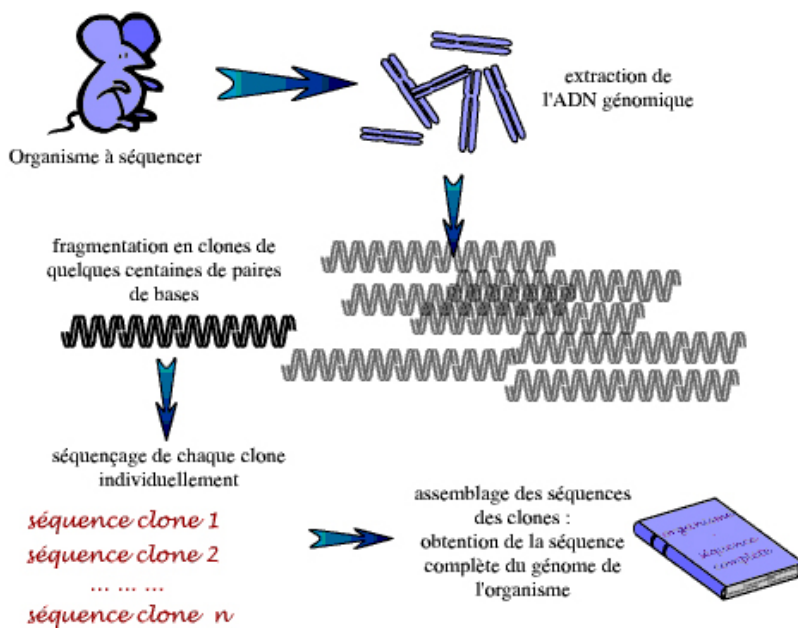


SV-F-1.1 poly 01

Génomes des cellules et des virus

génom	Taille approximative du génom haploïde	Nature chimique / nombre de chromosomes
Virus		
ΦX174	5 kb	ADN circulaire simple brin (forme répliquative double brin)
M13	6 kb	
VIH	9,1 kb	
"T pair" : T2, T4, T6	165 kb	ADN linéaire simple brin
★ Phage λ	48,5 kb	
★ VMT	6,5 kb	ADN linéaire double brin
★ SARS-Cov2	29,9 kb	
Procaryotes		
Methanococcus jannaschii	1600 kb	ADN circulaire double brin
★ Escherichia coli	4600 kb = 4,6 Mb	
Borrelia burgdorferi	910 kb	ADN linéaire double brin
Eucaryotes		
		Nombre haploïde de chromosomes (n)
Saccharomyces cerevisiae (levure)	13 Mb	16
Caenorhabditis elegans (nématode)	97 Mb	6
Arabidopsis thaliana (petite Angiosperme herbacée)	100 Mb	5
Drosophila melanogaster (drosophile)	180 Mb	4
★ Homo sapiens	3000 Mb = 3 Gb	23 (22 autosomes + 1 gonosome)
Zea mais (maïs)	4500 Mb = 4,5 Gb	10
Fugu rubripes (poisson)	400 Mb	22
Amphiuma means (amphibien)	90000 Mb = 90 Gb	14

Fig. 1. La taille de quelques génomes. ★ ☆=★★★



Un peu de calculs mathématiques ☺

Soit L la taille d'un génome ($L=3 \cdot 10^9$ pour l'Homme) et N le nombre total de nucléotides lus : par exemple $N_1=3 \cdot 10^{10}$ et $N_2=9 \cdot 10^9$. On obtient alors une profondeur de lecture, correspondant au « nombre de fois que le génome est lu » : $P=N/L$; ici : $P_1=10$ et $P_2=3$.

On peut alors déterminer la proportion d'ADN lu au moins une fois : $\text{ADN lu} = 1 - e^{-N/L}$. Ici, on obtient un taux de lecture de 99,99% dans le cas 1 et de seulement 95% dans le cas deux, ce qui montre qu'une profondeur de 3 est largement insuffisante au séquençage complet du génome.

Concrètement, on considère qu'une profondeur de lecture supérieure à 10 est nécessaire pour obtenir une séquence complète de bonne qualité (= complète et avec un minimum de « trous », restant de petites tailles).

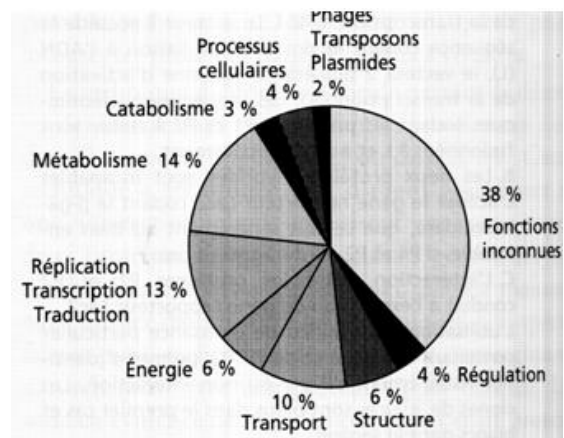
Fig. 2. Le séquençage d'un génome nécessite de séquencer un grand nombre de fragments d'ADN. ★ [2]

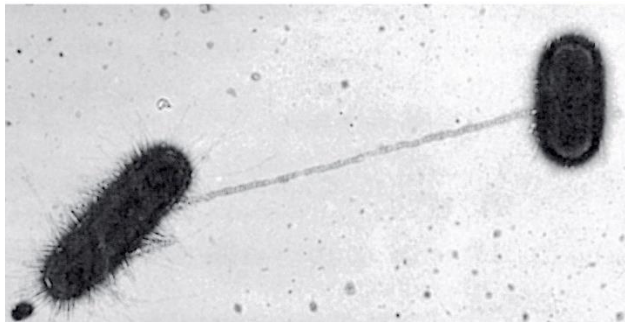
Taille du génome : 4 639 221 paires de bases (soit 4,6 Mb)

Nombre d'ORF : 4 288

Fraction codante : 88 %

Nombre d'ARNt : 86

Fig. 3. Le génome d'*Escherichia coli* : données issues du séquençage ★★



↑ Observation au MET d'un début de conjugaison bactérienne.

