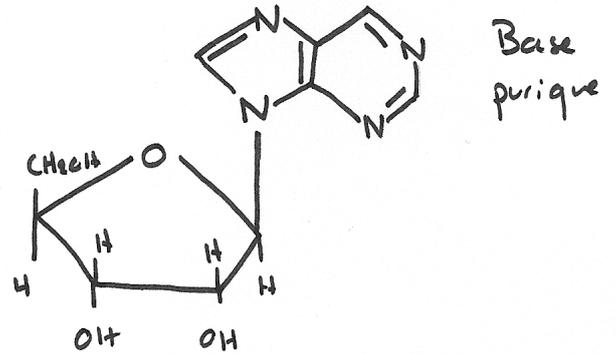
β -D. ribose
 \rightarrow ARN



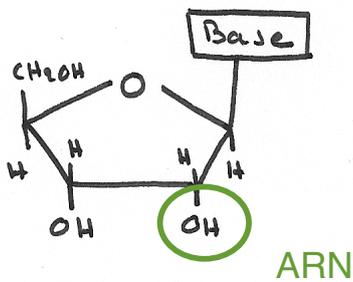
β -D. désoxyribose
 \rightarrow ADN



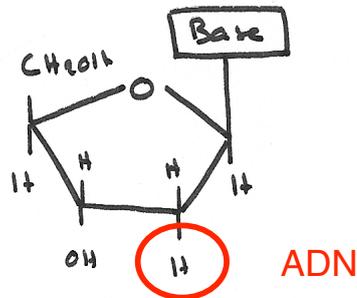
nucléoside



nucléoside



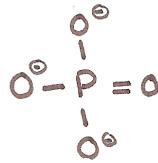
nucléoside



désoxynucléoside

Figure 6 : les nucléotides (à savoir)

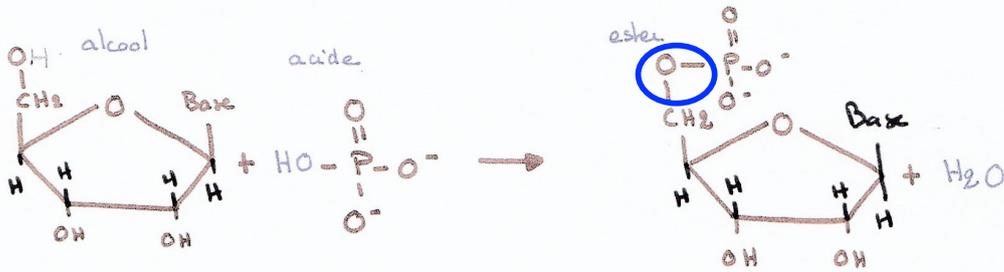
Phosphate inorganique : Pi



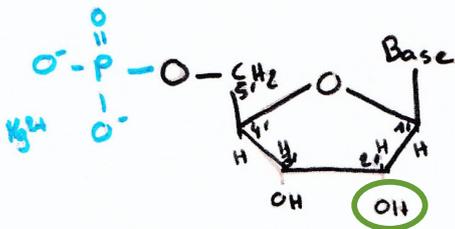
Pi = phosphate inorganique

Réaction d'estérification entre un groupement alcool et un phosphate inorganique

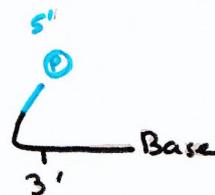
liaison phosphoester



Représentation d'un nucléotide



Nucléoside monophosphate
= nucléotide



Représentation schématique
d'un nucléotide

Les nucléotides de l'ADN et de l'ARN : nomenclature à connaître

Base azotée	nucléotide de l'ADN = desoxyribonucléotide	nucléotide de l'ARN = ribonucléotide
Adénine	desoxy adénosine monophosphate (dAMP)	adénosine monophosphate (AMP)
Guanine	desoxy guanosine monophosphate (dGMP)	guanosine monophosphate (GMP)
Thymine	desoxy thymidine monophosphate (dTMP)	
Cytosine	desoxy cytidine monophosphate (dCMP)	cytidine monophosphate (CMP)
Uracile		uridine monophosphate (UMP)

Figure 7 : ATP et AMPc (à savoir, AMPc → cours de spé communication cellulaire)

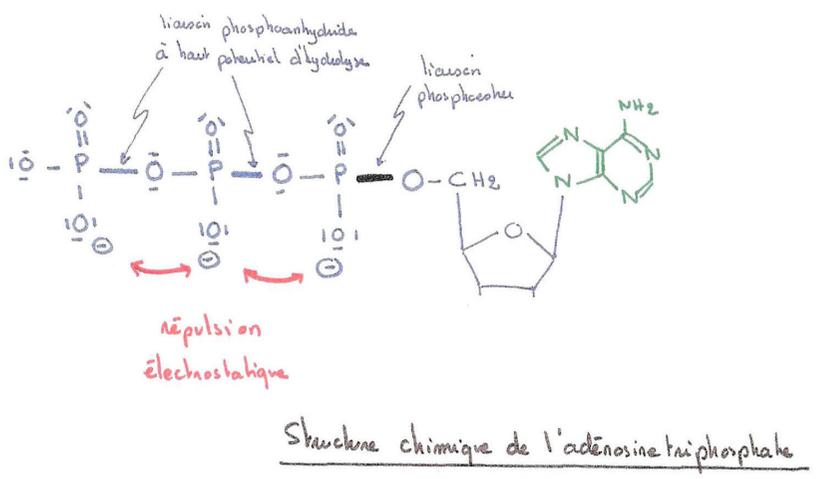
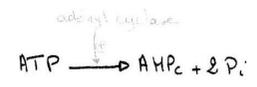
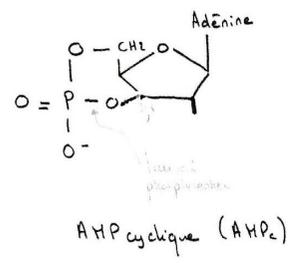
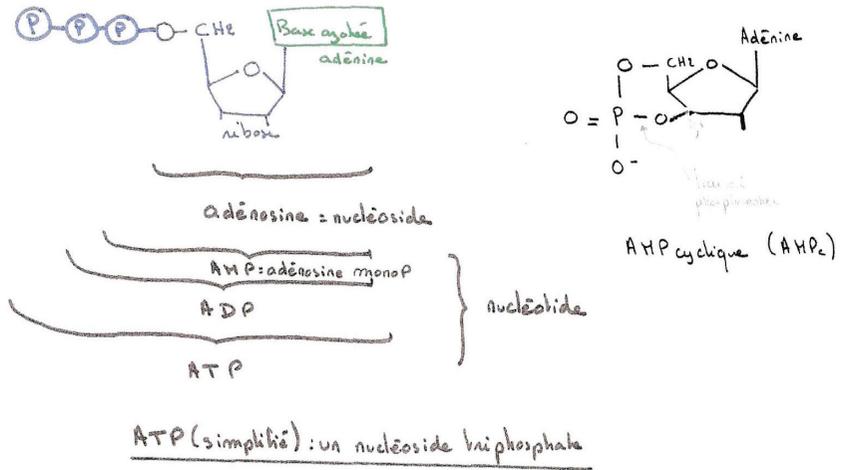
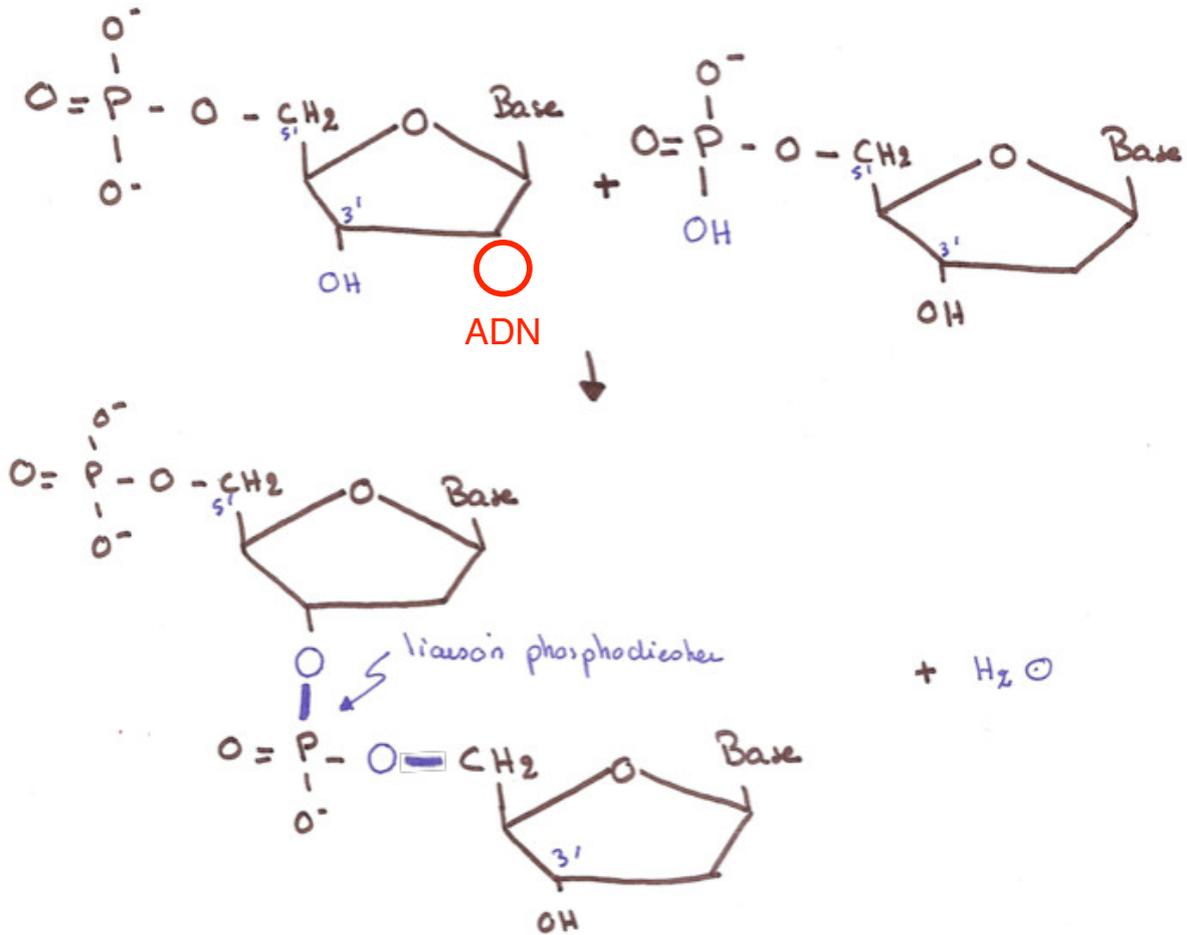


Figure 8 : formation d'une liaison phosphodiester : condensation avec élimination d'une molécule d'eau (à savoir)



Taille exprimée selon 3 unités :

- (la longueur)
- la masse moléculaire en **Daltons** : 1 paire de bases = environ 600 daltons
- le **nombre** de nt ou de leurs **bases** pour les molécules simple brin ou de paires de bases quand la molécule est constituée de 2 brins. Le multiple le plus courant est le **kilobase kb** (= 1000 pb ou b).

ADN	ARN
<p>Selon règnes et espèces : de 5000 à plus de 100 millions de paires de bases (pb ou bp) :</p> <ul style="list-style-type: none"> - masse moyenne d'une bp = 600 Da - 1 mm = 3 000 bp = 2. 10⁶ Da (2m dans une cellule humaine !!!) 	<p>Nombre de nt allant de plusieurs dizaines à plusieurs milliers :</p> <ul style="list-style-type: none"> - ARNr : une centaine à 5 000 - ARNt : 75-90 - ARNm : en fonction du gène codant

**Figure 9 : polymérisation de nucléotides (ADN ici)
(in Lodish)**

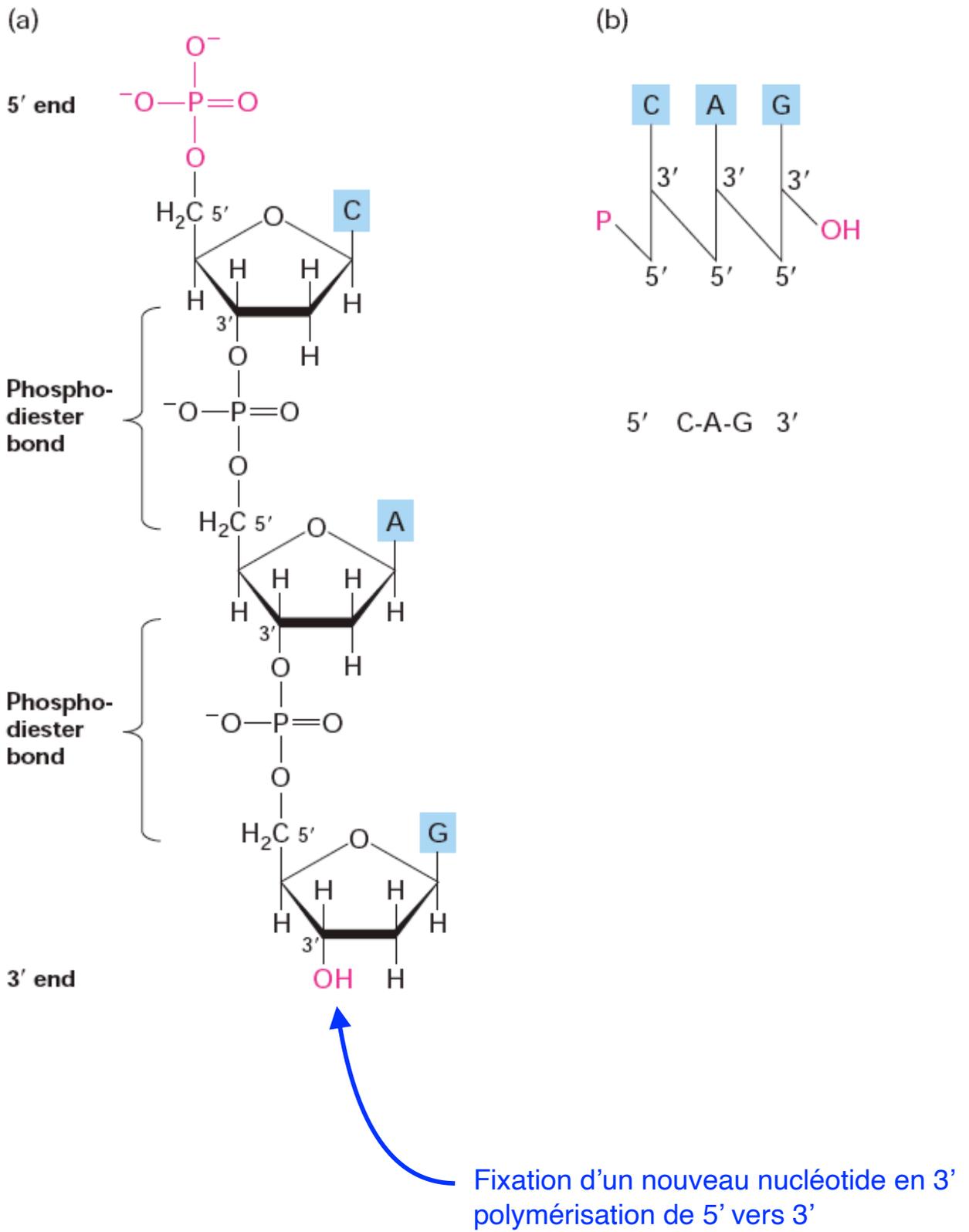
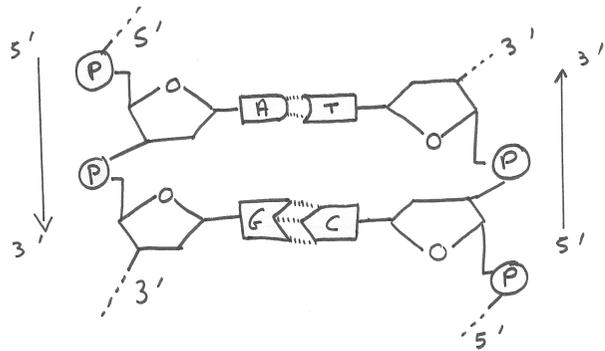
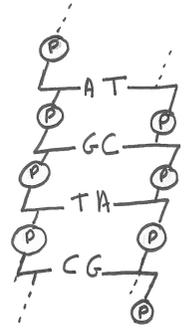


