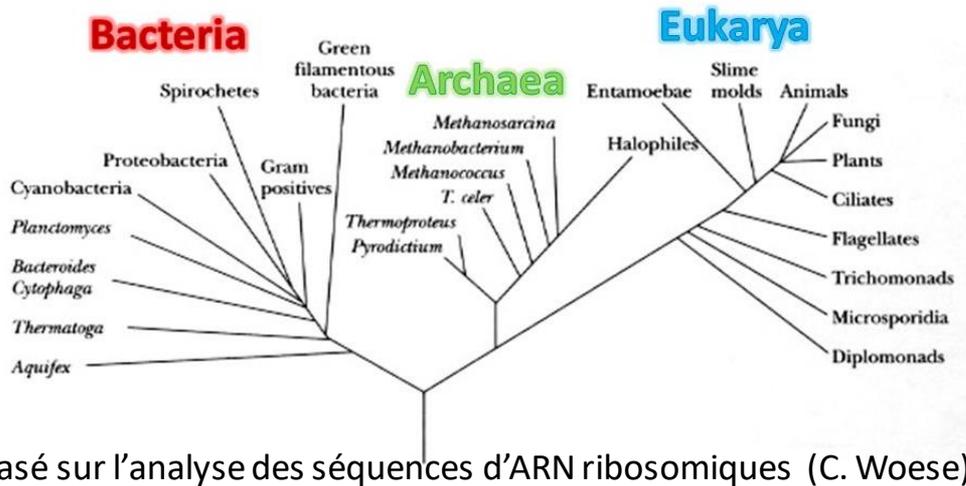


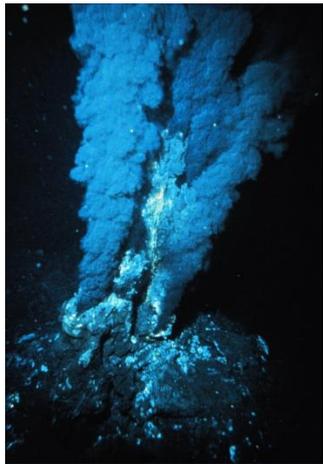
Introduction : les trois domaines du vivant



Les archaea ont été tout d'abord découvertes dans des environnements extrêmes comme :

- les sources hydrothermales (Thermococcales, Archeoglobales)
- les sources chaudes volcaniques
- les environnements avec des concentrations en sel très élevées (Halobacteriales)

Aujourd'hui découvertes dans de nombreux biotopes pas forcément extrêmes comme le sol, les océans (très nombreuses), la flore intestinale ...



Submarine geothermal vent, the habitat of *N. equitans*



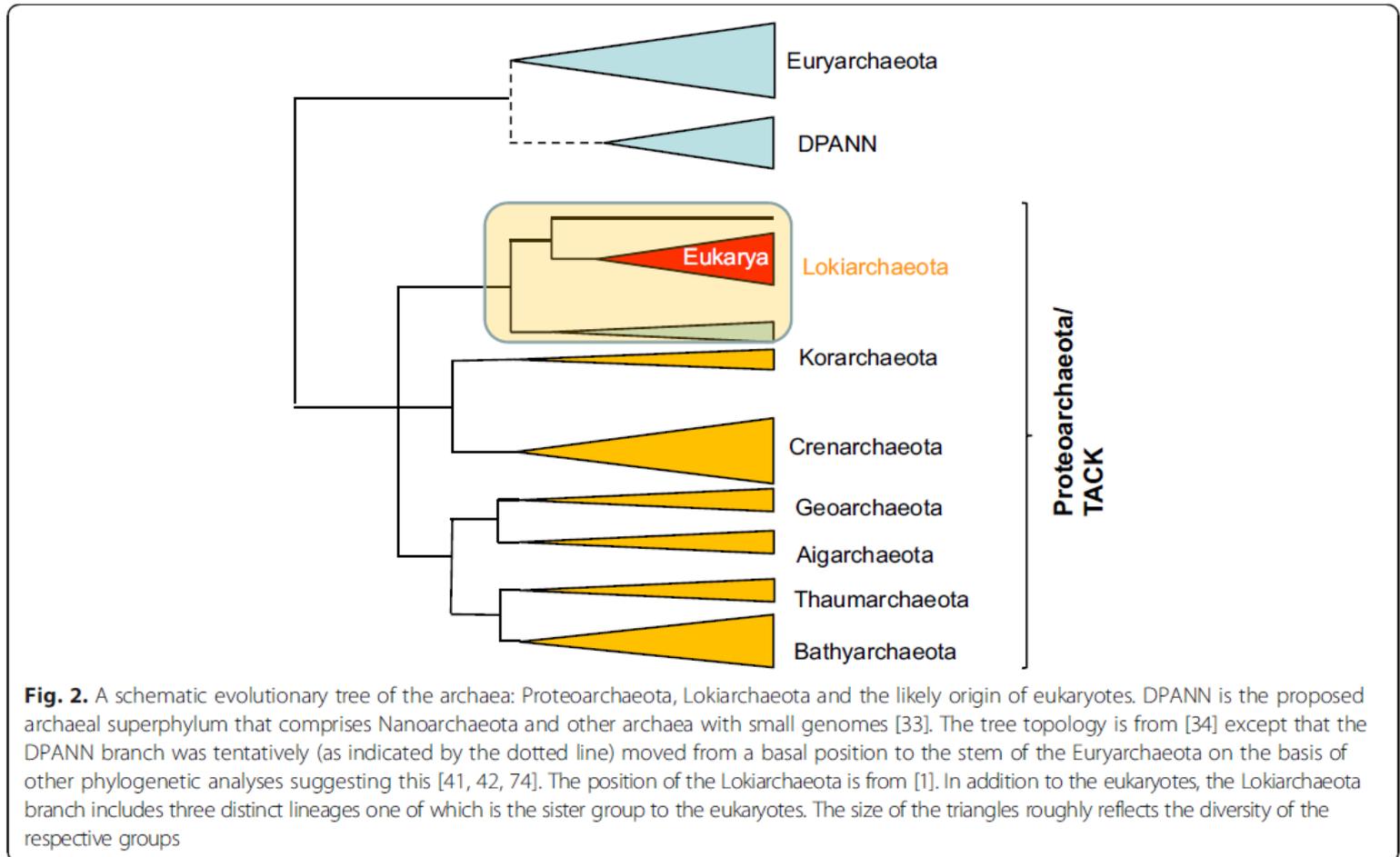
L'image représente " Grand Prismatic Spring" du parc national de Yellowstone