Pour information, quelques étapes de la conjugaison bactériennes. →

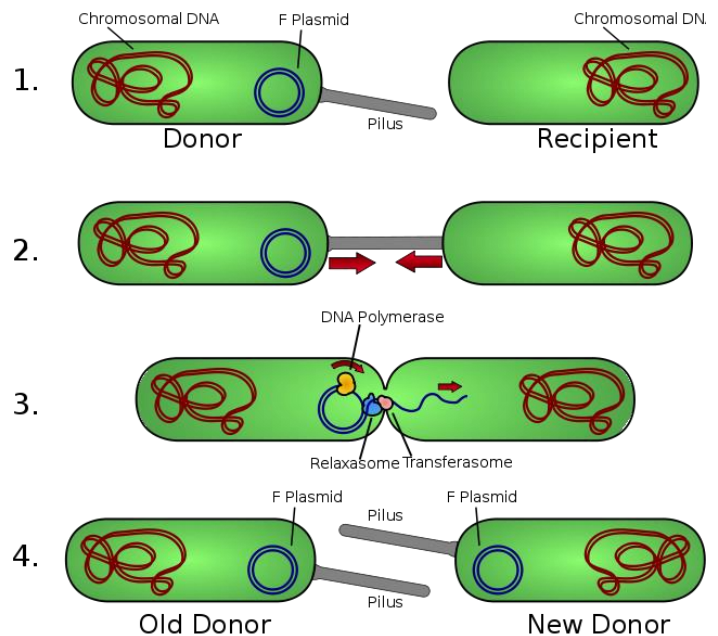


Fig. 4. La conjugaison bactérienne. ★

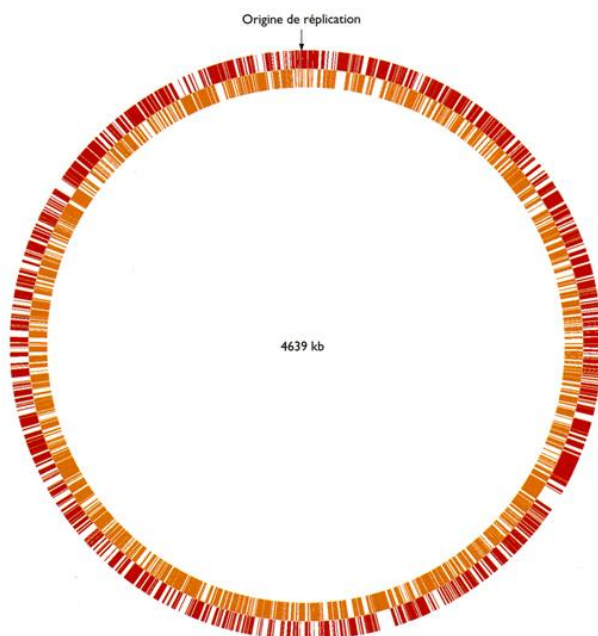


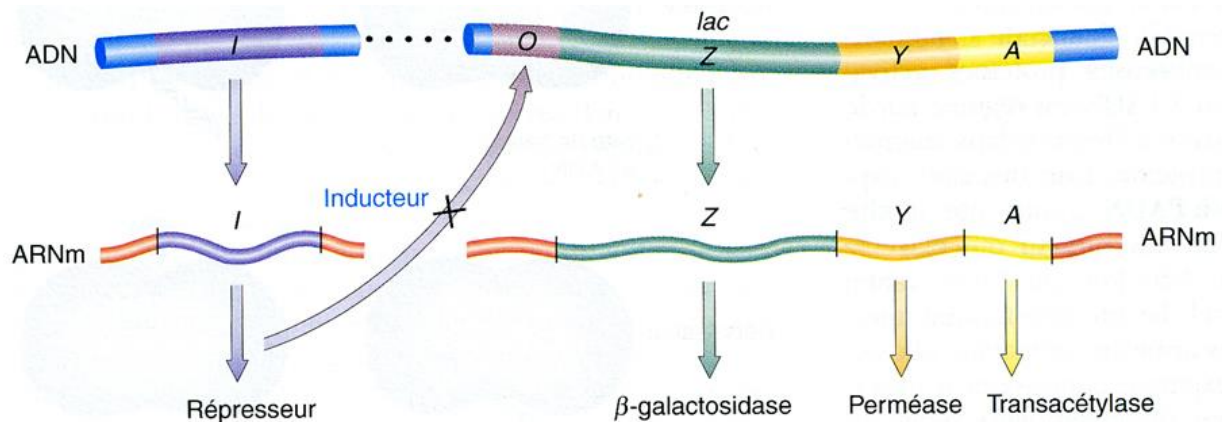
Fig. 5. Le génome d'*E. coli*. ★

Représentation du chromosome circulaire de la souche K12 d'*Escherichia coli*.

Les séquences transcrites sont indiquées, sur le cercle externe ou le cercle interne en fonction de leur sens de transcription (sens horaire pour le cercle externe et sens anti-horaire pour le cercle interne) : cette distinction correspond au brin portant la séquence codante.

La réplication débute chez les eubactéries en un point précis, l'origine de réplication, placée en haut par convention.

On peut noter que très peu de séquences ne sont pas transcrites.

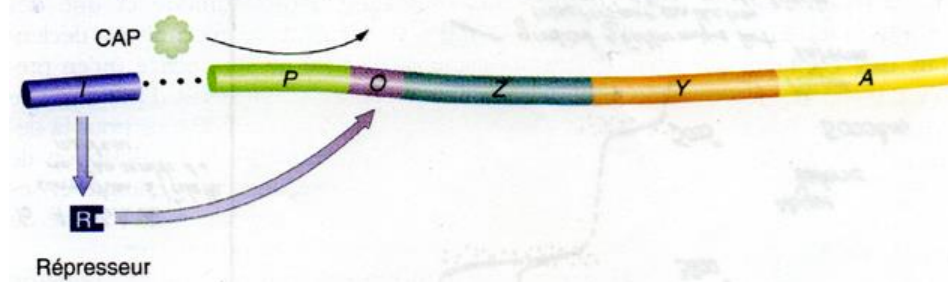


Un modèle simplifié de l'opéron *lac*. Les trois gènes, *Z*, *Y* et *A* sont exprimés de façon coordonnée. Le produit du gène *I*, le répresseur, bloque l'expression des gènes *Z*, *Y* et *A* en interagissant avec l'opérateur (*O*). L'inducteur peut inactiver le répresseur, empêchant ainsi son interaction avec l'opérateur. Dans ce cas, l'opéron est exprimé intégralement.

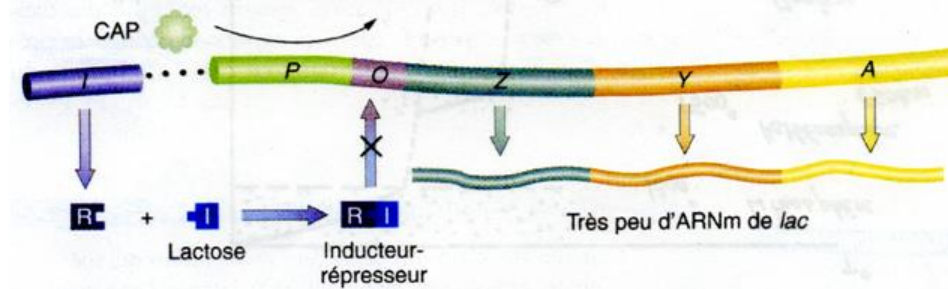
Fig. 6. Structure de l'opéron lactose. ★★★

Fig. 7. Fonctionnement de l'opéron lactose : une adaptation du niveau de transcription des gènes de structures aux besoins de la cellule, selon les conditions du milieu. ★★★

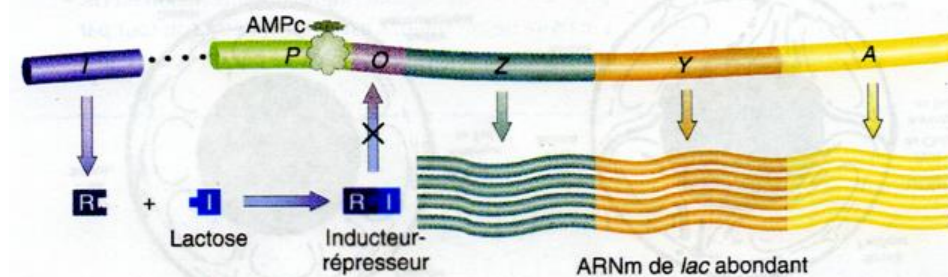
(a) Glucose présent (AMPc) ; pas de lactose ; pas d'ARNm de *lac*



(b) Glucose présent (AMPc faible) ; lactose présent



(c) Pas de glucose (AMPc élevé) ; lactose présent



[a]

En **absence de lactose**, le gène *lacI* (I) code une protéine à action de type répresseur : il s'agit d'un tétramère qui se fixe sur l'ADN au niveau de l'opérateur (O) et **empêche** la transcription par l'ARN polymérase.