Figure 10 : structure de la molécule d'ADN



Représentation schématisée des deux brins d'ADN en plan



Schématisation

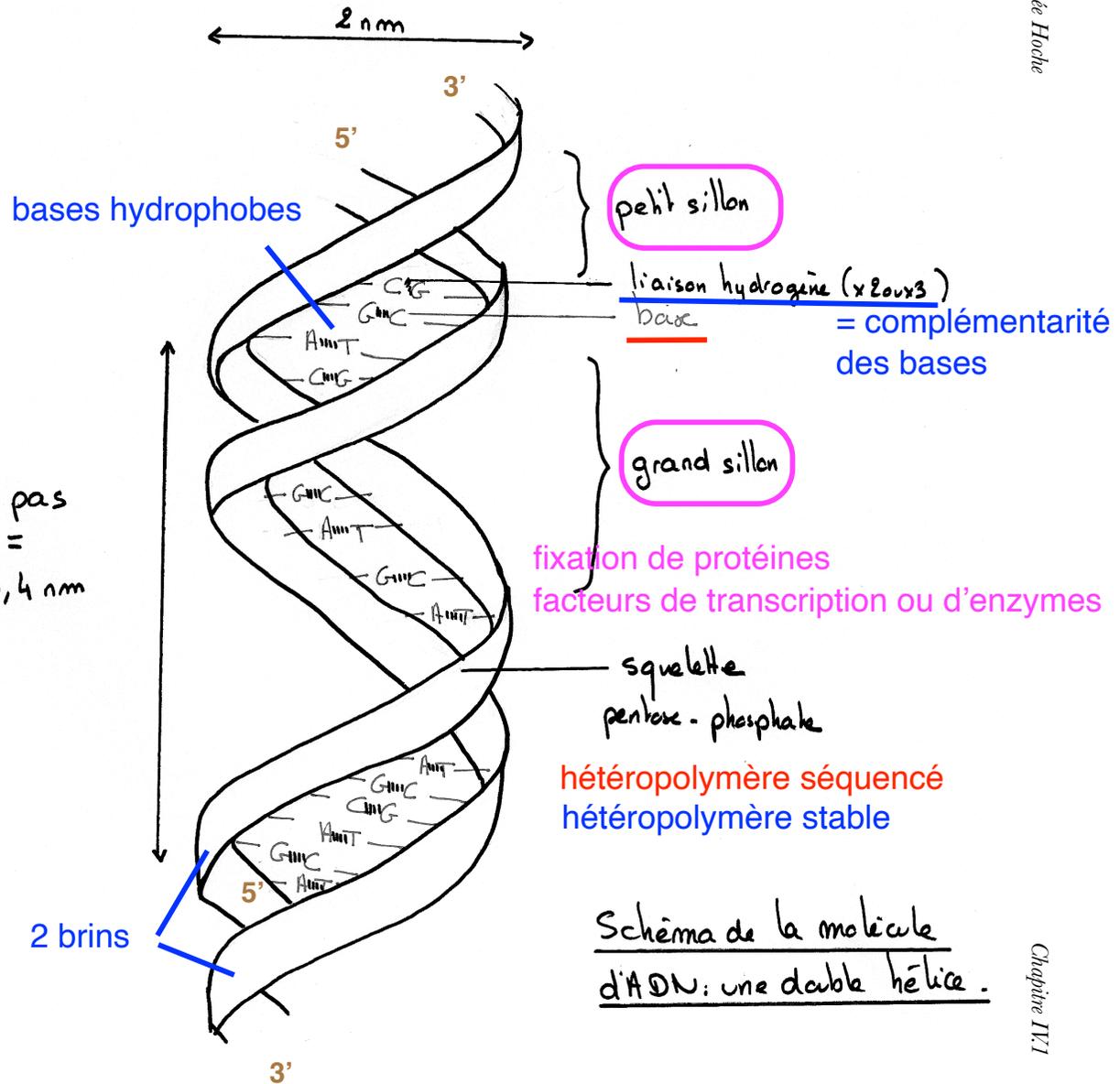
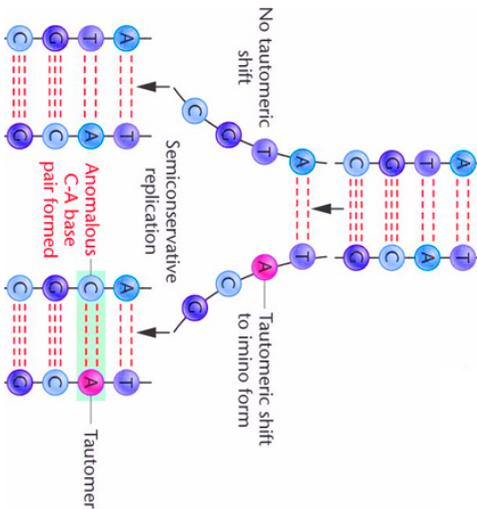
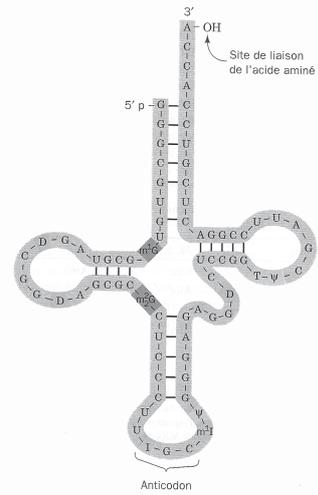
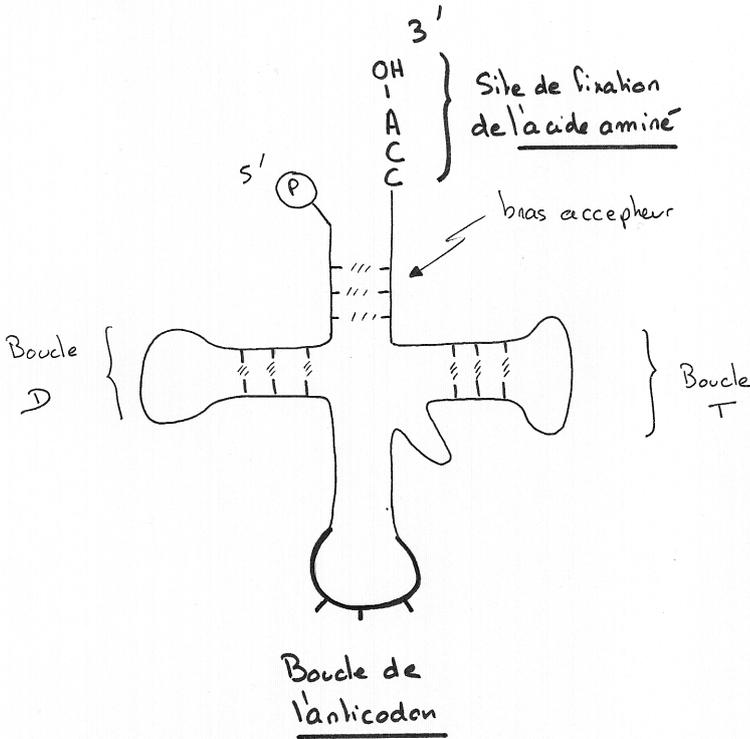


Schéma de la molécule d'ADN: une double hélice.

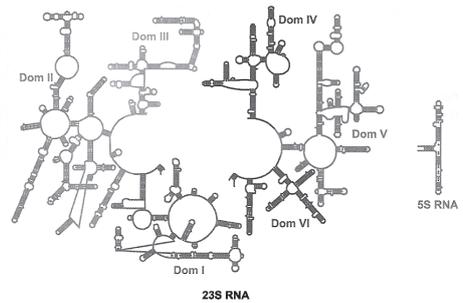
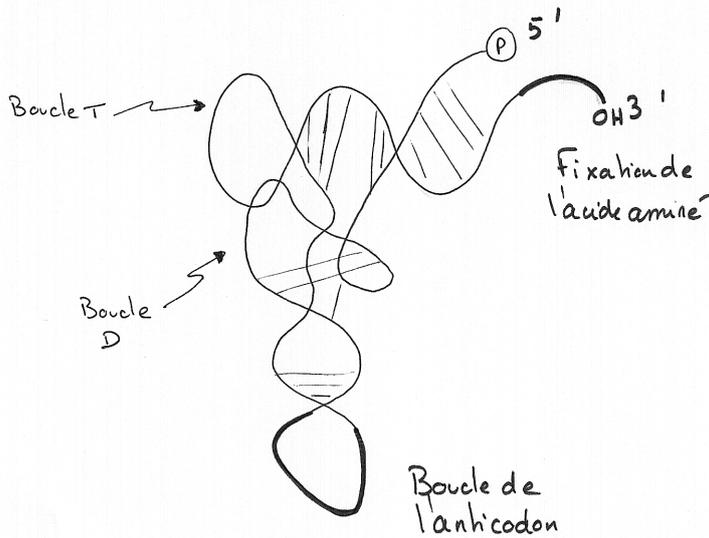
Figure 11 : structure secondaire des ARNt et ARNr (importance des structures tige-boucle) et tertiaire

Structure secondaire d'un ARNt



La séquence des bases de l'ARN^{Phe} de levure selon la représentation en feuille de trèfle. Les symboles des nucléosides modifiés (en couleur) sont expliqués dans la Fig. 32-10.

Structure tertiaire d'un ARNt

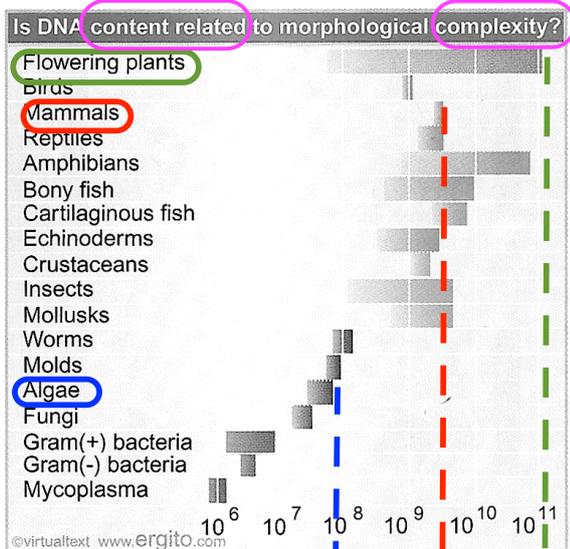


Structure 2D en feuille de trèfle et 3D d'un ARNt (in Voet et Voet, De Boeck, 2005)

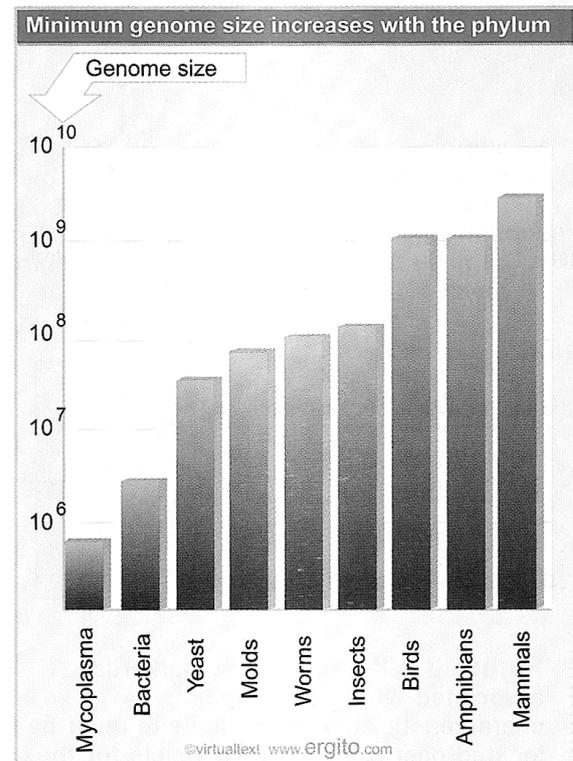
Structure 2D et 3D de l'ARNr 23s (in Voet et Voet, De Boeck, 2005)

Figure 12 : Taille des génomes de quelques êtres vivants procaryotes (*Mycoplasma* et *E. coli*) et eucaryotes (les autres)