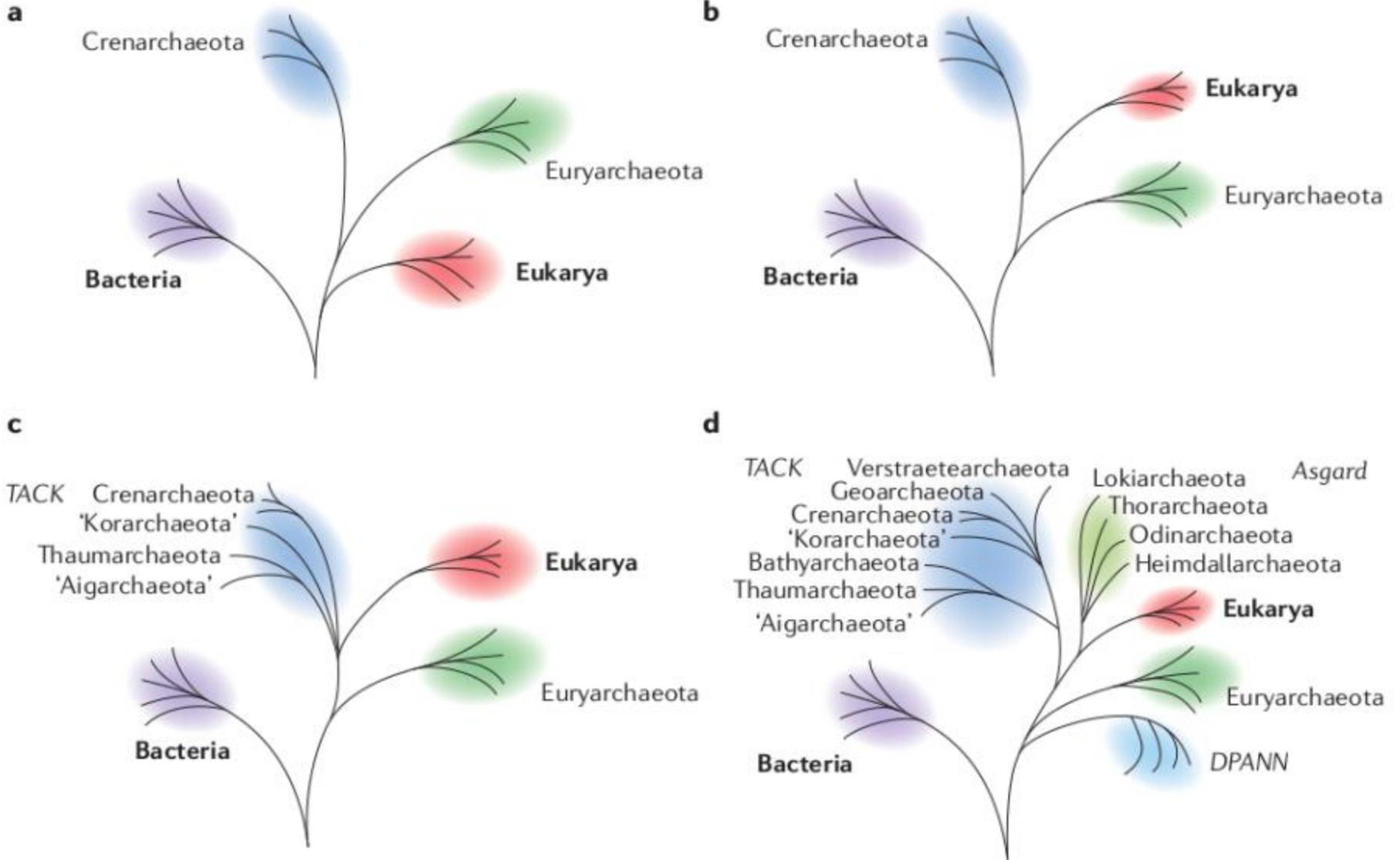
Mer morte, presque huit fois plus salée que les océans (275 g/l de NaCl). Habitat de *Haloferax volcanii*

Cheminée hydrothermale dans l'océan Atlantique

Les archés à l'origine des eucaryotes: hypothèse basée sur des données génome récentes



Archaea and the evolution of the tree of life



Les procaryotes

Les bactéries et les archaeabactéries

Organismes unicellulaires dont la structure cellulaire ne comporte pas de noyau ni d'organites contrairement aux eucaryotes.

Le matériel génétique chez les procaryotes est regroupé dans une zone appelée nucléoïde qui n'est pas séparée physiquement du reste de la cellule. Chez les eucaryotes, ce matériel est contenu dans le noyau.

Cellule bactérienne typique : taille de 0,5 à 5 μm (mais des exception comme l'espèce *Thiomargarita namibiensis* qui peut mesurer jusqu'à 0,5mm).

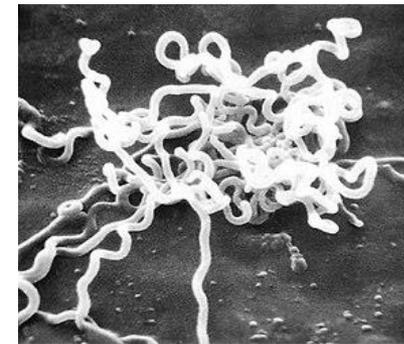
La plupart des bactéries sont soit sphériques (coques) ou en forme de bâtonnets (bacilles) mais d'autres formes existent.