[b+c]

En **présence de lactose**, sa fixation sur le répresseur inactive ce dernier. L'opérateur est donc « libéré » et l'ARN polymérase peut réaliser une transcription de l'ARNm polycistronique (gènes *lacZ*, *lacY* et *lacA*).

En présence de glucose, le **taux de transcription est faible** : la bactérie « préfère » utiliser le glucose.

[c]

En **présence de lactose ET absence de glucose**, l'absence de glucose conduit à la présence d'AMPc (« signal de faim ») : ce nucléotide peut s'associer avec une protéine CAP. Le complexe AMPc-CAP peut alors agir en favorisant le recrutement de l'ARN polymérase sur son promoteur : le **taux de transcription est alors élevé**, ce qui permet à la bactérie une utilisation importante du lactose présent dans le milieu.

Fig. 8. Le chromosomes eucaryote métaphasique. ★★★

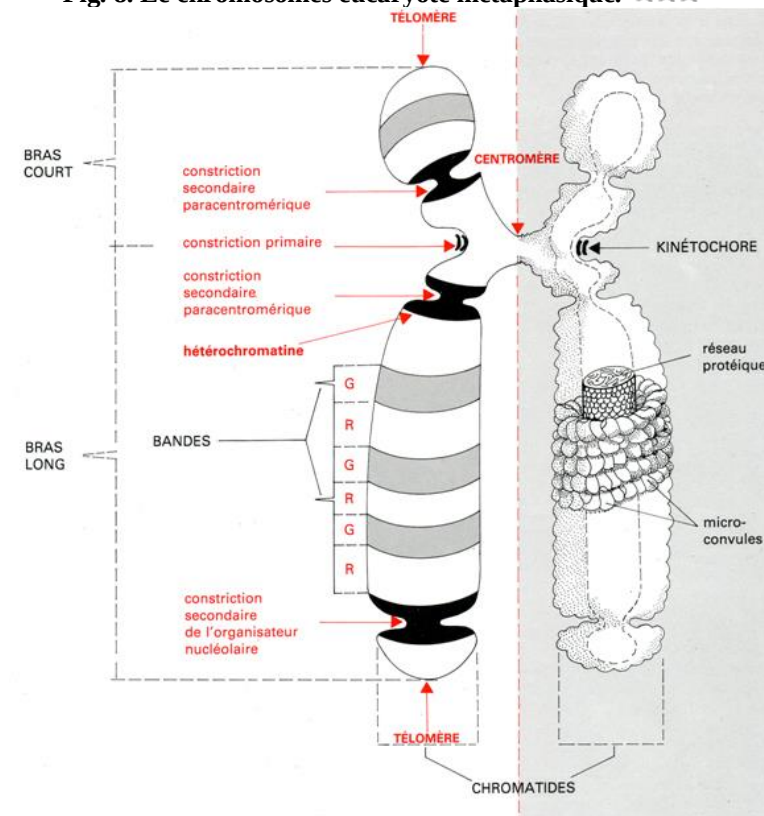


Fig. 9. Diversité morphologique des chromosomes. ★★

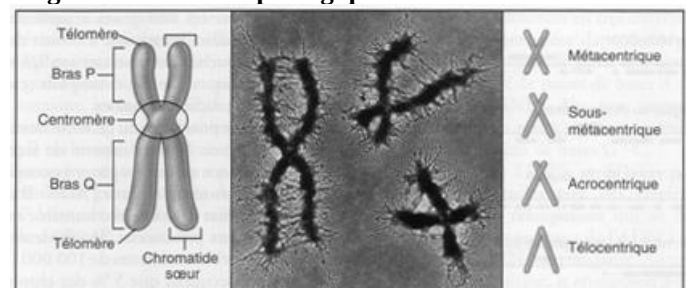


Fig. 10. Un caryotype humain : 46,XY ★★★

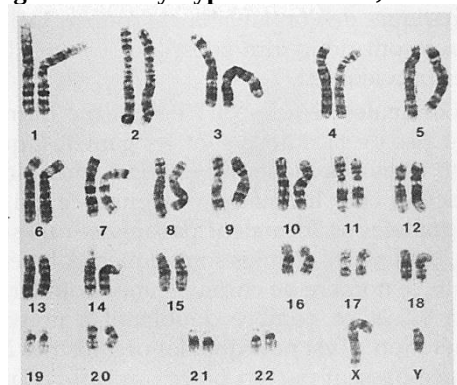


Fig. 13. Les différents niveaux de condensation de la chromatine eucaryote. ★★★ [4]