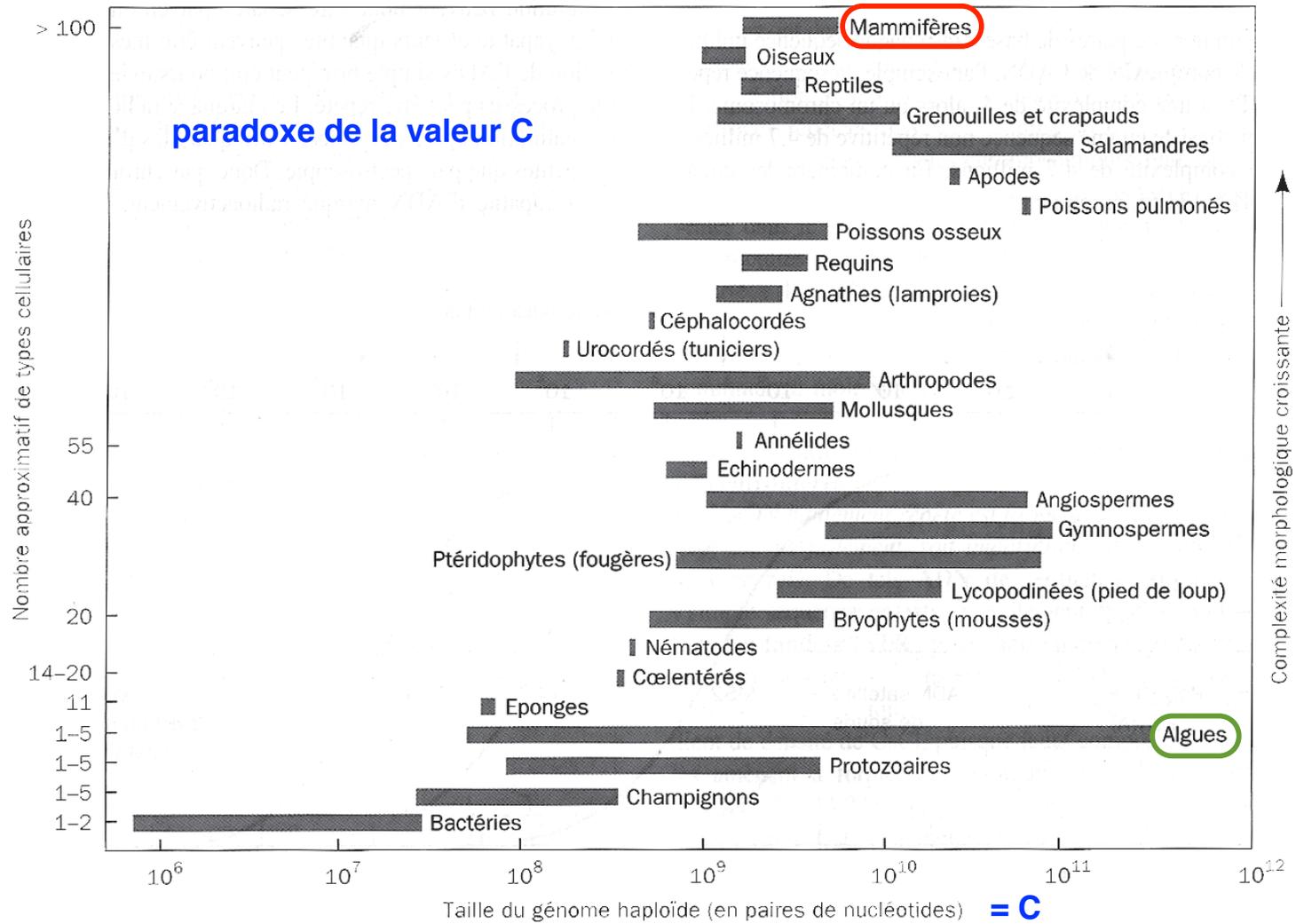
Organisme	Nombre de paires de bases	Nombre de gènes estimés	Densité génique (Gènes/Mb)
<i>Mycoplasma</i>	0,6 . 10 ⁶	450	860
<i>E. coli</i>	4 . 10 ⁶	4 000	950
<i>Saccharomyces cerevisiae</i>	12 . 10 ⁶	6 000	480
<i>Arabidopsis thaliana</i>	142 . 10 ⁶	26 000	220
<i>Caenorhabditis elegans</i>	97 . 10 ⁶	19000	190
<i>Drosophila melanogaster</i>	137 . 10 ⁶	14 000	82
<i>Homo sapiens</i>	3 . 10 ⁹	20 à 25 000	6.25



DNA content of the haploid genome is related to the morphological complexity of lower eukaryotes, but varies extensively among the higher eukaryotes. The range of DNA values within a phylum is indicated by the shaded area.



The minimum genome size found in each phylum increases from prokaryotes to mammals.



Tailles comparées des génomes haploïdes de différentes catégories d'organismes mettant en évidence le paradoxe de la valeur C. La complexité morphologique des organismes, estimée par le

nombre de types cellulaires, est croissante du bas de la figure vers le haut. [D'après Raff, R.A. et Kaufman, T.C. *Embryos, Genes, and Evolution*, p. 314, Macmillan (1983).]

Taille des génomes haploïdes
(in Voet et Voet, De Boeck, 2005)

Figure 13 : caractéristiques des 3 ARN

	ARNr	ARNt	ARNm
Localisation cellulaire	Intégrés dans le ribosome dans le cytoplasme	Cytoplasme	Cytoplasme
Abondance (% des ARN cytoplasmiques totaux)	71 %	15 %	5 %
Taille	Les plus grands	petits (75-95 nt)	variables
Structure	Tige-boucle → structure III complexe	Tige-boucle → trèfle	linéaire
Stabilité	Stables	Stables	Instables (qq min à qq h)

Figure 14 : structure 3D d'un ribosome procaryote

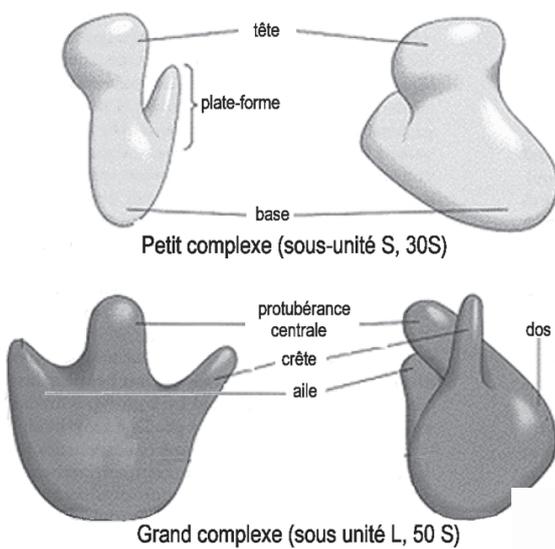


Figure 15 : les protéines sont synthétisées à partir d'ARNm
(in Bordas, programme de 1ère S)

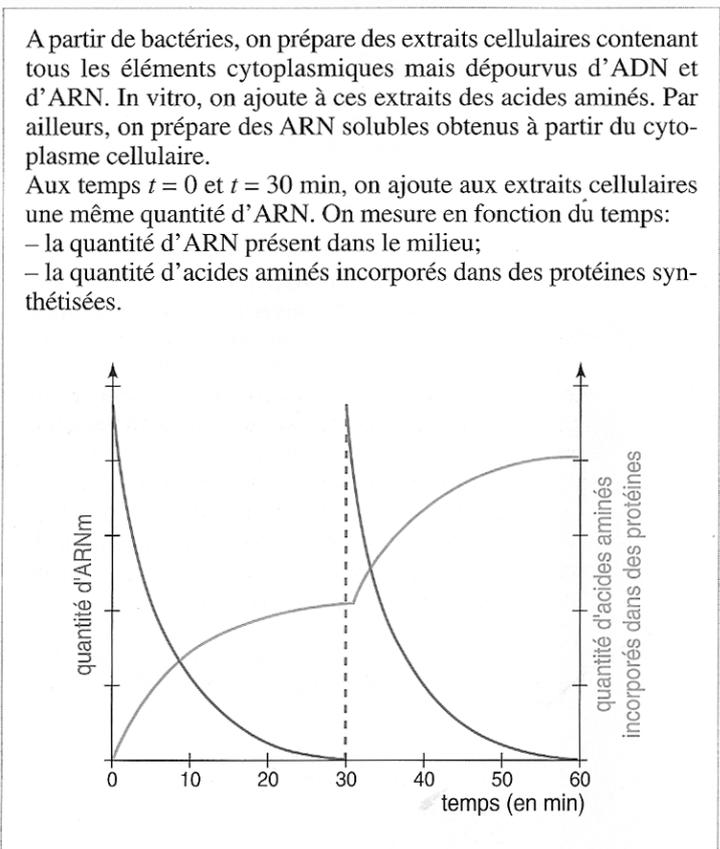


Figure 16 : l'ARN migre du noyau vers le cytoplasme
(in Bordas, programme de 1ère S)
Expérience à connaître parfaitement et à savoir dessiner!

Les deux photographies ci-dessous présentent des autoradiographies* de cellules qui ont été cultivées en présence d'un précurseur radioactif spécifique de l'ARN. Chaque tache noire repère un endroit où se trouve de l'ARN ayant incorporé le précurseur radioactif :

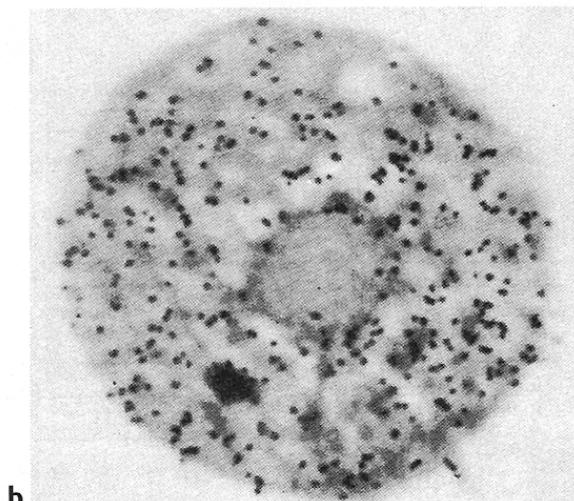
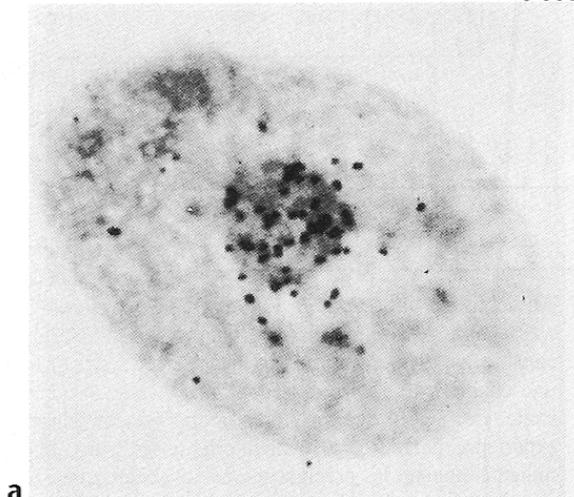
Le **cliché a** présente l'autoradiographie d'une cellule après 15 minutes de culture sur un milieu contenant le précurseur radioactif de l'ARN.

Le **cliché b** correspond à l'autoradiographie d'une cellule qui a été d'abord cultivée pendant 15 minutes sur un milieu contenant un précurseur radioactif de l'ARN puis pendant une heure et demie sur un milieu contenant des précurseurs non radioactifs de l'ARN.

*Autoradiographie : technique qui permet de localiser un élément radioactif par mise en contact et impression d'un support photographique

Représentation schématique de l'expérience :

× 6 000



Principe de l'expérience (préciser le précurseur utilisé):

Résultats et interprétation :

Figure 17 a: Découverte de la correspondance entre la séquence de l'ADN et la séquence des protéines

Exemple de la Drépanocytose (rappels du lycée) : travaux d'Ingram

Ingram, vers **1956** a été l'un des premiers à établir la relation précise entre :

- une **mutation** bien caractérisée ayant comme manifestation phénotypique une maladie héréditaire grave : **l'anémie falciforme** ou drépanocytose
- ET
- une **protéine** spécifique : **l'hémoglobine**.

Par des méthodes d'analyse basée sur de la chromatographie, il montra que l'hémoglobine S diffère de la normale A par un seul acide aminé (le sixième) dans les chaînes β : **l'acide glutamique est remplacé par de la valine** (Glu \rightarrow Val). C'est la première fois que l'on fit le lien entre une mutation et la modification de la séquence en acides aminés d'une protéine.

✓ **A utiliser en colles ou à l'écrit pour montrer que la séquence en nucléotides code pour la séquence en acides aminés.**

Exemple chez les Procaryotes (*E. coli*) : travaux de Yanofsky

En 1966, **Yanofsky** montra que les positions des mutations dans un gène correspondent parfaitement aux positions des acides aminés modifiés dans la séquence de la protéine codée.

Par les techniques de cartographie à haute résolution adaptées aux bactéries, il positionne plusieurs mutations supposées ponctuelles dans une séquence codant pour une enzyme (sous unité A de la tryptophane synthétase) d'*E. coli*. Il parvient également à séquencer la protéines synthétisée à partir du gène sauvage et des gènes mutés.

Yanofsky peut ainsi confronter la carte des mutations dans la séquence protéique à la carte des mutations dans le gène. : les cartes correspondent!

Il montre ainsi qu'il existe une excellente **colinéarité entre la séquence de l'ADN et la séquence de la protéine**.

Figure 17 b : détermination du code génétique (in Griffith et al., De Boeck 2002)

On synthétise un ARNm en mettant des proportions connues de nucléotides : on peut donc déterminer la probabilité de formation d'un codon. Si un codon code pour un acide aminé, les fréquences des aa dans le peptide synthétisé doivent correspondre aux fréquences des codons

Les fréquences attendues de différents codons dans des ARNm synthétiques constitués de $\frac{3}{4}$ d'uracile et de $\frac{1}{4}$ de guanine

Codon	Probabilité	Rapport*
UUU	$p(\text{UUU}) = \frac{3}{4} \times \frac{3}{4} \times \frac{3}{4} = \frac{27}{64}$	1,00
UUG	$p(\text{UUG}) = \frac{3}{4} \times \frac{3}{4} \times \frac{1}{4} = \frac{9}{64}$	0,33
UGU	$p(\text{UGU}) = \frac{3}{4} \times \frac{1}{4} \times \frac{3}{4} = \frac{9}{64}$	0,33
GUU	$p(\text{GUU}) = \frac{1}{4} \times \frac{3}{4} \times \frac{3}{4} = \frac{9}{64}$	0,33
UGG	$p(\text{UGG}) = \frac{3}{4} \times \frac{1}{4} \times \frac{1}{4} = \frac{3}{64}$	0,11
GGU	$p(\text{GGU}) = \frac{1}{4} \times \frac{1}{4} \times \frac{3}{4} = \frac{3}{64}$	0,11
GUG	$p(\text{GUG}) = \frac{1}{4} \times \frac{3}{4} \times \frac{1}{4} = \frac{3}{64}$	0,11
GGG	$p(\text{GGG}) = \frac{1}{4} \times \frac{1}{4} \times \frac{1}{4} = \frac{1}{64}$	0,03

* La fréquence de UUU est utilisée comme la fréquence de base par rapport à laquelle les fréquences des autres codons sont mesurées pour établir les rapports respectifs. Le rapport pour UUG, par exemple, est calculé par : $p(\text{UUG})/p(\text{UUU}) = 0,33$.