Streptococcus pneumoniae

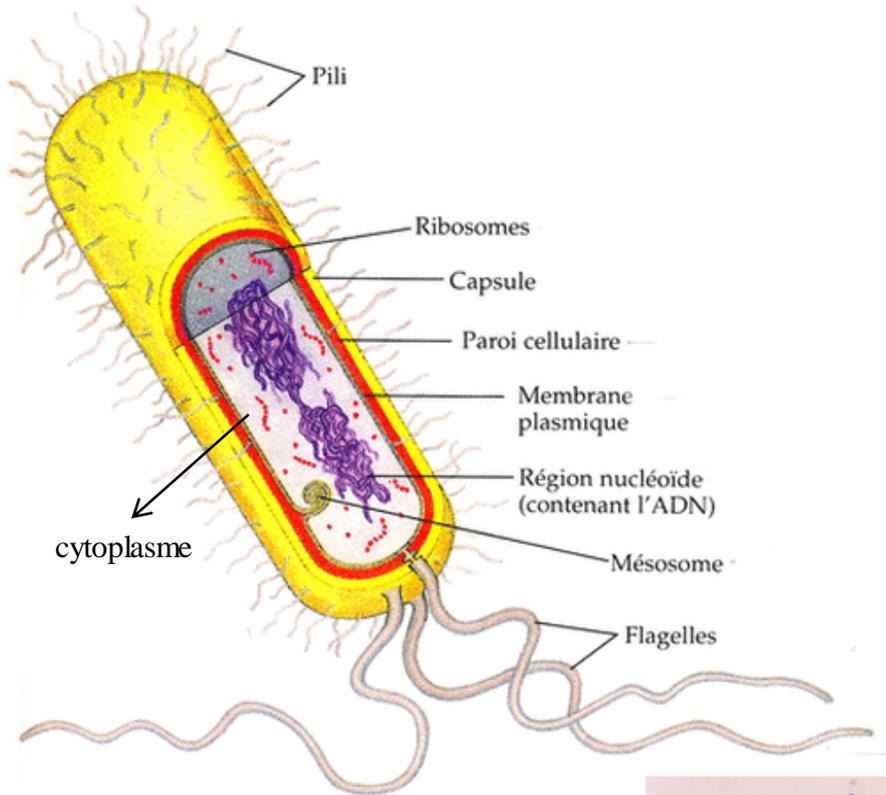


Escherichia coli



Treponema pallidum

La cellule procaryote



Plasmide : élément autonome pour la réplication. Code pour des fonctions d'adaptation à l'environnement.



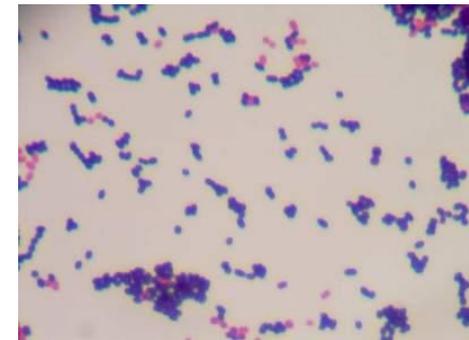
Bactérie Gram⁺ : *Bacillus subtilis*

Mésosome : invagination de la membrane plasmique.

La nature de la paroi cellulaire permet de diviser les bactéries en deux groupes (mise en évidence par la coloration gram) :

Bactérie à Gram positif (Gram⁺)

Bactérie à Gram négatif (Gram⁻)



Mélange :

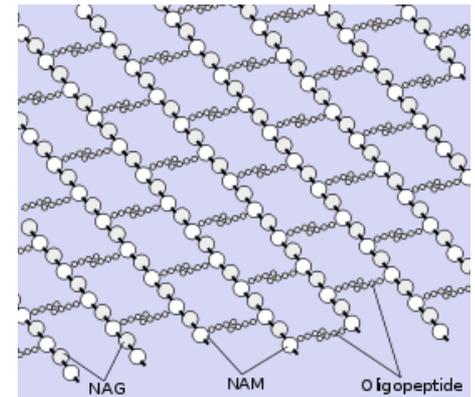
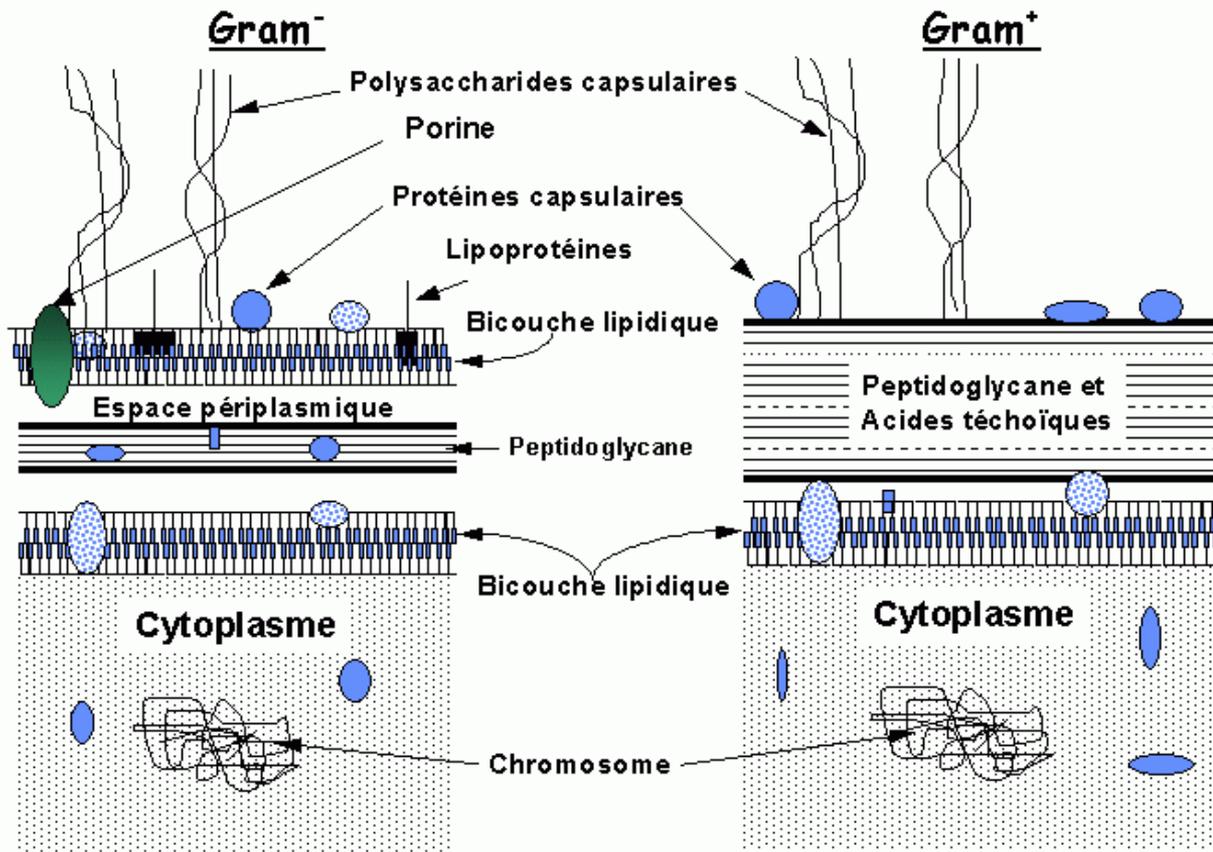
bactérie Gram⁺ : enterocoques

Bactéries Gram⁻ : Acinetobacter

La cellule procaryote

Bactérie Gram⁺ possède une paroi cellulaire formée d'une couche épaisse de peptidoglycane et d'acides téchoïques.