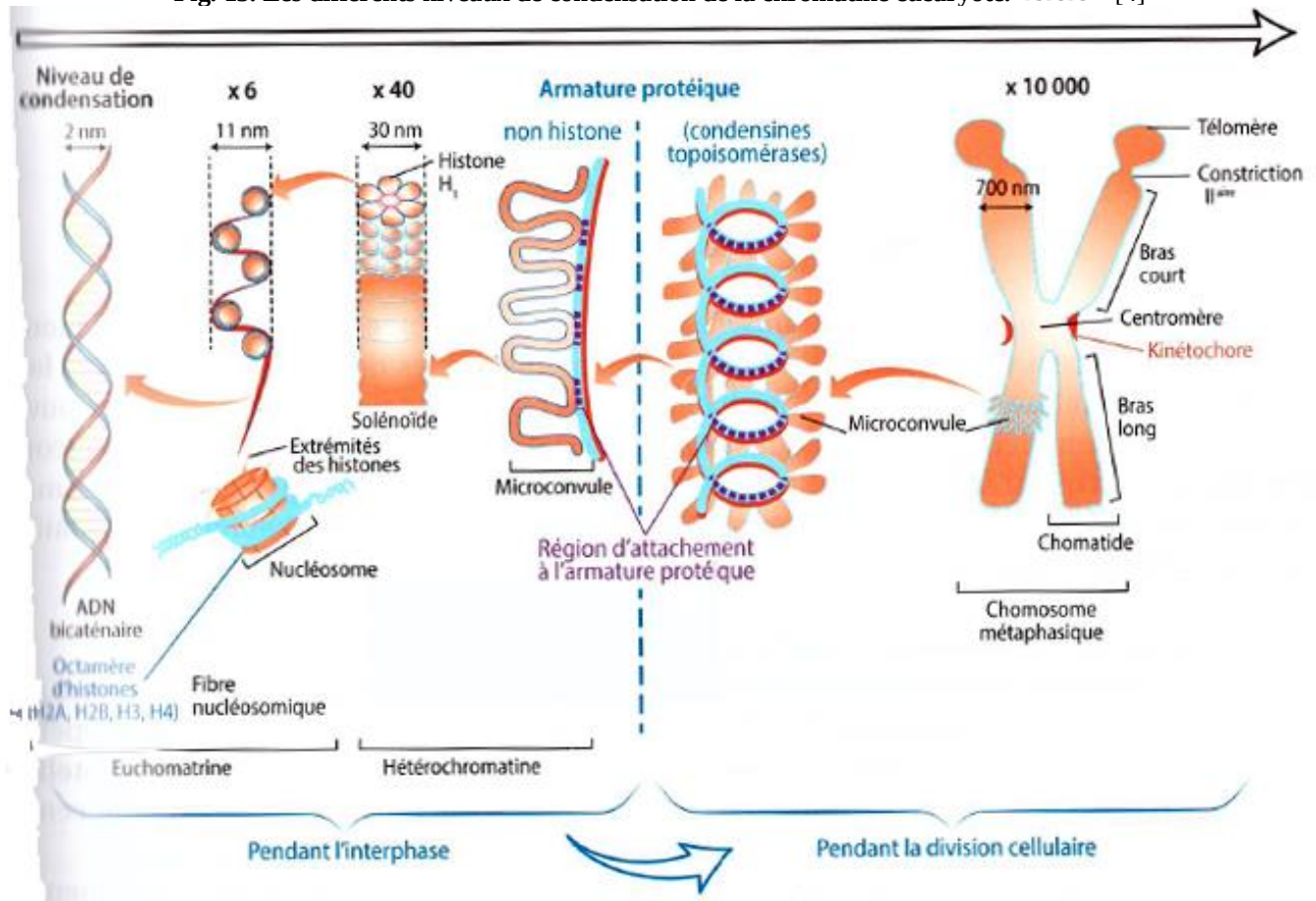


Fig. 14. Organisation du chromosome bactérien. ★★★ [3]

★★★ [3]

Le chromosome bactérien est condensé grâce à des protéines jouant un rôle analogue aux histones. L'ADN est ainsi surenroulé, c'est-à-dire qu'il forme des boucles repliées sur elles-mêmes (on dit que l'ADN fait des super-tours). Les domaines sont décondensés au niveau des portions transcrites.

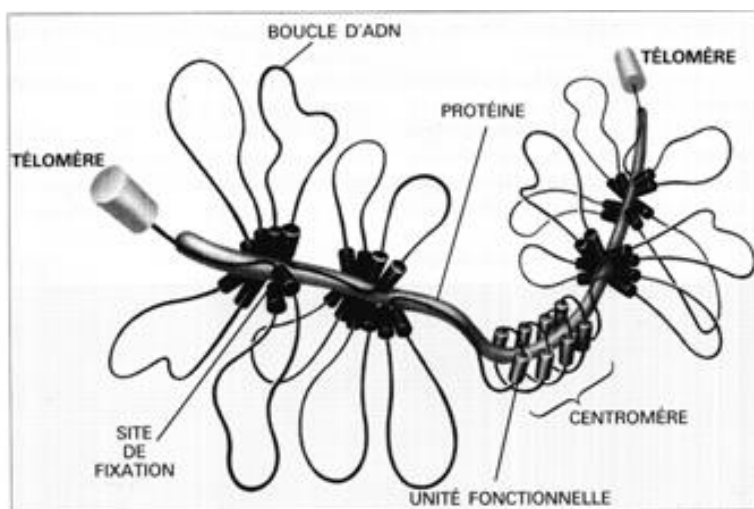
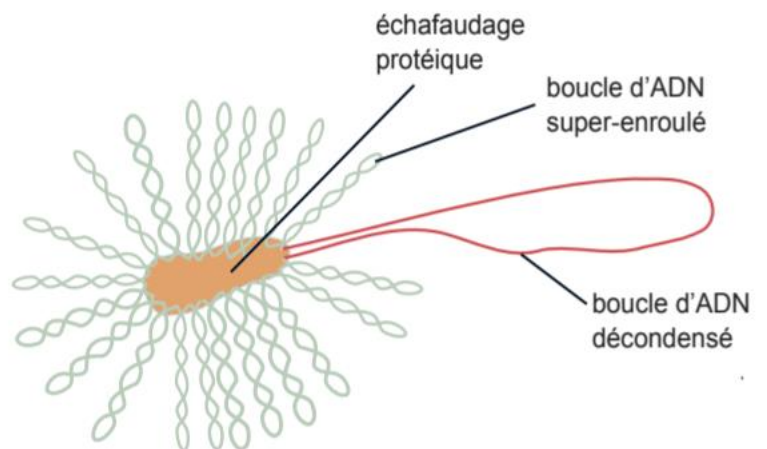


Fig. 15. Représentation schématique des segments essentiels d'un chromosome humain. ★★

Chaque boucle contient un ou plusieurs gènes. Les cylindres représentent des régions d'ADN qui ne codent pas pour les protéines, mais indispensables à la stabilité des chromosomes et à leur activité ; il s'agit en particulier des télomères et du centromère.

Fig. 16. Taille et fraction codante de quelques génomes eucaryotes. ★★

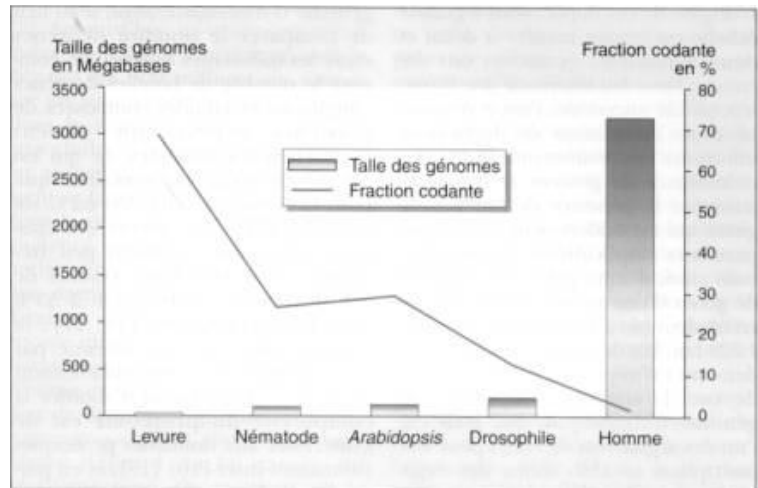


Fig. 17. Mise en évidence de séquences répétées dans le génome. ★★★ [3]