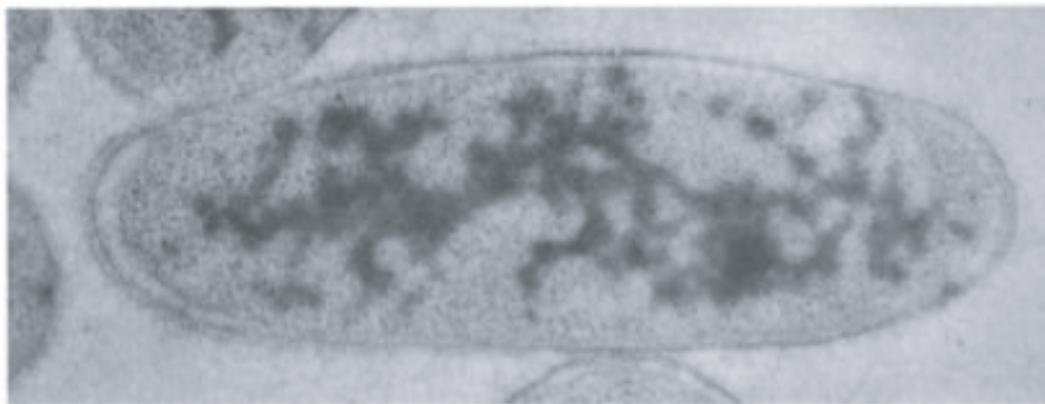
Figure 18 : le code génétique et quelques exceptions à son universalité
(in Peycru et al., Dunod, 2005)

		seconde base du codon			
		U	C	A	G
première base du codon	U	UUU Phe UUC Phe UUA Leu UUG Leu	UCU Ser UCC Ser UCA Ser UCG Ser	UAU Tyr UAC Tyr UAA Stop UAG Stop	UGU Cys UGC Cys UGA Stop UGG Trp
	C	CUU Leu CUC Leu CUA Leu CUG Leu	CCU Pro CCC Pro CCA Pro CCG Pro	CAU His CAC His CAA Gln CAG Gln	CGU Arg CGC Arg CGA Arg CGG Arg
	A	AUU Ile AUC Ile AUA Ile AUG Ile	ACU Thr ACC Thr ACA Thr ACG Thr	AAU Asn AAC Asn AAA Lys AAG Lys	AGU Ser AGC Ser AGA Arg AGG Arg
	G	GUU Val GUC Val GUA Val GUG Val	GCU Ala GCC Ala GCA Ala GCG Ala	GAU Asp GAC Asp GAA Glu GAG Glu	GGU Gly GGC Gly GGA Gly GGG Gly

DES CODES GÉNÉTIQUES PARTICULIERS.

Codons	Code universel	Code mitochondrial		
		Mammifères	Levure	Végétaux
UGA	Stop	Tryptophane	Tryptophane	Stop
AGG et AGA	Arginine	Stop	Arginine	Arginine
AUA	Iso-leucine	Méthionine	Méthionine	Iso-leucine
CUA	Leucine	Leucine	Thréonine	Leucine

Figure 19 : le nucleoïde bactérien



E. coli marquée avec un anticorps anti ADN.

Figure 20 : représentation schématique du génome bactérien d'Escherichia coli (bactérie gram-)

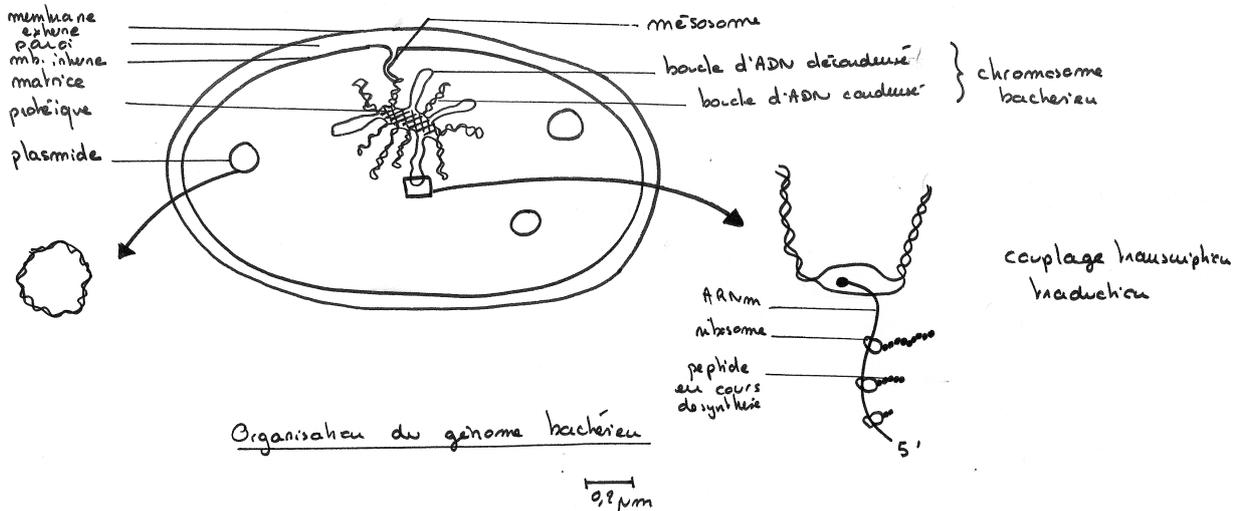
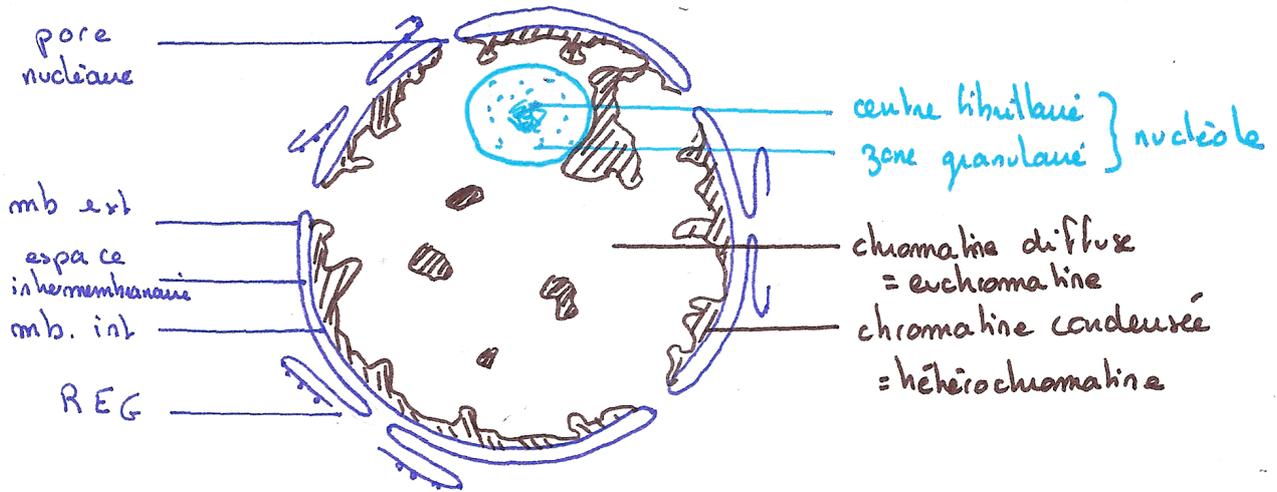


Figure 21 : organisation du noyau eucaryote



Noyau - nucléole et chromatine (1µm)

Figure 22 : coopération entre génome nucléaire et génome chloroplastique pour la synthèse de la RubisCO

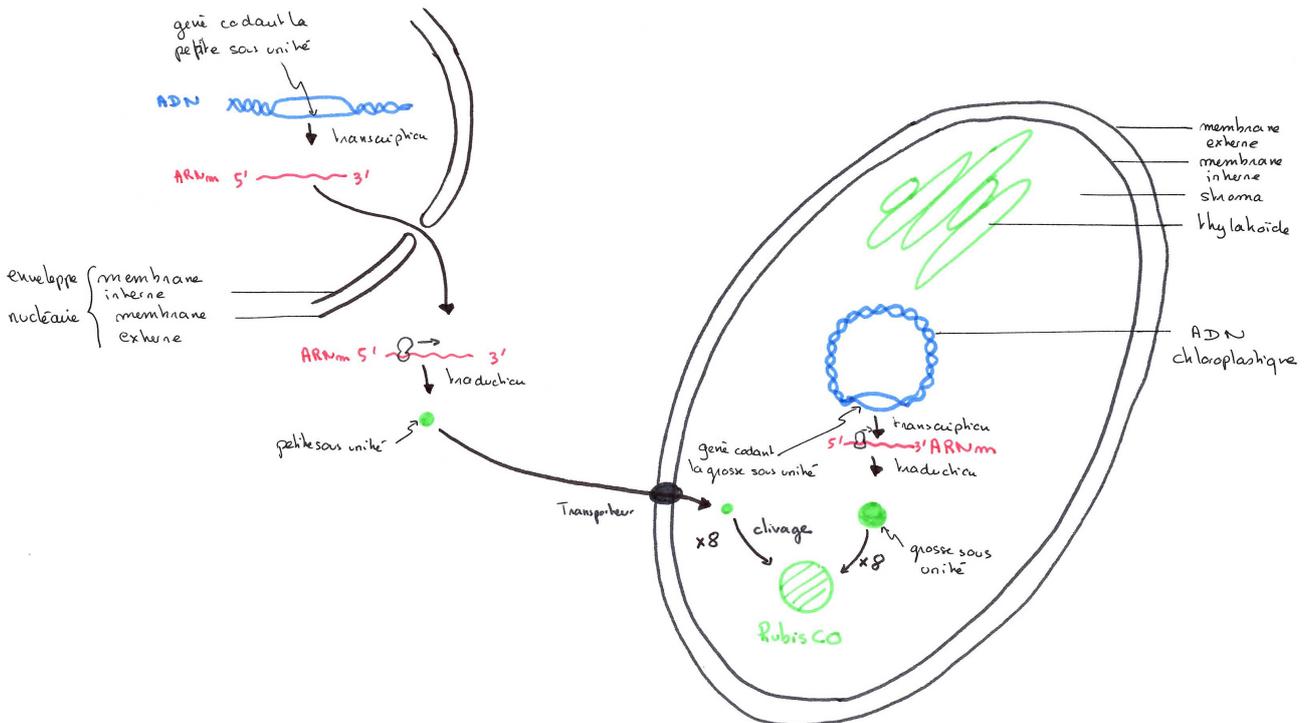


Figure 23 : les plasmides, des molécules d'ADN facultatives conférant un avantage sélectif

Type de plasmide	Propriétés
Plasmide de résistance	Résistance à un antibiotique
Plasmide de fertilité	Conjugaison et transfert d'ADN
Plasmide killer	Synthèse de toxines tuant d'autres bactéries
Plasmide métabolique	Protéines impliquées dans des phénomènes de transport de molécules
Plasmide de virulence	Pathogénicité

Figure 24 : fibre nucléosomique

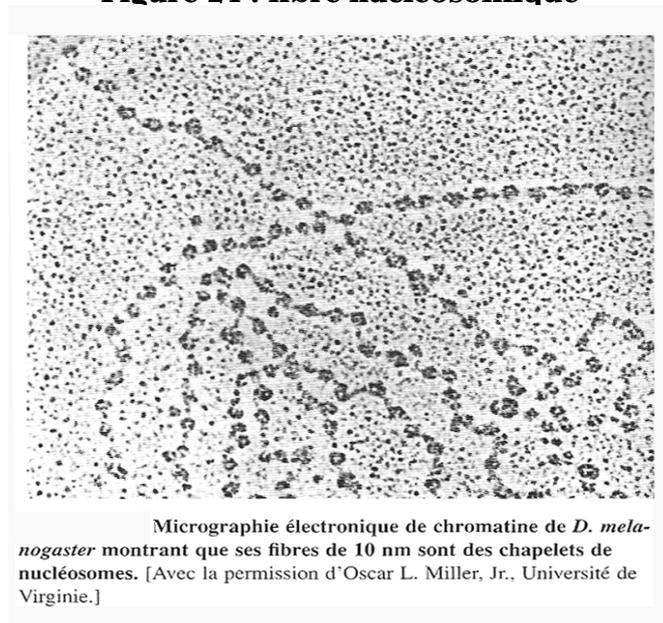
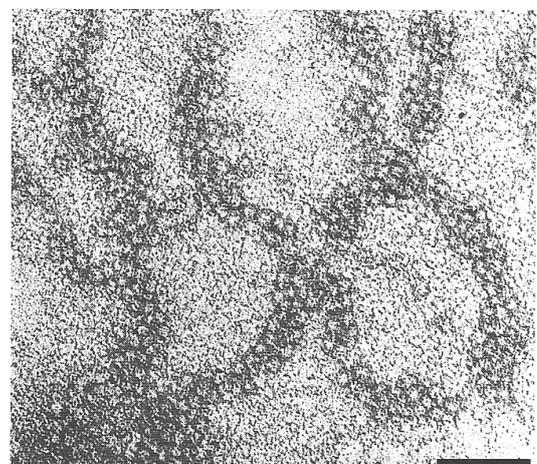