Bactérie Gram⁻ possède une mince couche de peptidoglycane entouré d'une membrane interne et d'une membrane externe.



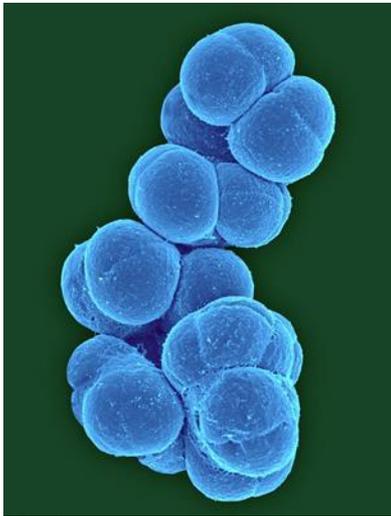
Structure du peptidoglycane

Structures de la membrane et de la paroi de peptidoglycane chez les bactéries Gram⁺/Gram⁻

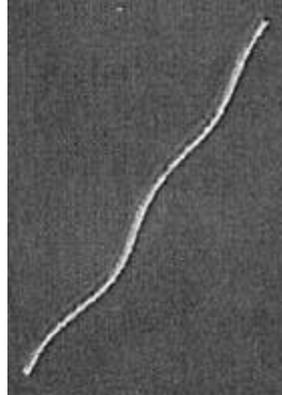
La cellule procaryote

Les archaea :

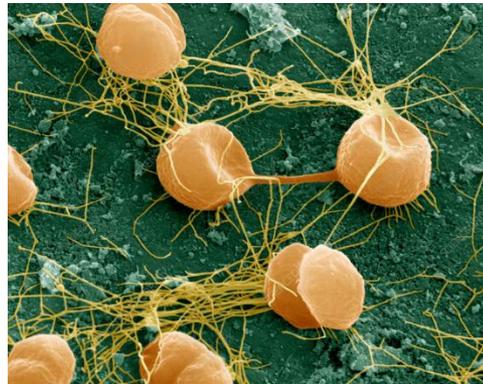
- composition chimique de la paroi différente de celle des bactéries (plusieurs types de composition existent)
- phospholipides membranaires ont des caractéristiques spécifiques ne se retrouvant ni chez les bactéries, ni chez les eucaryotes



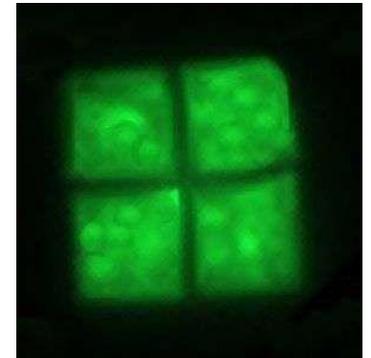
Methanosarcina barkeri



*Methanobacterium
thermoautotrophicum*



Pyrococcus furiosus

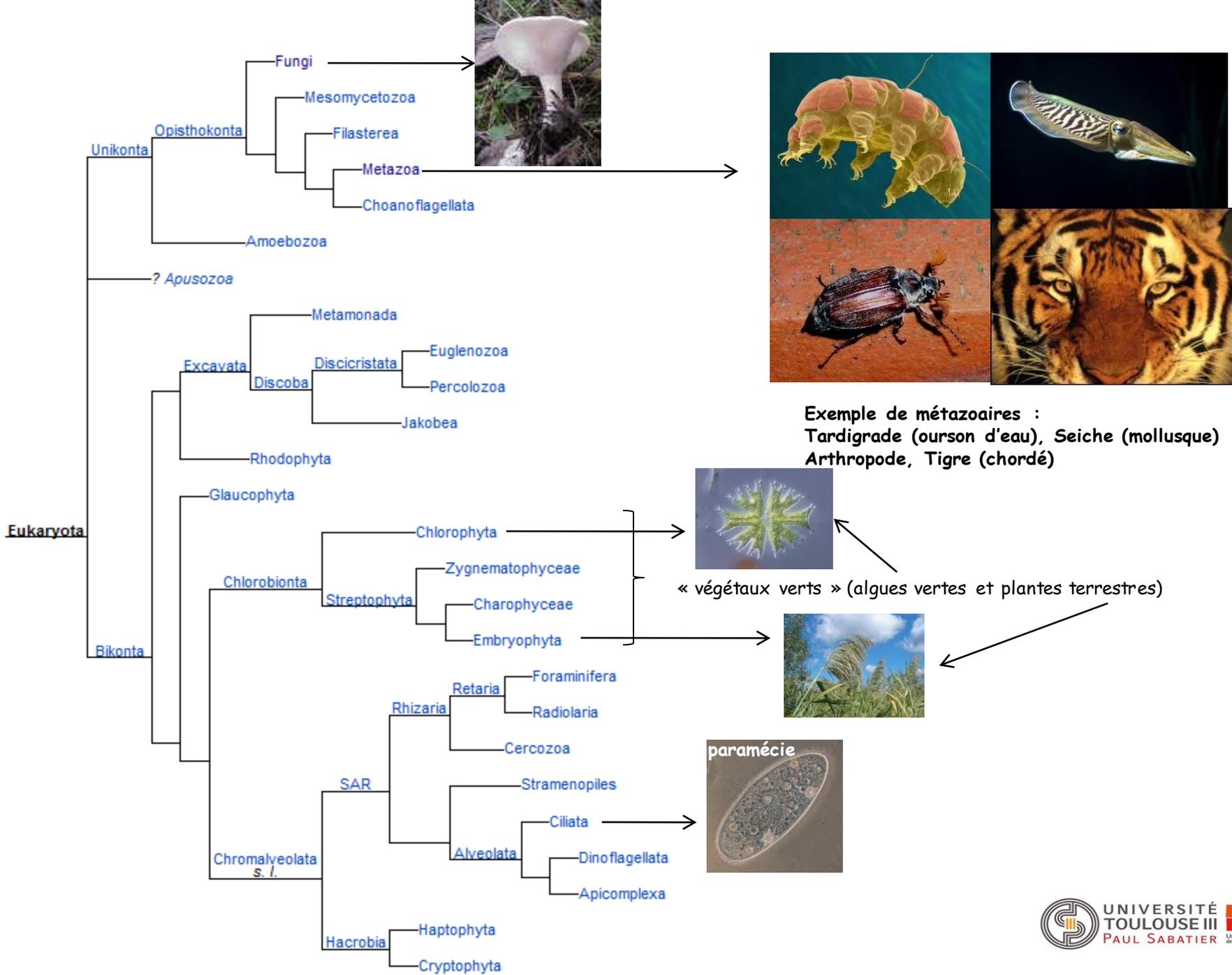


Haloquadratum walsbyi

Les eucaryotes

Domaine regroupant tous les organismes uni- ou pluricellulaire se caractérisant par la présence d'un noyau et de mitochondries dans les cellules.