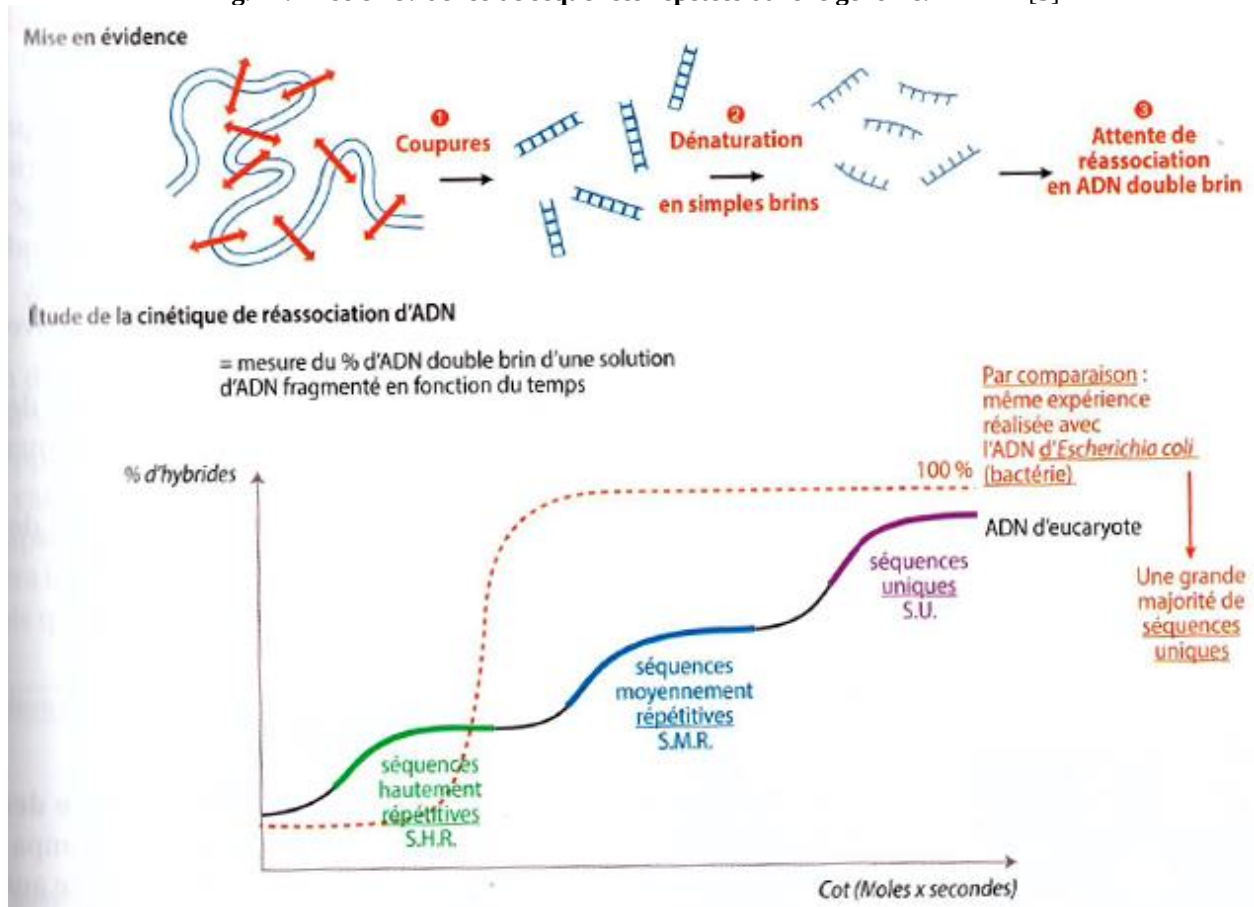


Fig. 18. Les différents types de séquences d'ADN au sein d'un génome eucaryote. ★★ [3]

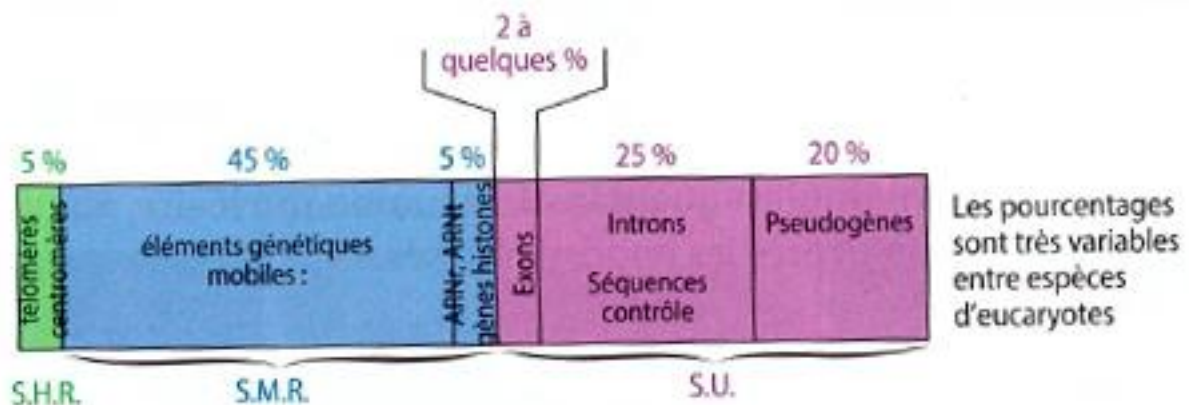
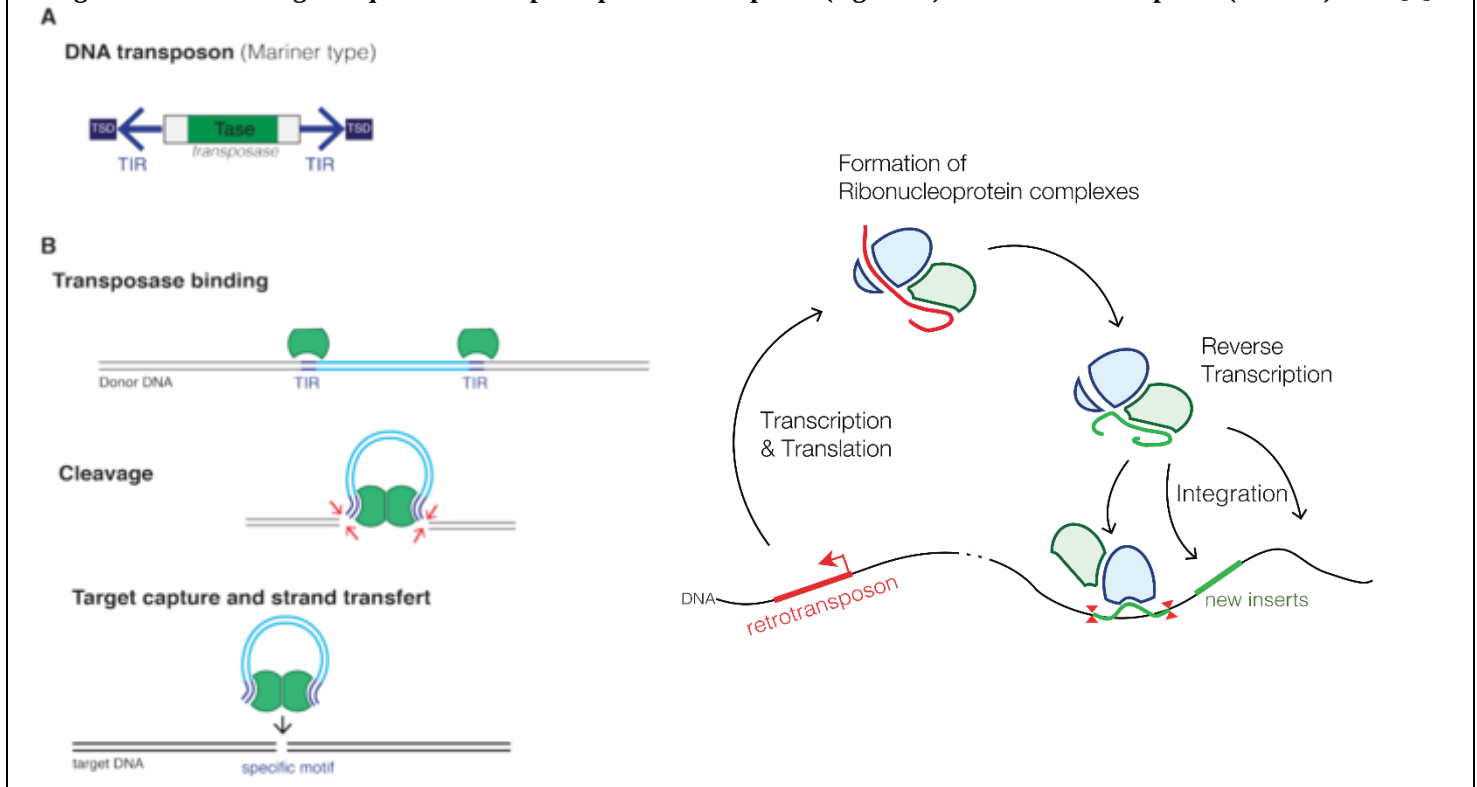
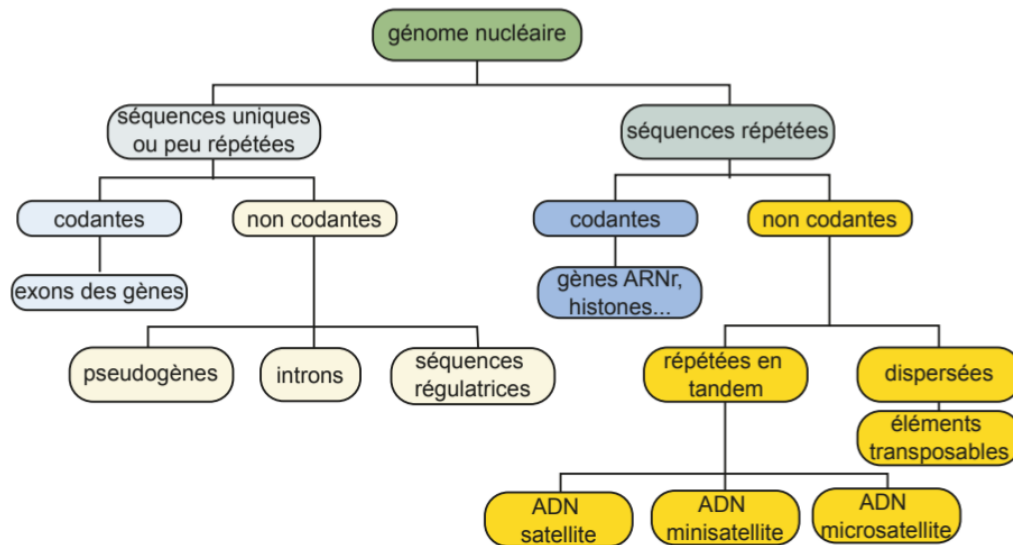
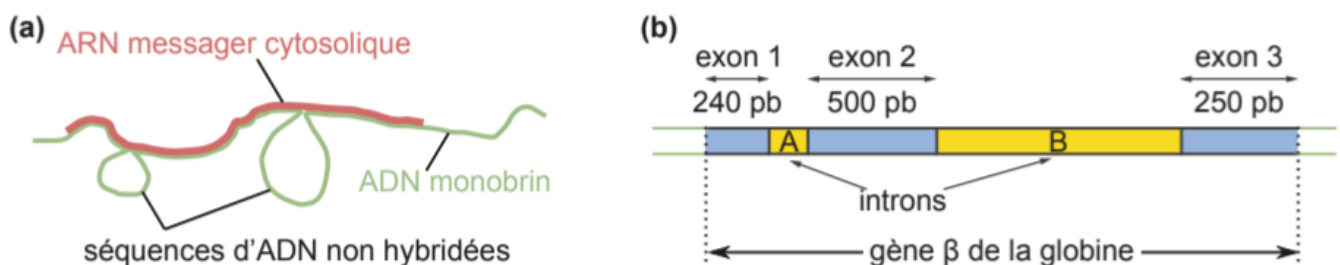