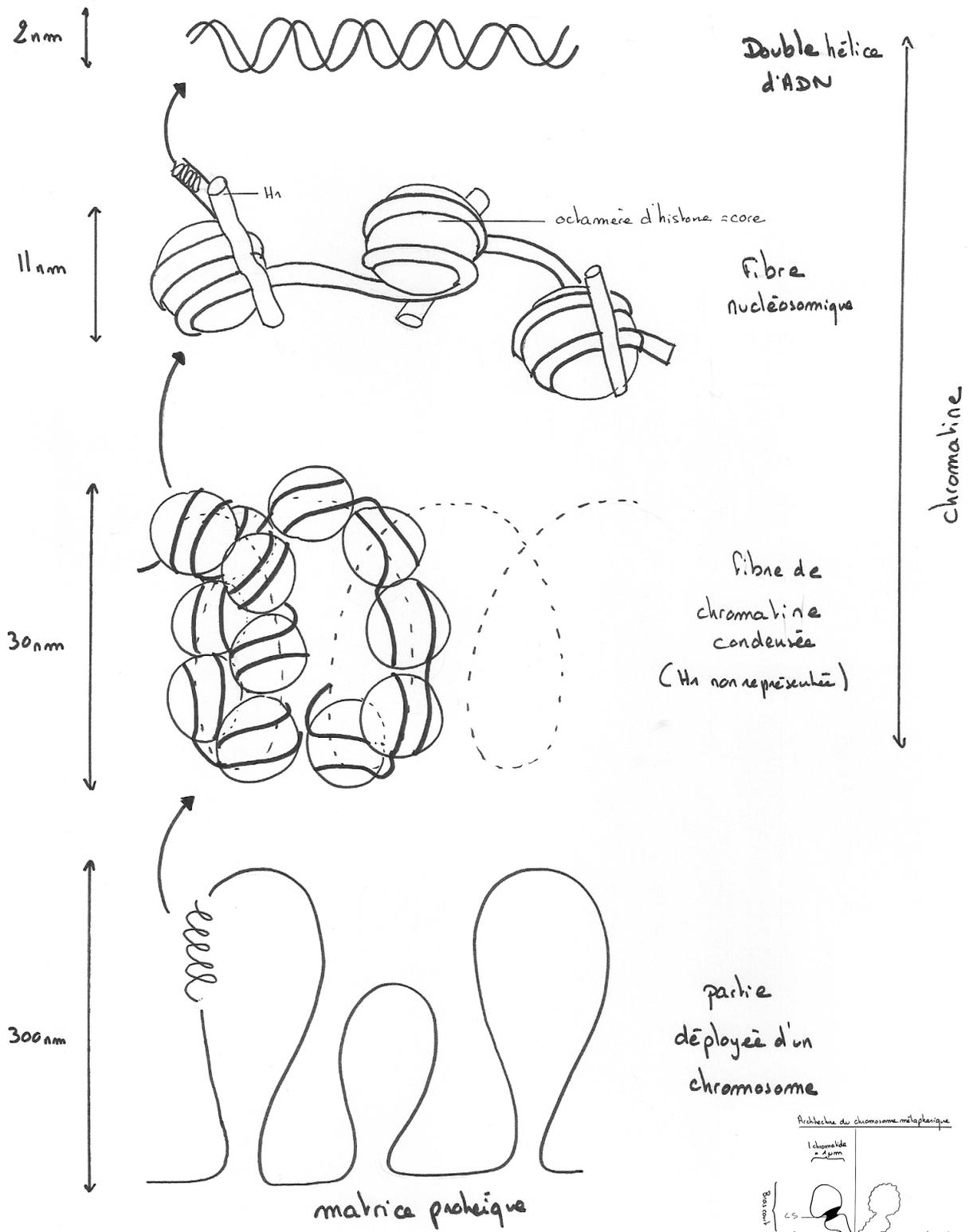
Figure 25 : : fibre de chromatine de 30 nm
(in Voet et Voet, De Boeck, 2005)

Décrire le cliché de MET

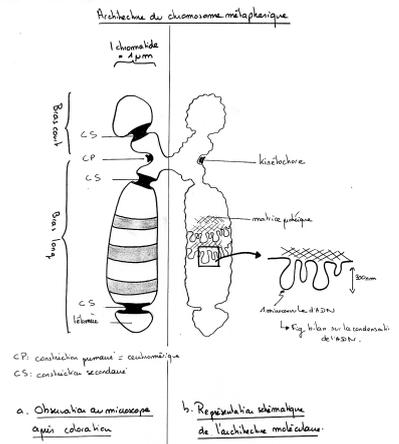


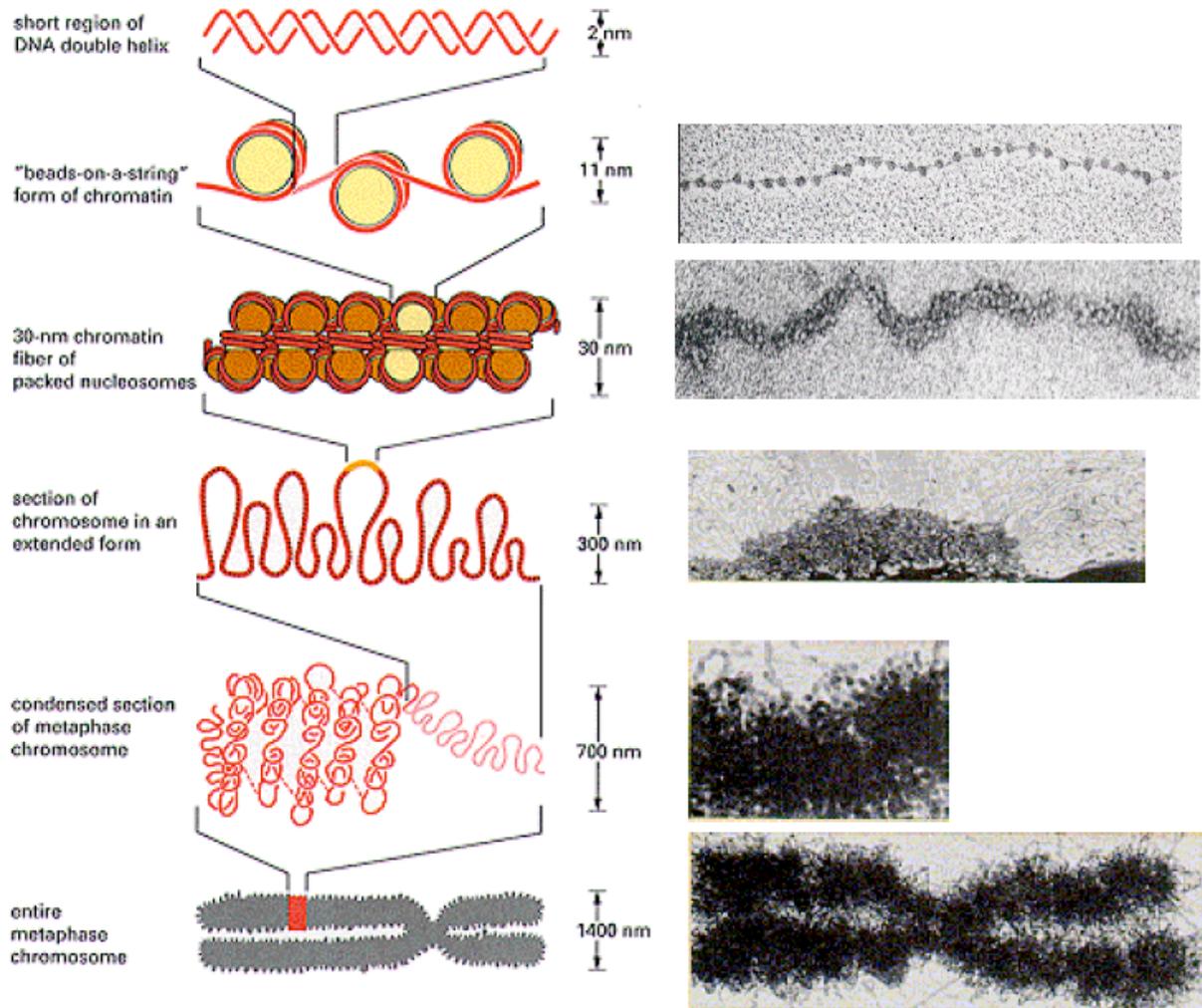
Micrographie électronique des filaments de chromatine de 30 nm. Notez que les filaments ont un diamètre correspondant à deux ou trois nucléosomes. La barre représente 1000 Å. [Avec la permission de Jérôme B. Rattner, Université de Calgary, Canada.]

Figure 26a : Bilan sur la condensation de la chromatine



partie déployée d'un chromosome





► **FIGURE 10-1 Overview of the structure of genes and chromosomes.**

DNA of higher eukaryotes consists of unique and repeated sequences. Only ~5% of human DNA encodes proteins and functional RNAs and the regulatory sequences that control their expression; the remainder is merely spacer DNA between genes and introns within genes. Much of this DNA, ~50% in humans, is derived from mobile DNA elements, genetic symbiots that have contributed to the evolution of contemporary genomes. Each chromosome consists of a single, long molecule of DNA up to ~280 Mb in humans, organized into increasing levels of condensation by the histone and nonhistone proteins with which it is intricately complexed. Much smaller DNA molecules are localized in mitochondria and chloroplasts.

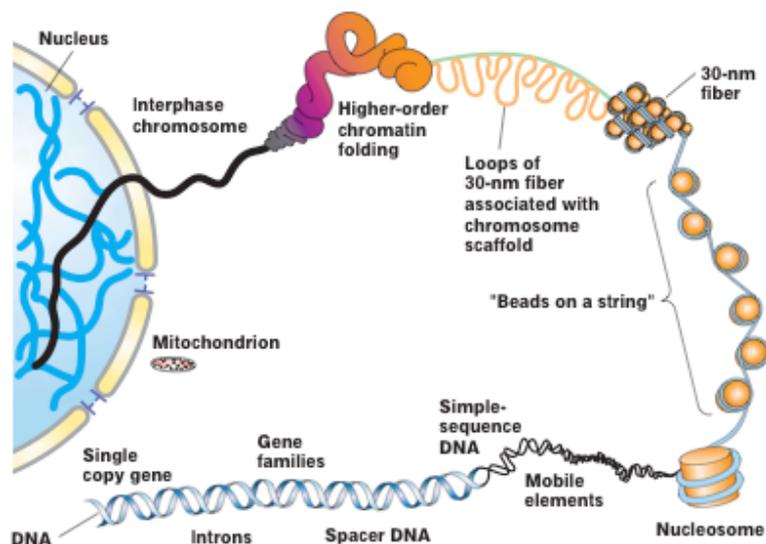


Figure 26c : Organisation du chromosome au cours du cycle cellulaire

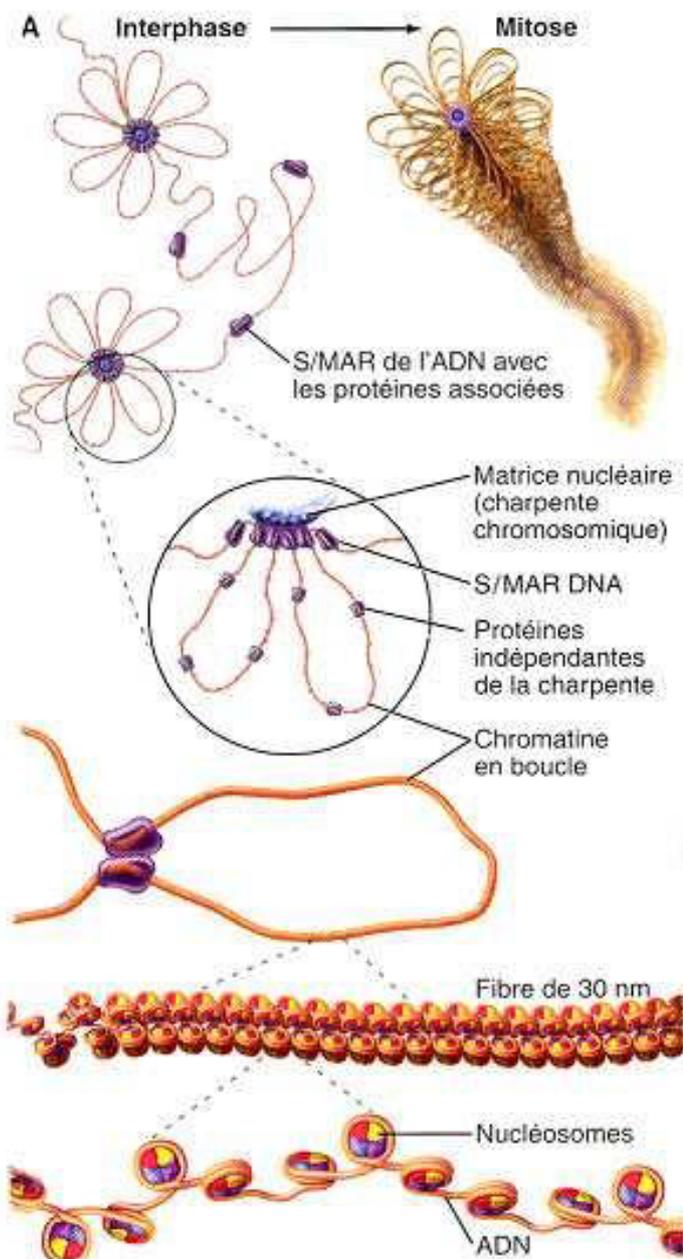
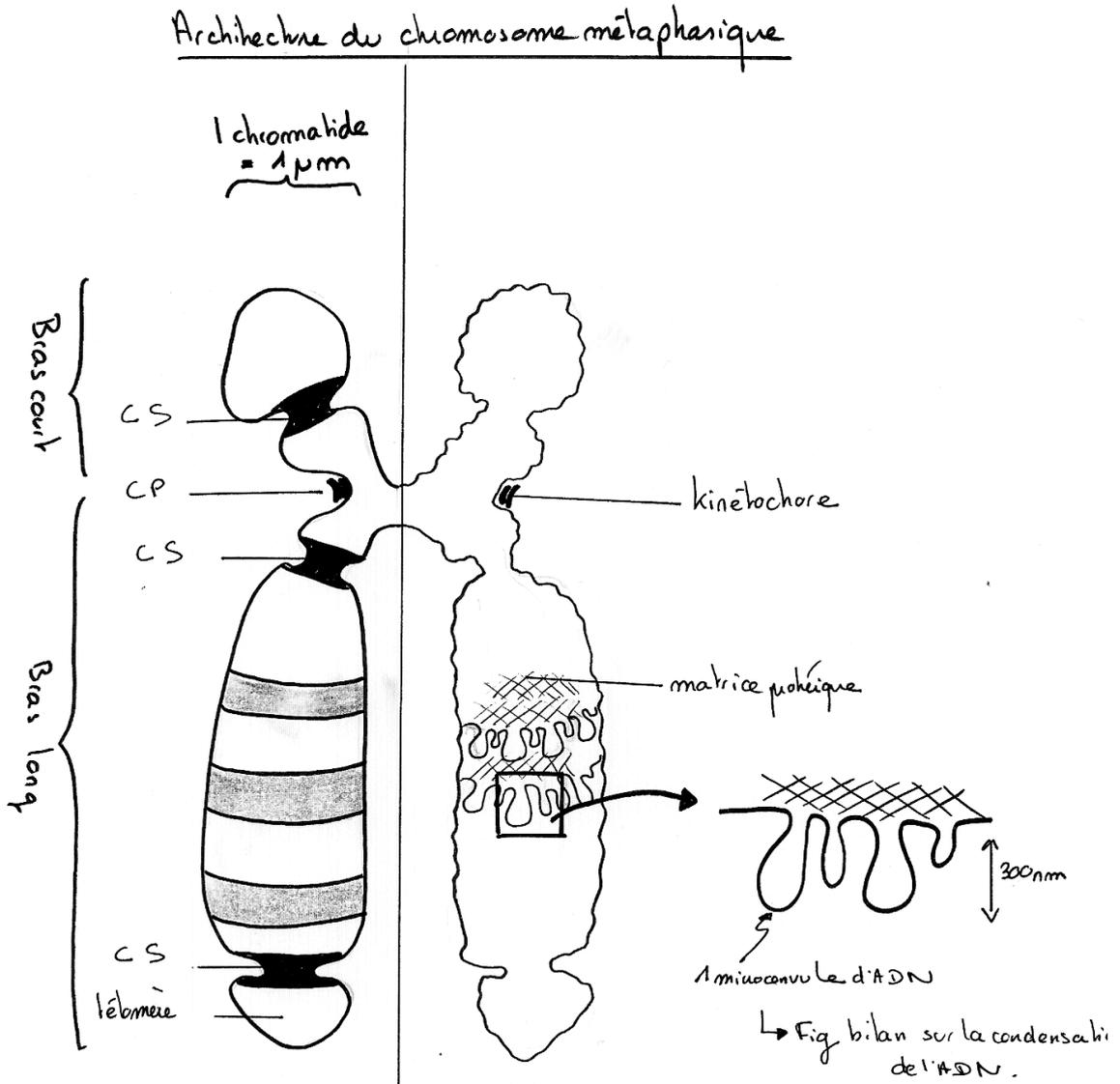