Reproduction sexuée



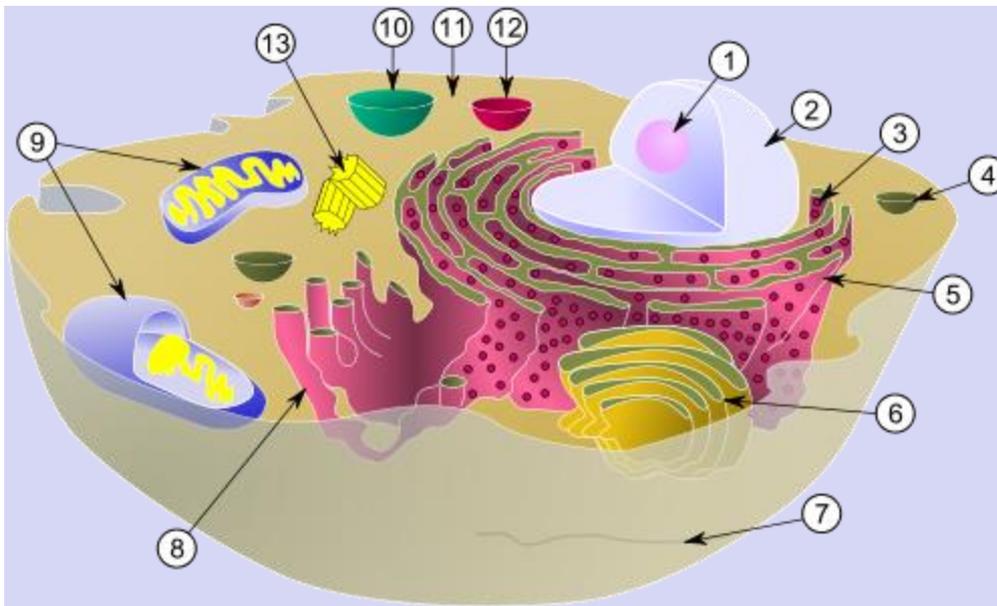
Exemple de métazoaires :
 Tardigrade (ourson d'eau), Seiche (mollusque)
 Arthropode, Tigre (chordé)

« végétaux verts » (algues vertes et plantes terrestres)

La cellule eucaryote

Elle se distingue de la cellule procaryotes notamment par l'existence de compartiments appelés organites :

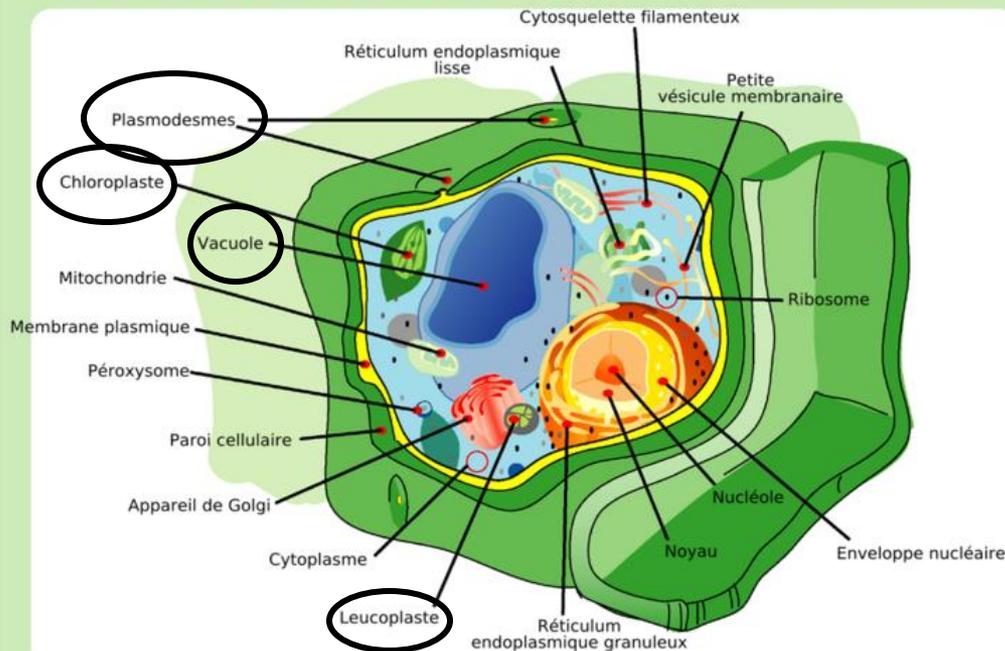
Cellule eucaryote typique



- 1 : nucléole : sous compartiment du noyau lieu de la transcription des ARN ribosomiques.
- 2 : noyau : organe contenant la majorité du matériel génétique
- 3 : ribosomes
- 4 : vésicule : Circulant dans le cytoplasme, elle peut stocker, transporter ou encore digérer des produits et déchets cellulaires.
- 5 : réticulum endoplasmique granuleux : site de la synthèse et du repliement des protéines
- 6 : appareil de Golgi : lieu où certaines protéines sont modifiées et rôle dans la sécrétion des protéines
- 7 : microtubule
- 8 : réticulum endoplasmique lisse : intervient dans plusieurs processus métaboliques (synthèse des lipides, métabolisme du calcium (stockage)), rôle dans la détoxification de la cellule...
- 9 : mitochondrie : « centrale énergétique », siège des deux dernières étapes de la respiration. Possède son propre matériel génétique (ADN mitochondrial).
- 10 : lysosome : digestion intracellulaire
- 11 : cytoplasme : désigne contenu de la cellule, émulsion colloïdale constituée par le cytosol, phase liquide où baigne les organites
- 12 : peroxyosome : détoxification de la cellule
- 13 : centrosome (centre organisateur des microtubules)

Caractéristiques de la cellule végétale :

Structure d'une cellule végétale



Grande vacuole centrale : maintient la turgescence de la cellule (eau entre dans la cellule, la vacuole se remplit et grossit. Ses membranes se tendent (turgescence des cellules). Cette eau exerce une pression sur la paroi des cellules et donne la rigidité aux parties sèches de la plante (tiges, feuilles ...). La vacuole sert aussi de poubelle à la cellule végétale. Afin de la vie de la cellule, la vacuole occupe 90% de l'espace cellulaire.