Fig. 19. Les éléments génétiques mobiles : principe d'un transposon (à gauche) et d'un rétrotransposon (à droite). ★ [5]**Fig. 20. Bilan : la diversité des séquences constituant le génome nucléaire eucaryote. ★★ [3]****Fig. 21. Un exemple de gène morcelé : le gène de la β -globine ★★ [3]**

Le gène de la β -globine comporte 3 exons et 2 introns (figure (b)). Cette structure est mise en évidence par l'hybridation de l'ARNm mature (cytosolique) avec l'ADN génomique (figure a).

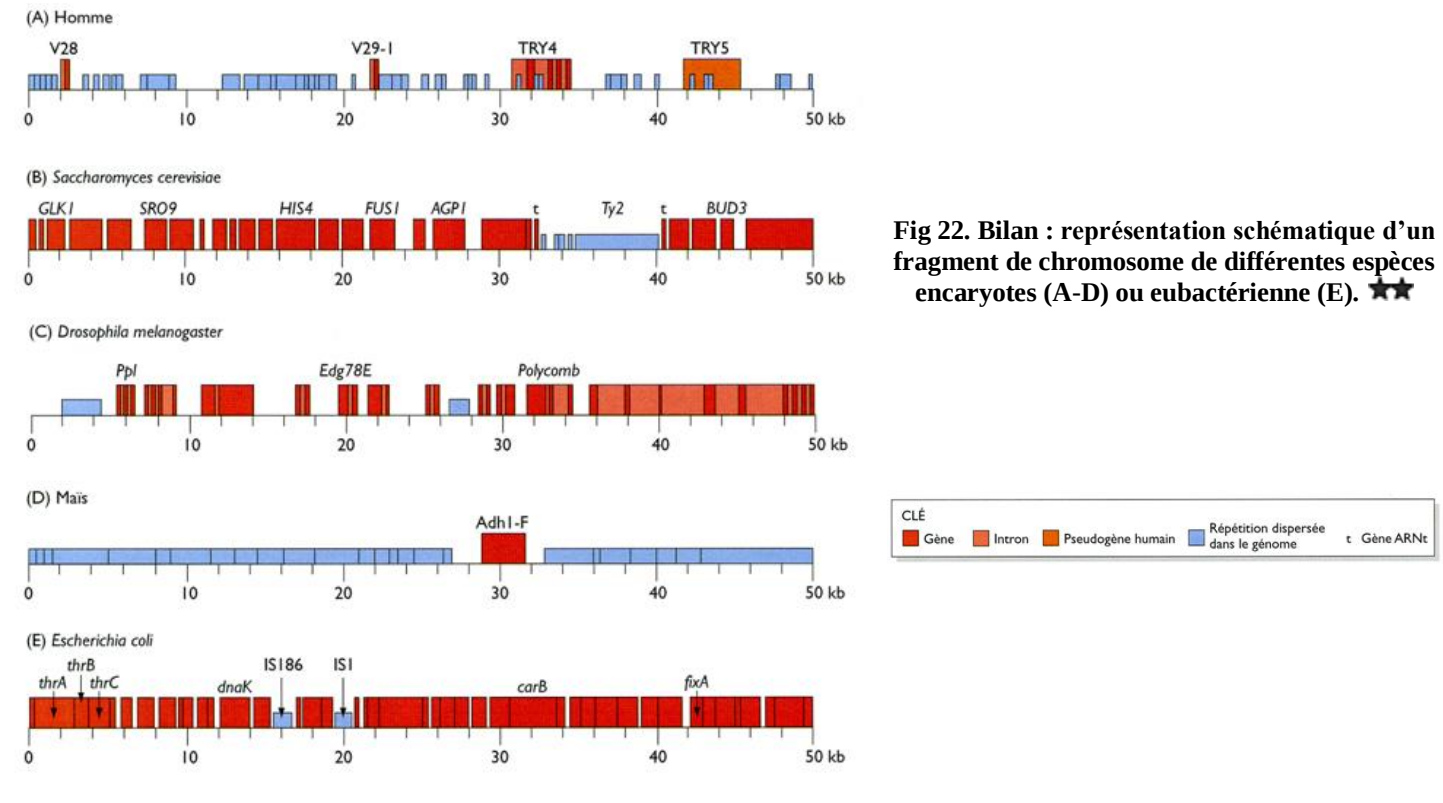
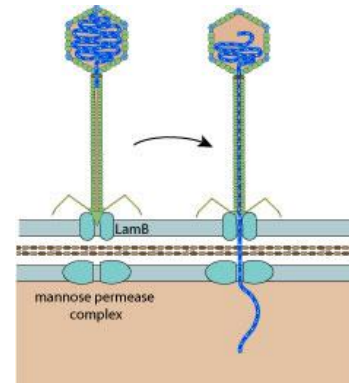
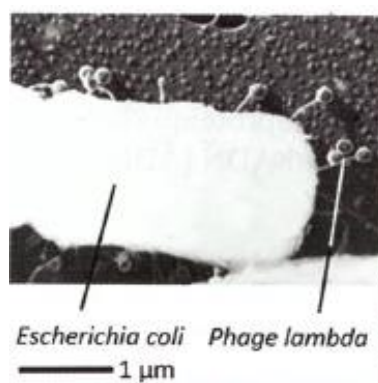