Figure 27a : le chromosome métaphasique des Eucaryotes



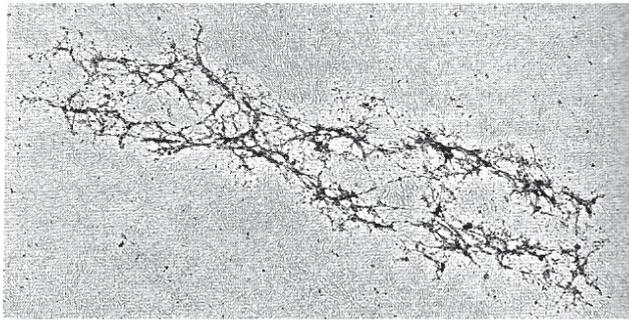
CP: constriction primaire = centromérique

CS: constriction secondaire

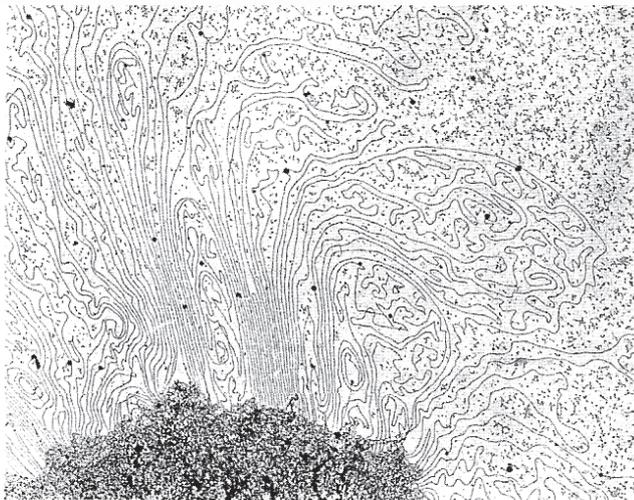
a. Observation au microscope
après coloration

b. Représentation schématique
de l'architecture moléculaire.

Figure 27b : : chromosome métaphasique observé au MET
(in Voet et Voet, De Boeck, 2005)

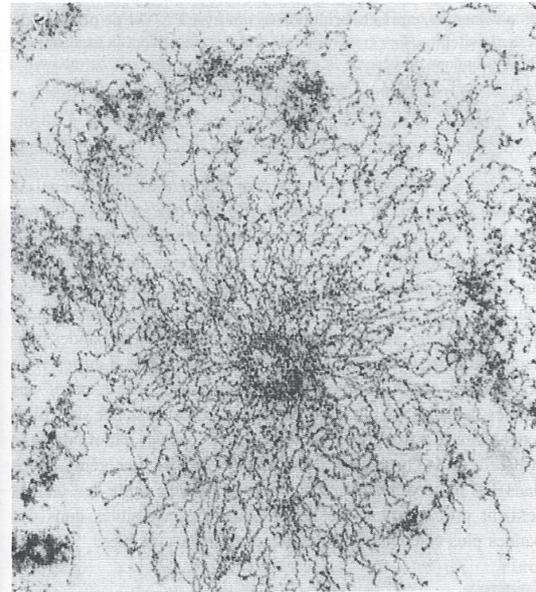


(b)

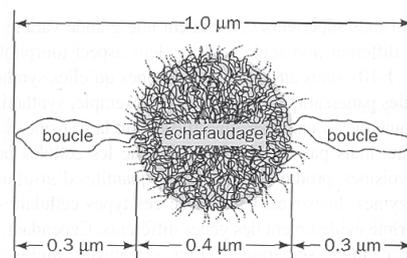


Micrographies électroniques d'un chromosome humain en métaphase, débarrassé de ses histones. (a) La matrice protéique centrale (échafaudage) sert à ancrer l'ADN environnant. (b) Un plus fort grossissement permet d'observer que l'ADN est attaché en boucles à l'échafaudage.

(a)



(b)



Organisation de l'ADN dans un chromosome métaphasique. (a) Micrographie électronique d'une section transversale d'un chromosome humain en métaphase. Notez la masse des fibres de chromatine rayonnant en étoile autour de l'échafaudage central. [Aimablement fourni par Ulrich Laemmli, Université de Genève, Suisse.] (b) Diagramme d'interprétation indiquant comment les boucles radiales de 0,3 μm de long sont supposées se combiner avec l'échafaudage de 0,4 μm de large pour former un chromosome en métaphase de 1,0 μm de diamètre.

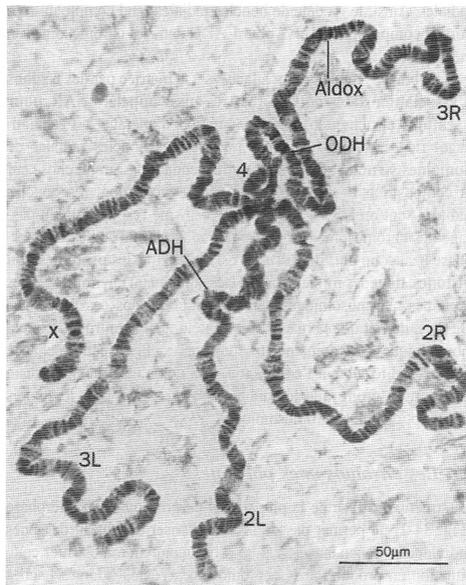
Figure 28 : Chromosomes polyténiques
(document à revoir pour le chapitre expression de l'IG)

Découverts par **Balbani en 1881** chez la larve de **Chironome**, ils sont présents dans les **glandes salivaires** des larves d'insectes (Chironome, Drosophile). Ces chromosomes sont formés d'un très grand nombre de molécules d'ADN qui sont restées solidaires après des réplifications multiples.

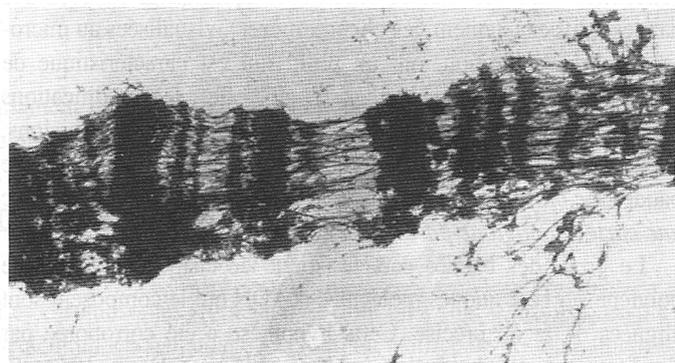
Un chromosome polytène (ou polyténique) peut contenir jusqu'à 1024 brins d'ADN (fonction non connue !)

Chez la Drosophile, on observe ainsi 4 chromosomes géants (2 mm de long) qui présentent une **alternance de bandes claires et sombres** :

- bandes claires : 15% de l'ADN, régions décondensées et exprimées
- bandes sombres : ADN condensé



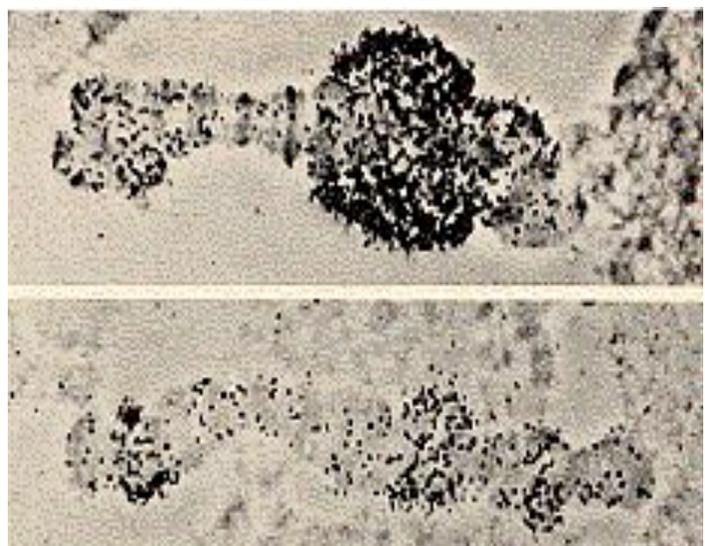
Micrographie photonique de chromosomes polyténiques colorés, de glande salivaire de *D. melanogaster*. Ces chromosomes présentent des bandes foncées séparées par des régions plus claires. Les quatre chromosomes d'une cellule sont reliés entre eux par leurs centromères. Les positions chromosomiques des gènes codant l'alcool déshydrogénase (ADH), l'aldéhyde oxydase (Aldox) et l'octanol déshydrogénase (ODH) sont indiquées.



Micrographie électronique d'un segment de chromosome polytène de *D. melanogaster*. Notez que les régions entre les bandes correspondent à des fibres de chromatine plus ou moins parallèles à l'axe du chromosome, alors que les bandes, qui contiennent environ 95 % de l'ADN chromosomique, sont beaucoup plus condensées. [Avec la permission de Gary Burkholder, Université du Saskatchewan, Canada.]

On observe des **puffs** ou encore **anneaux de Balbani** : ce sont des zones décondensées. On peut montrer que l'incorporation d'uracile tritiée se fait à leur niveau. Ceci montre que la transcription a lieu au niveau d'ADN décondensé.

Vue du chromosome IV de larve de Chironome après un traitement à l'ecdysone en présence d'uridine tritiée.



Même expérience mais en présence d'actinomycine D, un inhibiteur de la transcription (elle bloque l'accès aux ARNpolymérases sur l'ADN).

D'après users.rcn.com/~jkimball.ma.ultranet/

Figure 29 : Chromosomes en écouillons
(document à revoir pour le chapitre expression de l'IG)

Ils sont observables dans les ovocytes d'Amphibiens au stade diplotène de la méiose. Les chromosomes homologues sont associés 2 par 2 et chacun présente des boucles disposées en vis à vis sur ses 2 chromatides (→ schéma), d'où leur nom. L'incorporation d'uracile tritiée se fait au niveau de ces boucles, montrant comme dans le cas des chromosomes polytènes, un lien étroit entre décondensation et transcription.