Les plasmodesmes : canaux traversant la paroi cellulaire. Ils relient les membranes plasmiques et les cytoplasmes des cellules adjacentes et permettent un passage contrôlé de molécules entre cellules adjacentes

Les plastes, en particulier les **chloroplastes** qui jouent un rôle dans la photosynthèse et qui contiennent la chlorophylle. Les plastes possèdent leur propre matériel génétique.

Une paroi pectocellulosique faite de cellulose et de protéines, ainsi que de lignine dans de nombreux cas. Elle protège la cellule végétale. Elle diffère de la paroi cellulaire des champignons, faite de chitine, et des procaryotes, faite de peptidoglycanes.

Ensemble du matériel génétique d'une espèce codé, dans la majorité des cas, par son ADN (exception de certains virus dont le génome est porté par des molécules d'ARN).

Chez les bactéries ou archaeae :

- génome en général contenu dans une molécule d'ADN circulaire -> un seul chromosome circulaire
- présence également de génomes extrachromosomiques contenu dans les plasmides
- des exceptions : existence de génomes linéaires (ex : *Streptomyces* pour les bactéries et *Methanobacterium thermoautotrophicum* pour les archaeae) et existence de procaryotes possédant plusieurs chromosomes (2 ou 3, ex: les *Burkholderia* (3 chromosomes))

Chez les eucaryotes :

- génome nucléaire contenu dans le noyau et formé de plusieurs chromosomes (46 chez l'homme).
- génomes non nucléaires :
 - le génome mitochondrial contenu dans les mitochondries (chez la quasi-totalité des eucaryotes)
 - le génome chloroplastique contenu dans les chloroplastes (chez les algues et plantes supérieures)

Taille des génomes

Exemple génomes procaryotes

Species	Genome size (Mb)
Bacteria	
<i>Mycoplasma genitalium</i>	0,58
<i>Heamophilus influenzae</i>	1,8
<i>Escherichia coli K12</i>	4,6
<i>Bacillus subtilis</i>	4,2
<i>Pseudomonas aeruginosa PAO1</i>	6,3
Archaea	
<i>Nanoarchaeum equitans</i>	0,49
<i>Aeropyrum pernix</i>	1,7
<i>Natronomonas pharaonis</i>	2,9
<i>Sulfolobus solfataricus P2</i>	3
<i>Methanosarcina mazei</i>	4,1

fritillaires



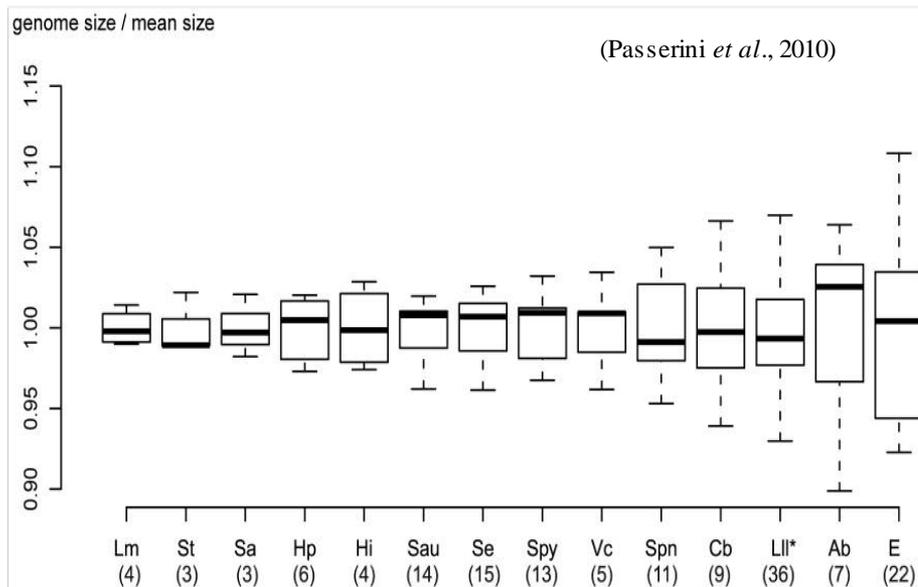
Exemple génomes eucaryotes

Species	Genome size (Mb)
Fungi	
<i>Saccharomyces cerevisiae</i>	12.1
<i>Aspergillus nidulans</i>	25.4
Protozoa	
<i>Tetrahymena pyriformis</i>	190
Invertebrates	
<i>Caenorhabditis elegans</i>	97
<i>Drosophila melanogaster</i>	180
<i>Bombyx mori</i> (silkworm)	490
<i>Strongylocentrotus purpuratus</i> (sea urchin)	845
<i>Locusta migratoria</i> (locust)	5000
Vertebrates	
<i>Takifugu rubripes</i> (pufferfish)	400
<i>Homo sapiens</i>	3200
<i>Mus musculus</i> (mouse)	3300
Plants	
<i>Arabidopsis thaliana</i> (vetch)	125
<i>Oryza sativa</i> (rice)	430
<i>Zea mays</i> (maize)	2500
<i>Pisum sativum</i> (pea)	4800
<i>Triticum aestivum</i> (wheat)	16 000
<i>Fritillaria assyriaca</i> (fritillary)	120 000

Taille des génomes

Cependant, au sein d'une même espèce bactérienne, la taille du génome peut varier considérablement :

Exemple parmi 42 souches d'*Escherichia coli*, on observe une variation de la taille du génome allant de 2,6 Mb à 5,7 Mb



Lm, *L. monocytogenes*; St, *S. thermophilus*; Sa, *S. agalactiae*; Hp, *H. pylori*; Hi, *H. influenzae*; Sau, *S. aureus*; Se, *S. enterica*; Spy, *S. pyogenes*; Vc, *V. cholerae*; Spn, *S. pneumoniae*; Cb, *C. botulinum*; Lll, *L. lactis* subsp. *lactis*; Ab, *A. baumannii*; Ec, *E. coli*.

L'ADN : support de l'information génétique

Le génome est composé de régions codantes, les gènes, et de régions non codantes.

Les gènes codent :

- des protéines
- des ARN ribosomiques
- des ARN de transfert

Le nombre de gènes dans un génome varie moins que sa taille, cependant la corrélation avec la complexité de l'organisme n'est pas parfaite.