Fig 22. Bilan : représentation schématique d'un fragment de chromosome de différentes espèces encaryotes (A-D) ou eubactérienne (E). ★★

Fig. 23. Comparaison structurale des trois virus au programme. ★★★

	Phage λ	VMT	Coronavirus
Enveloppe ?	Non	Non	Oui
Capside	Constituée d'une tête (contenant l'IG) et d'une queue	Constituée d'une protéine répétée un grand nombre de fois (2300 !) Forme de bâtonnet	Nucléocapside formée par l'association entre la protéine N et l'ARN viral
Taille	150 à 200 nm Dont 50 nm pour la tête	Longueur : 300 nm Largeur : 17 nm	60 à 140 nm
Nature du génome	ADN double brin	ARN (+) simple brin	ARN (+) simple brin
Taille du génome	48,5 kb	6,4 kb	~ 30 kb
Nombre de gènes	~ 70 → nombreuses protéines	4 → 3 protéines (dont 1 avec 2 sous-unités)	7 à 15 → 29 à 33 protéines
Cellules infectées	Bactéries Gram (-) Ex : E. coli	Angiospermes (~150 espèces différentes, surtout des Fabacées et Solanacées)	Différents Mammifères, dont Homo sapiens
Fixation sur la cellule hôte	Fixation sur une porine de la paroi	Pénétration par blessure d'un organe	Fixation sur ACE2 (enzyme liée à la face externe de la membrane plasmique de plusieurs types cellulaires)

Fig. 24. Fixation du phage λ sur la bactérie et injection de l'ADN viral. ★★ [4][5]Fig. 25. Le génome du phage λ . ★

A l'intérieur de la carte sont indiquées la direction de la transcription et la localisation des promoteurs de gauche et de droite (PL et PR).

A noter : présence de gènes pour la synthèse de la tête et de la queue, pour le contrôle du cycle, de l'intégration du génome dans l'ADN hôte.

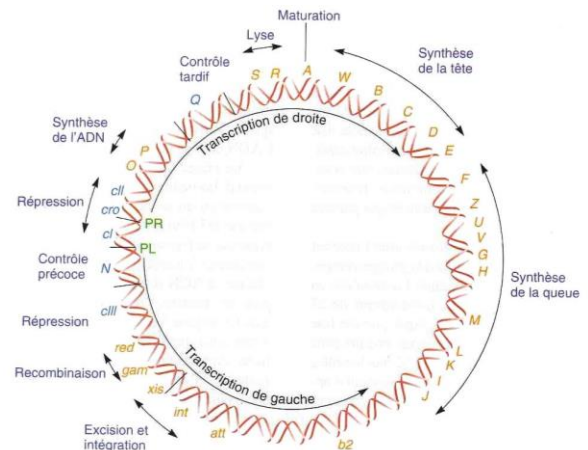


Fig. 26. Structure du virion et génome (simplifié) pour les trois virus au programme. ★★★ [4]

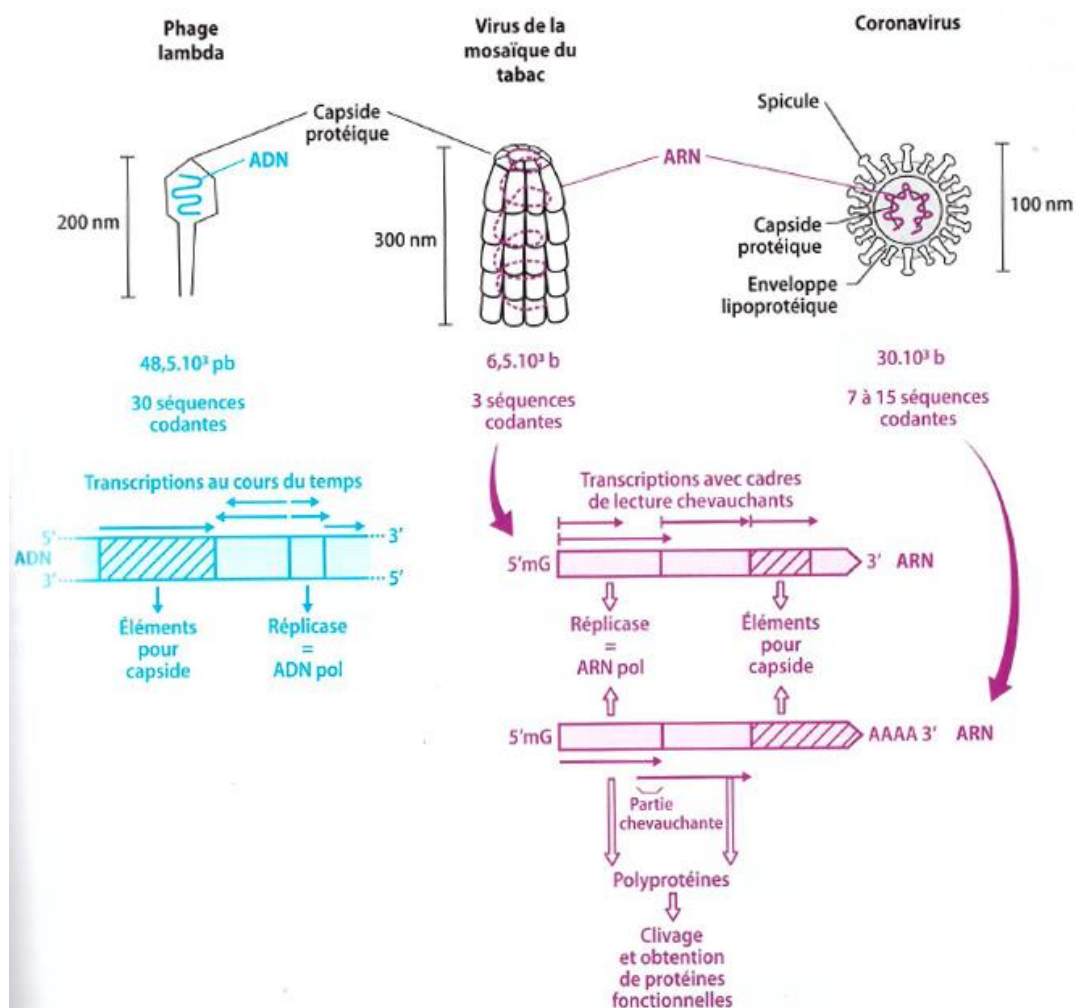


Fig. 27. Le Virus de la Mosaïque du Tabac (VMT) ★★★ [3][4]