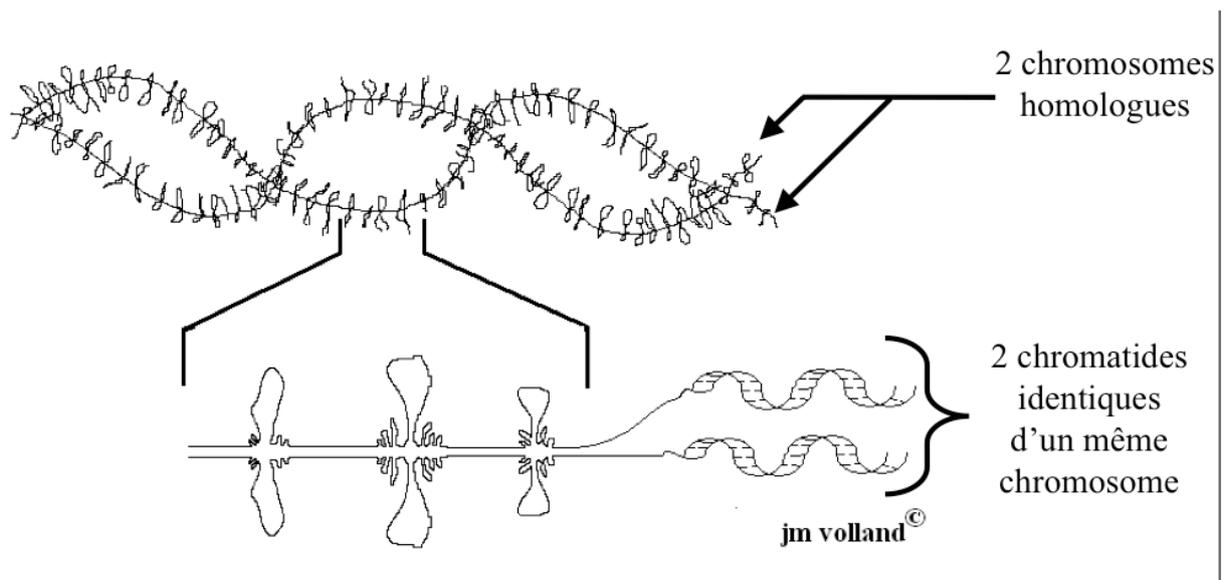
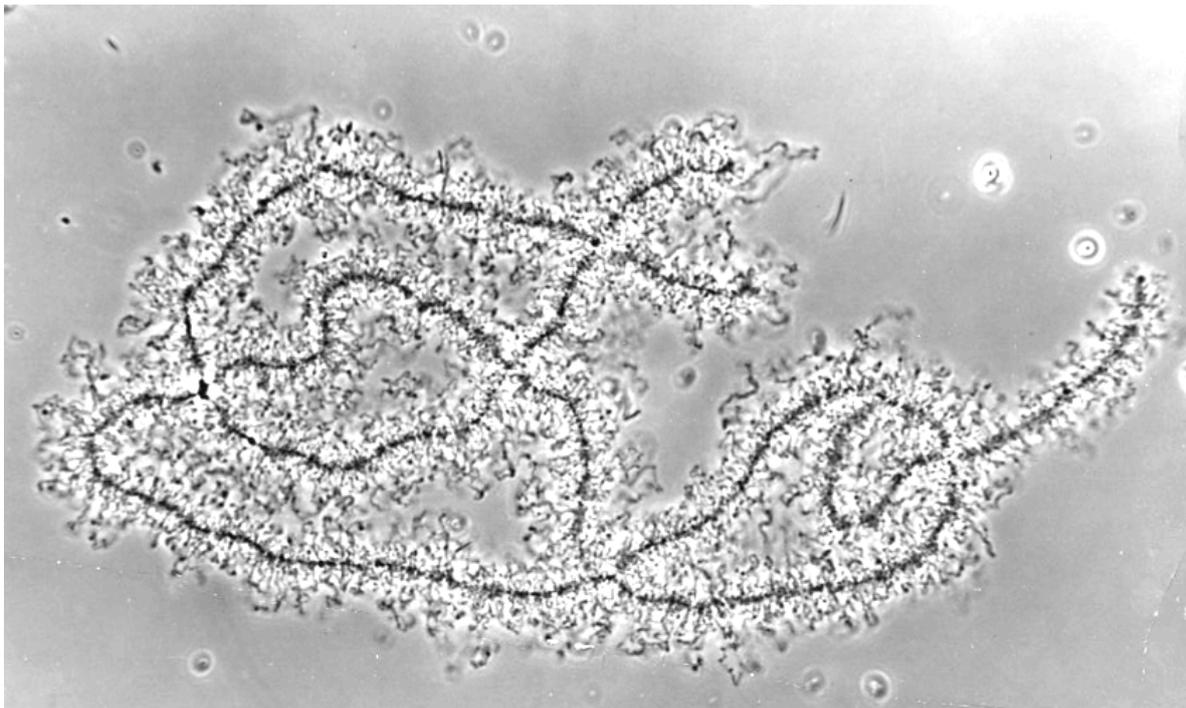


Figure 30 : Notion de gène

Cette notion a beaucoup évolué au cours du 20^{ème} siècle, en fonction des découvertes réalisées. Il est donc nécessaire de retracer les grandes étapes de la génétique pour comprendre ce qu'est un gène.

Les débuts de la génétique : la génétique formelle avec les travaux de Mendel (1860)

Mendel mena ses études sur le Pois (*Pisum*), une plante qui possède des caractéristiques intéressantes :

- elle possède des caractères constants et facilement différenciables
- il est possible de contrôler les croisements
- les hybrides sont aussi fertiles que les parents

Ses travaux montrent que les caractères sont déterminés par des facteurs héréditaires qui sont transmis de manière indépendante. Ces facteurs ont été appelés **gènes** pour la première fois par Johannsen.

Pour la petite histoire, Mendel avait commencé ses travaux sur la Souris, mais un évêque en visite déclara qu'un monastère augustin n'était pas un lieu approprié à la reproduction des souris...d'où le Pois!

La théorie chromosomique de l'hérédité : travaux de Weisman puis Boveri et Sutton (1903) et enfin Morgan (photo)

Leurs travaux ont montré le lien entre la transmission des gènes et celle des chromosomes à partir d'études sur la Drosophile et le Maïs. Ils confirment la localisation chromosomique des gènes. Ils mettent également en évidence que certains gènes sont sur le même chromosome : ils sont **liés**, alors que d'autres sont portés par des chromosomes différents : ils sont **indépendants**.



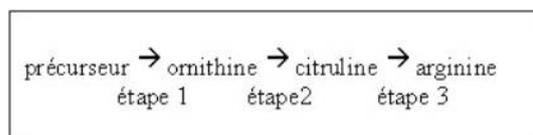
Un gène = une enzyme : Beadle et Tatum (1941) avec les travaux sur *Neurospora crassa*

Leur hypothèse de travail était que si les gènes codent pour des enzymes alors il devait être possible d'obtenir des souches déficientes dans des gènes qui contrôlent des réactions enzymatiques spécifiques. Ils ont obtenu de telles souches. Celles-ci poussent sur du milieu qu'ils ont appelé complet (c.a.d. qui contient les métabolites finaux qui sont nécessaires pour que le champignon croisse, comme les acides aminés, vitamines, bases etc.) mais ne poussent pas sur du milieu dit minimum (qui ne contient pas les métabolites). La supplémentation du milieu minimum par une substance précise qui restaure la croissance permet de savoir quel enzyme est déficiente. De telles souches sont dites **auxotrophes** (versus **prototrophe** pour la souche qui pousse sur milieu minimum).

Au cours de leurs études, ils ont par exemple obtenu trois souches qui ne pouvaient pas pousser sur du milieu dépourvu d'arginine :

souches	phénotype	gène muté
A	pousse sur milieu minimum +arginine	1
B	pousse sur milieu minimum +arginine ou +citruline	2
C	pousse sur milieu minimum+ arginine ou + citruline ou + ornithine	3

Or dans la cellule, des analyses biochimiques ont permis d'élucider la chaîne de biosynthèse de l'arginine qui est:



Ici pour expliquer les résultats, il suffit de supposer que :

- le gène 1 contrôle l'étape 3
- le gène 2 contrôle l'étape 2
- le gène 3 contrôle l'étape 1

Leurs travaux ont permis de mettre en évidence la **relation entre une mutation et l'absence d'une enzyme**.

Le lien entre génotype et phénotype s'expliquait enfin!!!! Le codage n'était cependant pas expliqué.

Le support moléculaire de l'IG est l'ADN :

Mise en évidence du support : Travaux de Griffith (notion de principe ou facteur transformant), Mc Carty puis Hershey et Chase *voir cours*

Mise en évidence de la structure en double hélice : travaux de Watson (photo, 1953) *voir cours*

Découverte de l'ARNm (1960) : travaux de F. Jacob et J. Monod (Prix Nobel)

Ces découvertes ont permis de définir un gène comme une portion d'ADN qui code pour un ARN généralement traduit en protéine

**Découverte des gènes mosaïques : (1977)**

On découvre qu'un gène peut permettre la formation de différents ARN (épissage différentiel) qui vont former des protéines apparentées mais différentes (= isoformes) : la définition "1 gène donne un ARN" n'est donc plus satisfaisante.

Test de complémentation : notion de cistron

Travaux sur le phage T4 par Benzer. La notion de cistron désignait une portion d'ADN codant une protéine ou une sous unité de protéine. La notion de cistron n'est plus utilisée : on sait que les cistrons correspondent aux gènes.

Séquençage des génomes et notion d'ORF :

La recherche de gènes dans un génome séquencé se fait en recherchant des **ORF = cadre ouvert de lecture (Open Reading Frame)** c'est-à-dire des séquences qui débutent par un codon d'initiation et se finissent par un codon stop. Ce sont des séquences qui possèdent potentiellement les éléments nécessaires à la synthèse d'une chaîne polypeptidique.

Ainsi, le nombre de gènes dans un génome donné par cette méthode correspond au nombre d'ORF.

Bilan : définition d'un gène en 2017

Un gène comprend l'ensemble des séquences nécessaires à la synthèse d'une chaîne polypeptidique ou d'un ARN fonctionnel. Il comprend des séquences codantes, des régions de contrôle et parfois des introns.

Figure 31 : structure de l'opéron lactose

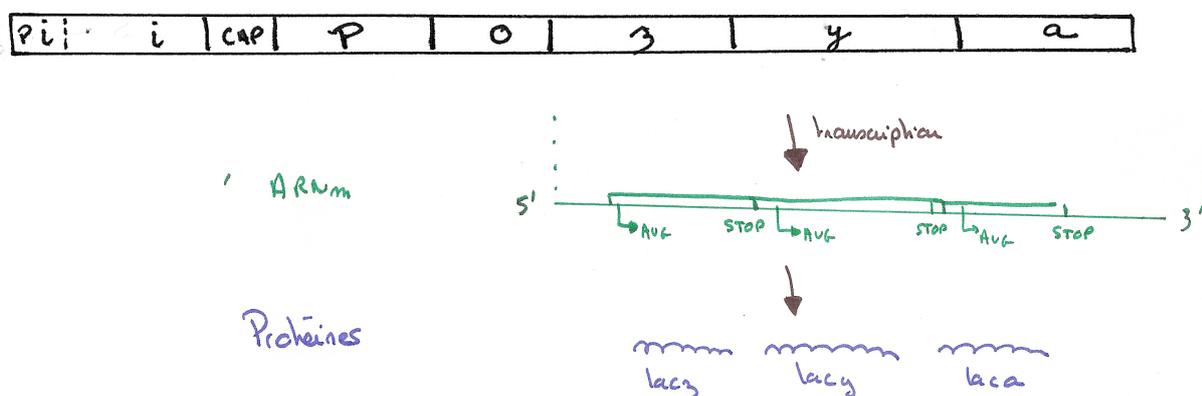


Figure 32 : étude de la complexité des génomes par des expériences de réassociation
(www.edu.upmc)

Méthode : l'ADN est découpé en fragments de 300 à 10 000 pb puis dénaturé. On le laisse ensuite se réhybrider.

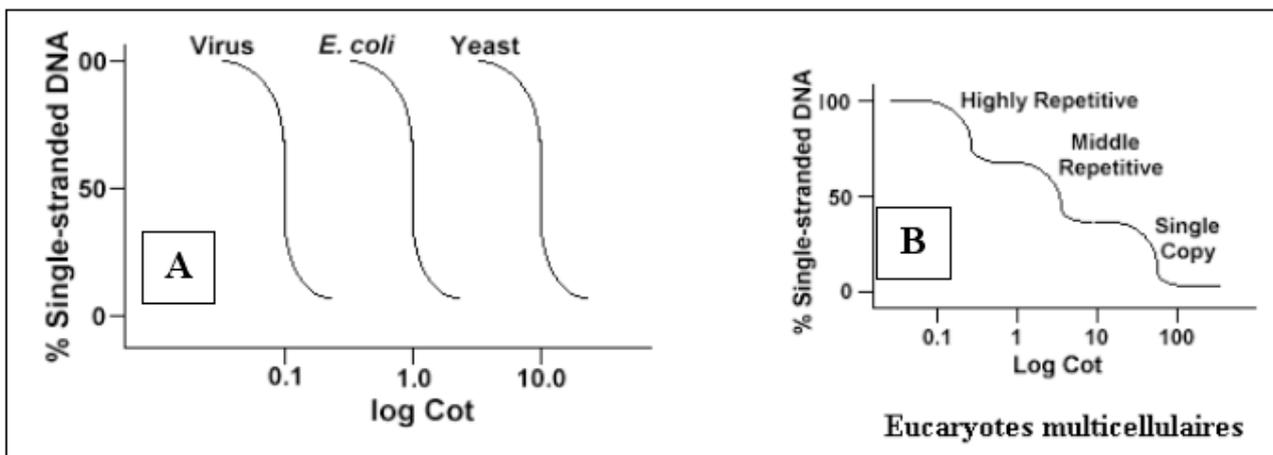
L'étape déterminant la vitesse de renaturation (=hybridation) est la rencontre entre séquences complémentaires : cette étape est directement dépendante du nombre de séquences identiques dans le génome, c'est-à-dire le nombre de séquences répétées. Une fois que les séquences complémentaires se sont rencontrées par diffusion au hasard, elles s'hybrident rapidement. et forment des molécules double brin.

A retenir : **plus une séquence est répétée, plus la réhybridation est rapide.**

Présentation des résultats : courbe Cot

Sur l'axe des Y on porte le pourcentage d'ADN qui reste simple brin par rapport à la concentration totale d'ADN. L'axe des X porte une échelle logarithmique du produit de la concentration initiale de l'ADN (en moles/litre) par le temps écoulé (en secondes). Ainsi une solution avec une forte concentration en ADN incubée pdt un temps court aura le même Cot qu'une solution ayant une faible concentration en ADN incubée pendant un temps long.

Ces courbes sont appelées courbes de Cot. Une courbe lisse est l'indice d'une renaturation progressive et régulière de l'ADN. C'est ce qui est observé pour des virus, des bactéries et pour des cellules de levure. Par contre lorsque l'on prend de l'ADN d'organismes multicellulaires (*Homo sapiens*), l'aspect est totalement différent. La courbe observée semble le résultat de la mise à bout de trois courbes. Ce qui indique que dans le premier cas il y avait un seul type d'ADN alors que dans le second trois types d'ADN différents doivent coexister : le premier type se renature très rapidement, le second se renature nettement plus lentement, quant au troisième il ne se renature que très lentement

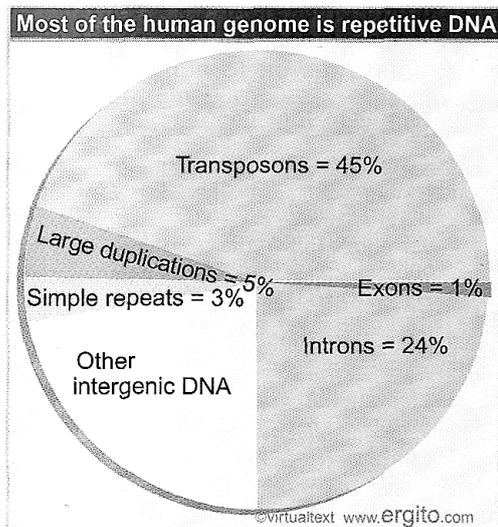


ADN hautement répété : ADN non codant des centromères et télomères

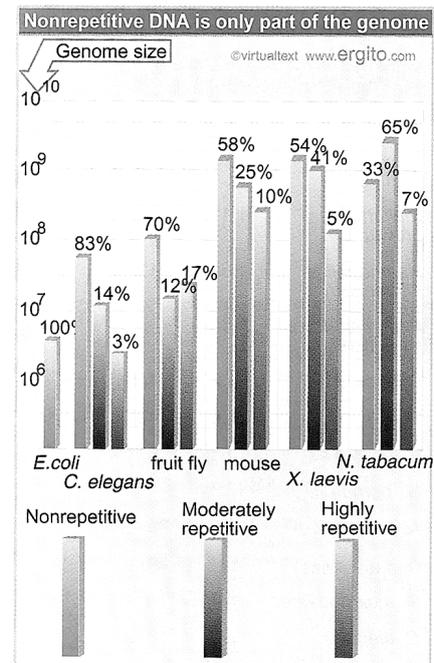
ADN moyennement répété : rétrotransposons, gènes codant pour les ARNr et ARNt

ADN non répété : essentiellement ADN codant

Figure 33 : Importance des séquences répétées dans les génomes
(in Lewin, International Edition 2004)



The largest component of the human genome consists of transposons. Other repetitive sequences include large duplications and simple repeats.