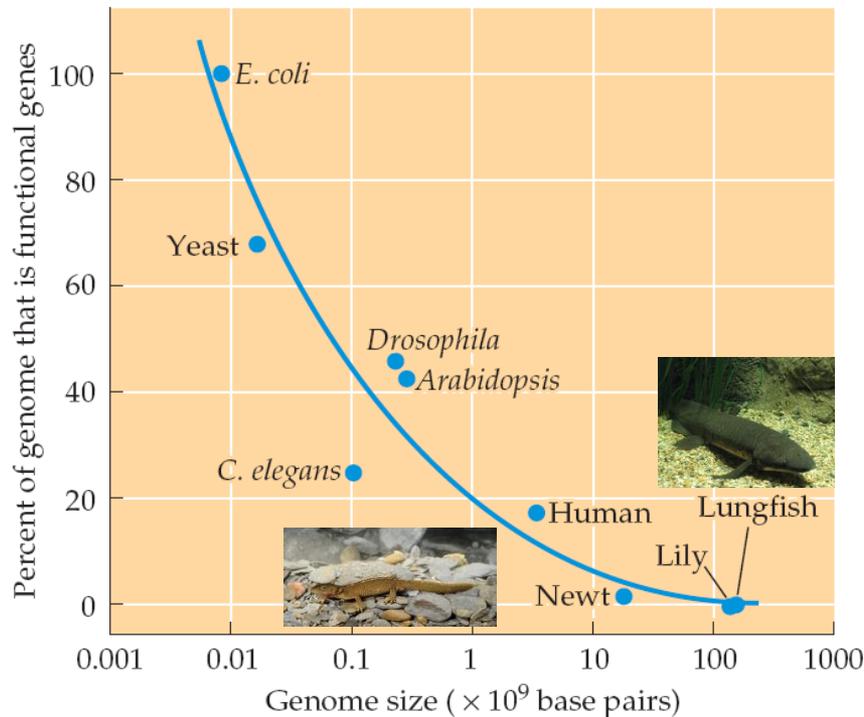
Species	Genome size (Mb)	Nombre approximatif de gènes
<i>Mycoplasma genitalium</i> (bactérie)	0,58	500
<i>Haemophilus influenzae</i> (bactérie)	1,8	1800
<i>Escherichia coli</i> K12 (bactérie)	4,6	4400
<i>Saccharomyces cerevisiae</i> (levure)	12	5800
<i>Drosophila melanogaster</i> (insecte)	180	14700
<i>Caenorhabditis elegans</i> (nematode)	103	20000

Species	Genome size (Mb)	Nombre approximatif de gènes
<i>Tetrahymena thermophila</i> (protozoaire)	125	2700
<i>Arabidopsis thaliana</i> (plante à fleur)	120	26500
<i>Oryza sativa</i> (riz)	430	~45000
<i>Zea mays</i> (maïs)	2200	> 45000
<i>Mus musculus</i> (souris)	2600	22000
<i>Homo sapiens</i>	3200	22000
Paramécie	72	40000

Différence de taille des génomes

Due bien sur à la différence du nombre de gènes codés par le génome mais pas seulement. En effet, il existe aussi dans le génome des régions non codantes. Cet ADN non codant peut être localisé soit :

- entre les gènes -> régions intergéniques (procaryotes et eucaryotes)
- à l'intérieur des gènes -> introns (eucaryotes)



ADN non codant comprend de multiples structures comme :

Pseudogènes

Éléments mobiles

Répétitions en tandem

« microARNs »

26.8 A Large Proportion of DNA Is Noncoding Most of the DNA of bacteria and yeasts encodes RNAs or proteins, but most of the DNA of more complex organisms is noncoding. Most noncoding DNA is probably nonfunctional.

Le « dogme central » de la biologie moléculaire



Les flèches indiquent le sens pour le transfert de l'information génétique. Terme « dogme central » employé en 1956 par Francis Crick.

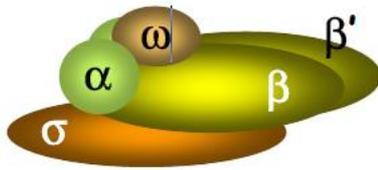
L'ARN peut parfois servir de matrice pour la synthèse d'un brin d'ADN complémentaire. Cependant, évènement rare comparé à la grande quantité de molécules d'ARN synthétisées à partir d'une matrice ADN.

Donc ce flux orienté de l'information reste d'actualité.

La transcription

Elle est réalisé par une enzyme : l'ARN polymérase qui réalise essentiellement les mêmes réactions de la bactérie à l'homme.

Chez les bactéries, une seule ARN polymérase constituée de 5 sous-unités, deux exemplaires de la sous-unité α , un exemplaire de chacune des sous-unité β , β' et ω . le facteur σ permet la reconnaissance des séquences spécifiques sur l'ADN, appelée **promoteur**.



ARN polymérase procaryote

Chez les eucaryotes, présence de trois ARN polymérases :

ARN polymérase	Synthèse	Constituants
I	- ARN pré-ribosomique	Ensemble de 12 sous-unités plus 2 autres polypeptides
II	- ARN messagers - snoRNA*	Ensemble de 12 sous-unités
III	- Les ARN de transfert - ARN ribosomique 5S - Petits ARN	Ensemble de 12 sous-unités plus 5 autres polypeptides

La transcription

3 étapes :

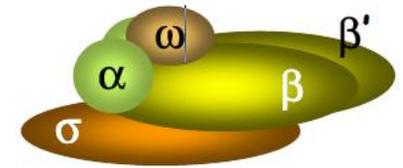
- l'initiation : reconnaissance de séquences spécifiques sur l'ADN : le promoteur
- l'élongation
- la terminaison

Transcription des gènes codant pour des protéines

Chez les bactéries :

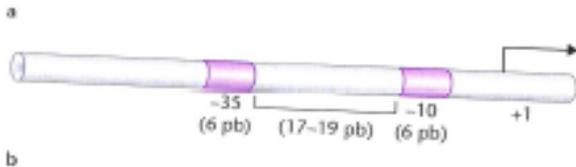
L'initiation :

Le facteur σ de la RNA polymérase reconnaît deux régions constituant le promoteur : la région -35 et la région -10 (TATAAT box) séparées par une région de taille variable (17 à 19 pb). Le promoteur se trouve en amont du début du (des gènes).