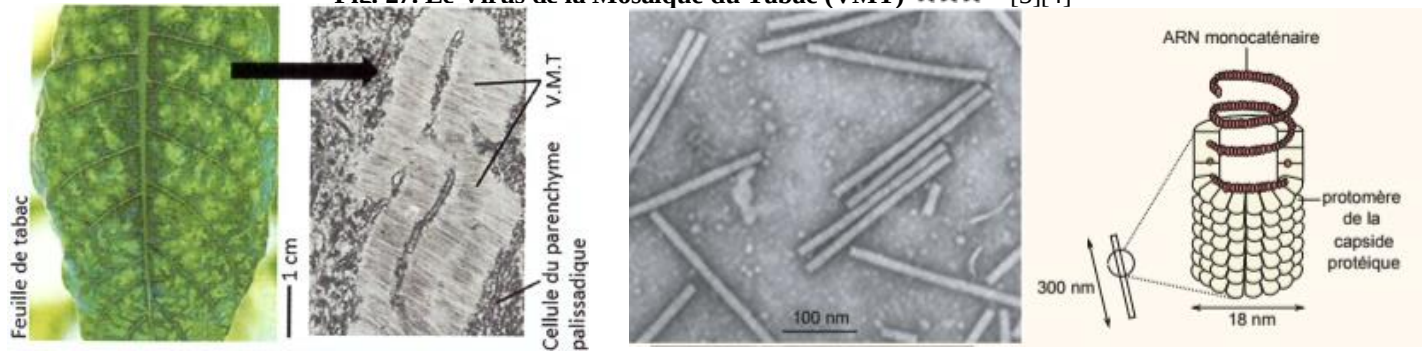


Fig. 28. Génome du VMT. ★★ [5]

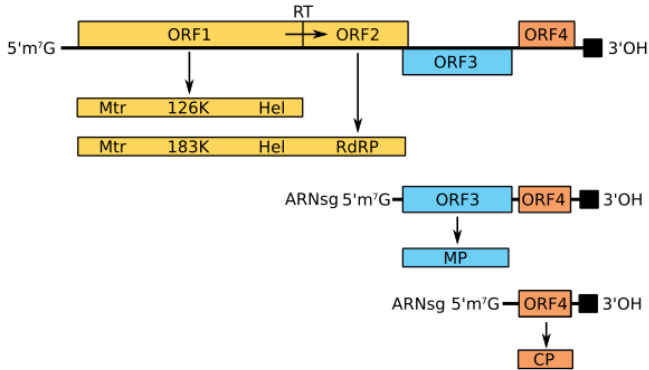
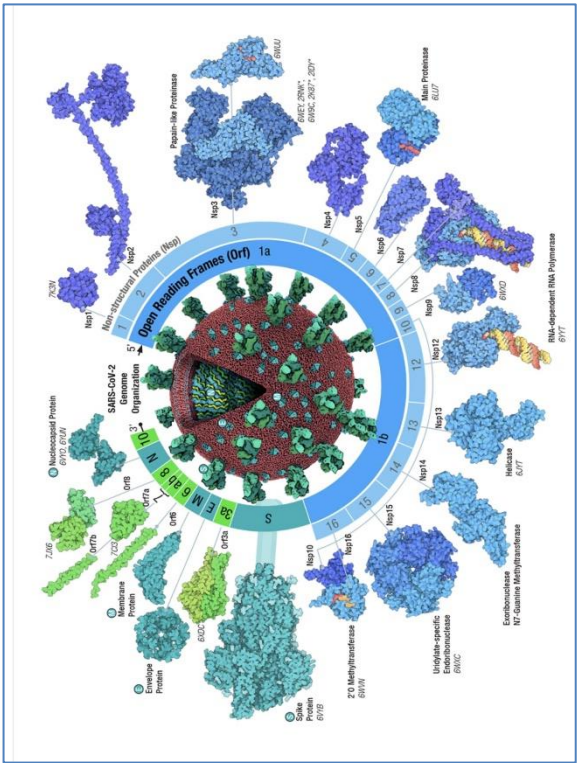
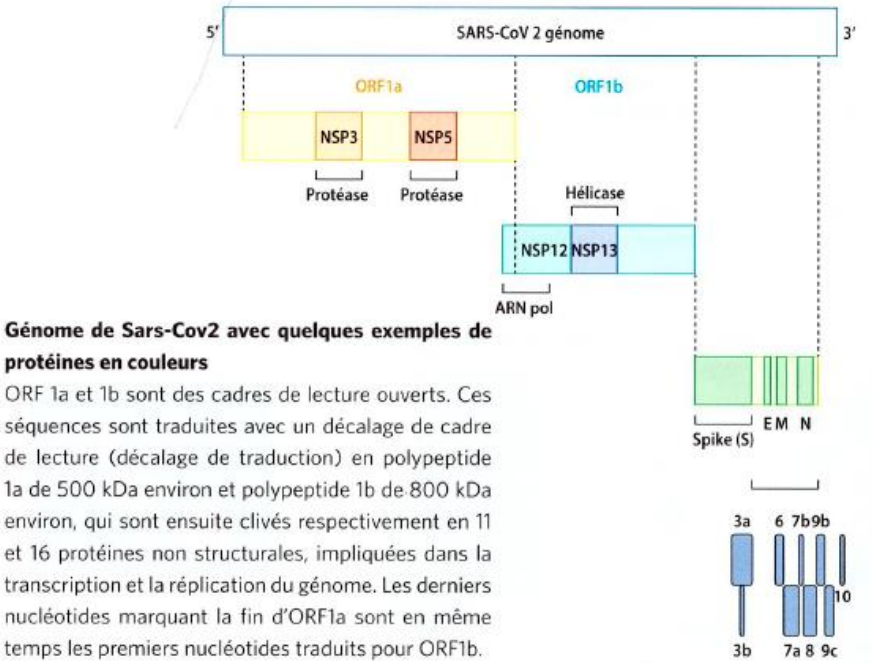


Fig. 29. Le génome du Sars-Cov-2. ★★ [4][1][5]



Références

- [1] ou absence de numéro : Ouvrage ou page Internet non précisé
- [2] G Furelaud, Y Esnault, F Fauzein. Le séquençage des génomes ; site Planet-Vie. <https://planet-vie.ens.fr/>
- [3] C. Perrier *et al.* Biologie BCPST1 tout-en-un. Dunod
- [4] O. Dautel *et al.* Biologie Géologie BCPST1. Vuibert
- [5] en.wikipedia.org