The proportions of different sequence components vary in eukaryotic genomes. The absolute content of nonrepetitive DNA increases with genome size, but reaches a plateau at $\sim 2 \times 10^9$ bp.

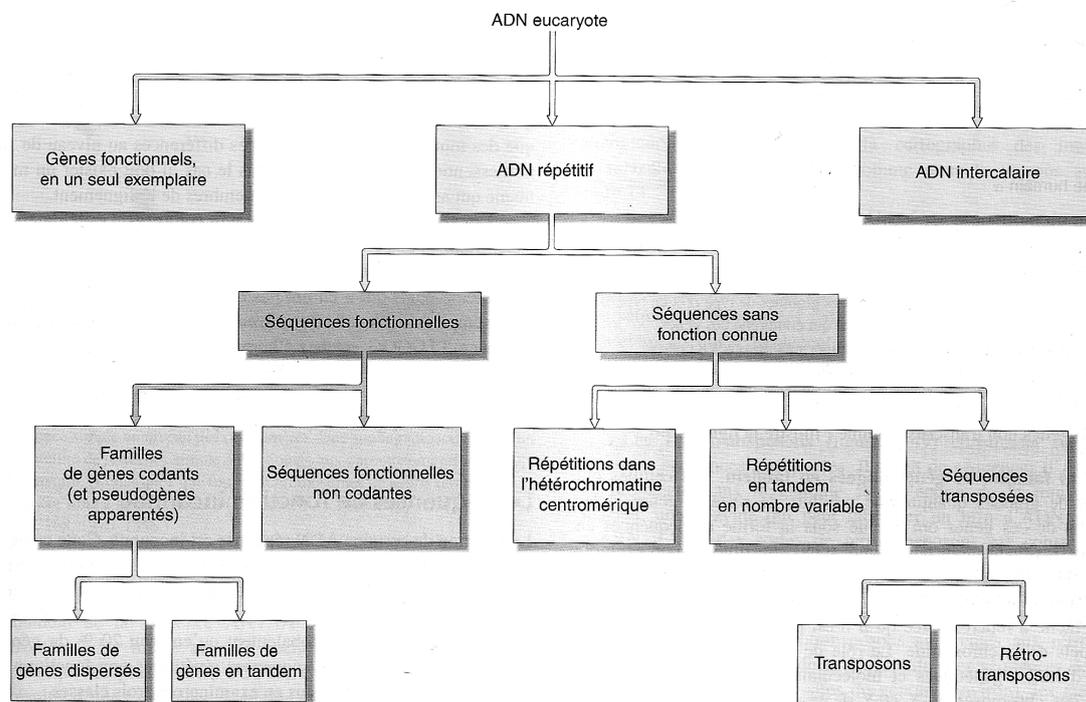


Figure 34 : Les séquences répétées des chromosomes eucaryotes.

Parmi les séquences répétées, celles dont les rôles sont les mieux décrits sont les suivantes. Chaque chromosome eucaryote possède deux types de séquences ayant un rôle principal sur la structure du chromosome, le centromères et les télomères. D'autres séquences répétées comme les transposons sont aussi très abondantes.

• Les séquences télomériques.

Il s'agit d'une séquence de 6 nucléotides répétées plus de 1 500 fois à chacune des extrémités d'un chromosome. Dans les cellules, ces télomères sont associés à des complexes protéiques qui neutralisent cette extrémité et la protègent. En effet, l'extrémité des chromosomes pourrait d'une part constituer un signal interprété comme une cassure au sein de l'ADN nécessitant une réparation et d'autre part cette extrémité serait plus sensible à l'action d'exonucléases qui hydrolyseraient les chromosomes.

Les séquences télomériques présentent un mode de réplication particulier où une enzyme spécifique, la télomérase, qui porte sa propre matrice d'ARN, synthétise de façon itérative la même séquence. La télomérase permet ainsi d'éviter un raccourcissement des chromosomes lors de la phase S, et cette enzyme n'est généralement pas active dans les cellules somatiques, mais active dans les cellules souches ou germinales. La réduction de la longueur des télomères lors des divisions successives serait une des raisons du vieillissement cellulaire par perte de matériel génétique aux extrémités des chromosomes.

• Les séquences centromériques.

Une séquence d'une longueur de 170 paires de bases chez l'homme est répétée plus de 10 000 fois pour constituer un centromère, où l'ADN est essentiellement sous forme d'hétérochromatine. Le rôle de ces centromères est de servir de zone de reconnaissance pour un complexe protéique, le kinétochore, qui sert d'interface entre le matériel génétique et les microtubules : lors des divisions cellulaires, ces centromères sont donc indispensables pour la fixation des chromosomes à la machinerie de ségrégation.

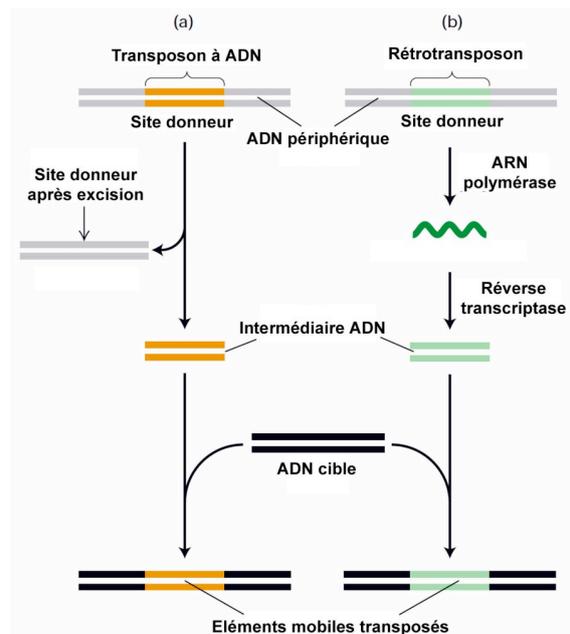
• L'ADN transposable.

Il s'agit de séquences d'ADN capables de se déplacer dans le génome. Ces séquences ont une longueur de quelques centaines à quelques milliers de paire de bases et sont considérées comme des symbiotes moléculaires dont la seule fonction connue serait de se maintenir dans le génome. Ils ont été découverts par Barbara Mac Clintock chez le maïs dans les années 1940, et sont classés en deux catégories :

- ceux qui s'excisent d'une région donnée et qui se déplacent d'un locus à l'autre sous forme d'une molécule d'ADN (transposons à ADN, mécanisme de couper/coller)
- ceux qui sont copiés ou transcrits en un ARN, lui-même rétro-transcrit en ADN puis inséré (rétro-transposons).

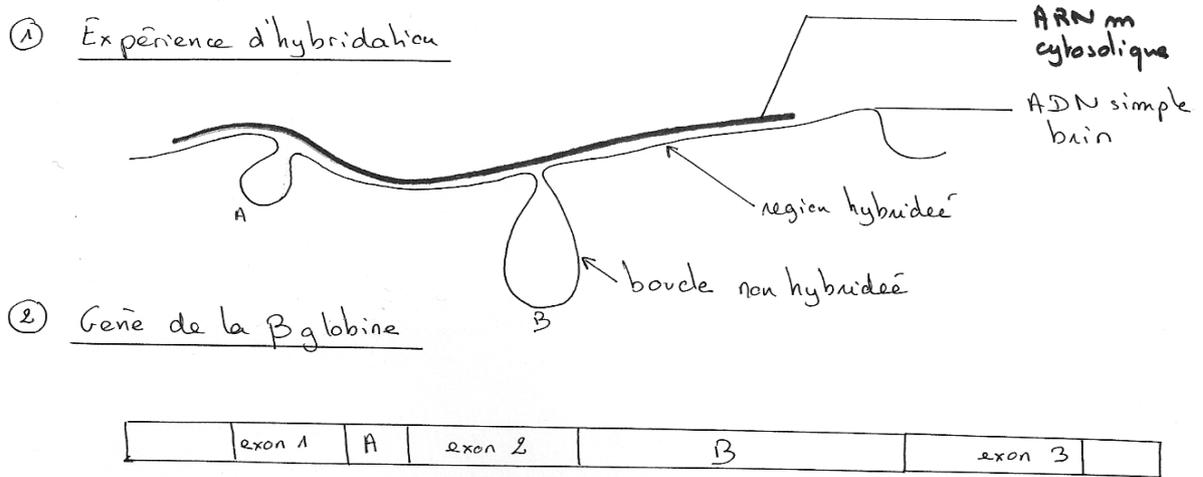
Ces derniers sont considérés comme apparentés aux rétrovirus qui utilisent aussi ce processus de rétro-transcription.

Tous ces éléments génétiques mobiles participent à l'évolution de plusieurs façons différentes. A la suite de leur déplacement, ils peuvent s'insérer dans des gènes et les inactiver. Du fait de leur grande abondance, ils peuvent aussi être l'origine de crossing-over homologues inégaux, ce qui conduit à des duplication de gènes, ou à l'ajout ou le retrait d'introns, d'exons. Enfin, leur insertion à proximité des *enhancers* provoque également des modifications du contrôle de la transcription.



Dans le cas des transposons à ADN, (a), le transposon à ADN s'excise, laissant le site d'insertion initial tel qu'à l'origine. En ce qui concerne les rétrotransposons, un ARN intermédiaire est produit, puis donne naissance à une nouvelle séquence d'ADN qui sera insérée ailleurs dans le génome. Le site initial conserve par ailleurs le transposon.

Figure 35 : les gènes eucaryotes sont morcelés



A et B: introns
 exon 1 : 240 pb
 exon 2 : 500 pb
 exon 3 : 250 pb

Le gené de la globine est morcelé

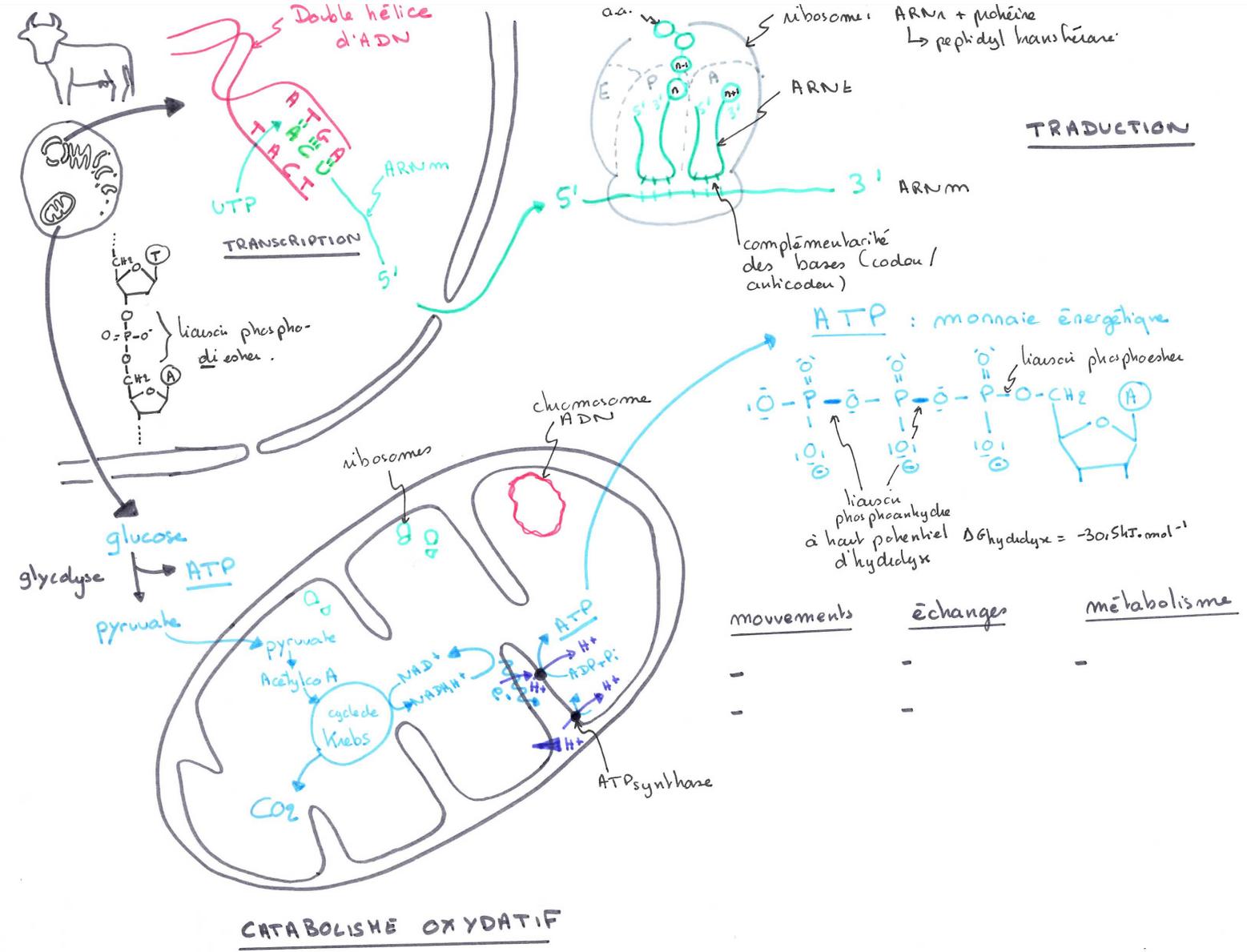


Figure 36 : SB sur les nucléotides