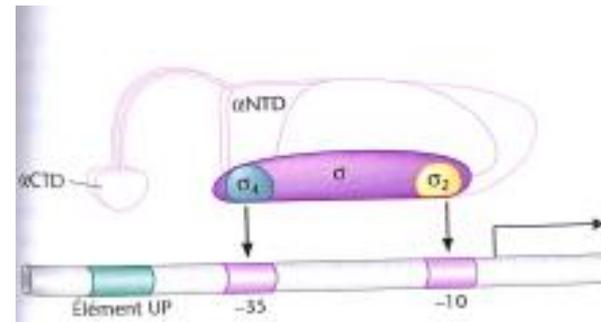


ARN polymérase procaryote

Promoteurs bactériens



Recrutement du cœur de l'ARN polymérase au promoteur

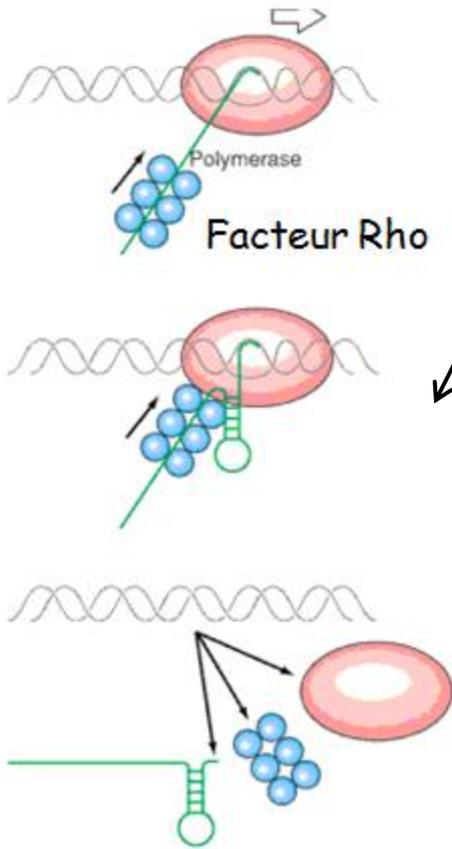


La transcription

Transcription des gènes codant pour des protéines

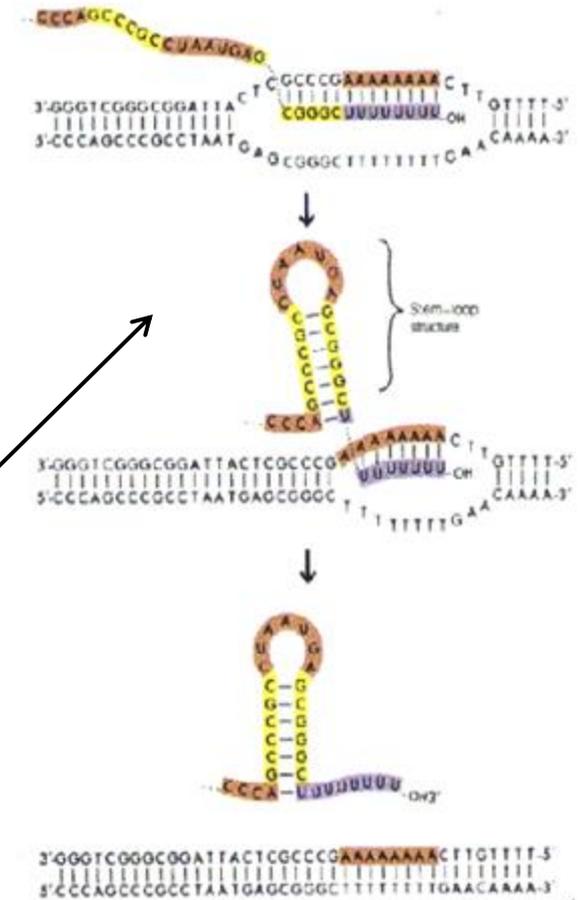
Chez les bactéries :

La terminaison : deux types de terminateurs Rho-dépendants et Rho-indépendants



Rho-dépendant : L'ATPase Rho, une protéine qui se déplace le long du transcrit naissant jusqu'à rattraper la polymérase, stimule la terminaison de la transcription

Rho-indépendant : deux éléments une courte séquence répétée inversée suivie d'une série d'environ 8 paires de bases A-T. Le transcrit est capable de former une structure secondaire tige-boucle qui provoque la désorganisation de l'ARN polymérase. La suite du transcrit est faiblement apparié au brin ADN ce qui conduit à sa libération.

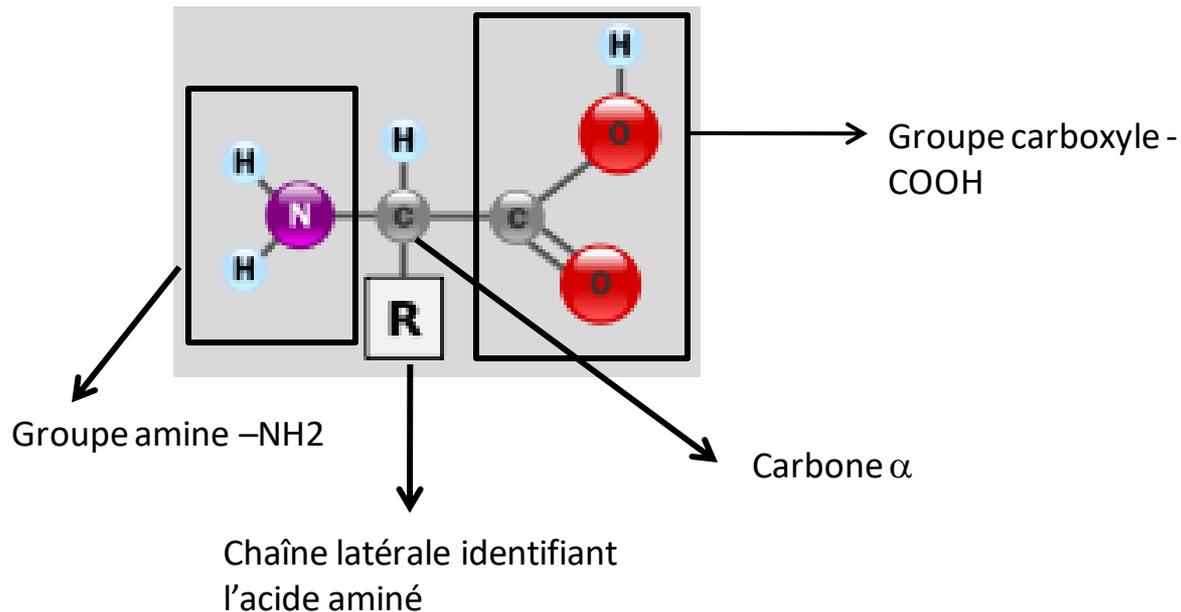


La traduction

Etape qui permet la synthèse des protéines à partir de l'ARN messager.

Une protéine est une macromolécule biologique composée d'acides aminés. Il existe 20 acides aminés qui sont rencontrés dans tous les organismes.

Structure générique d'un acide aminé



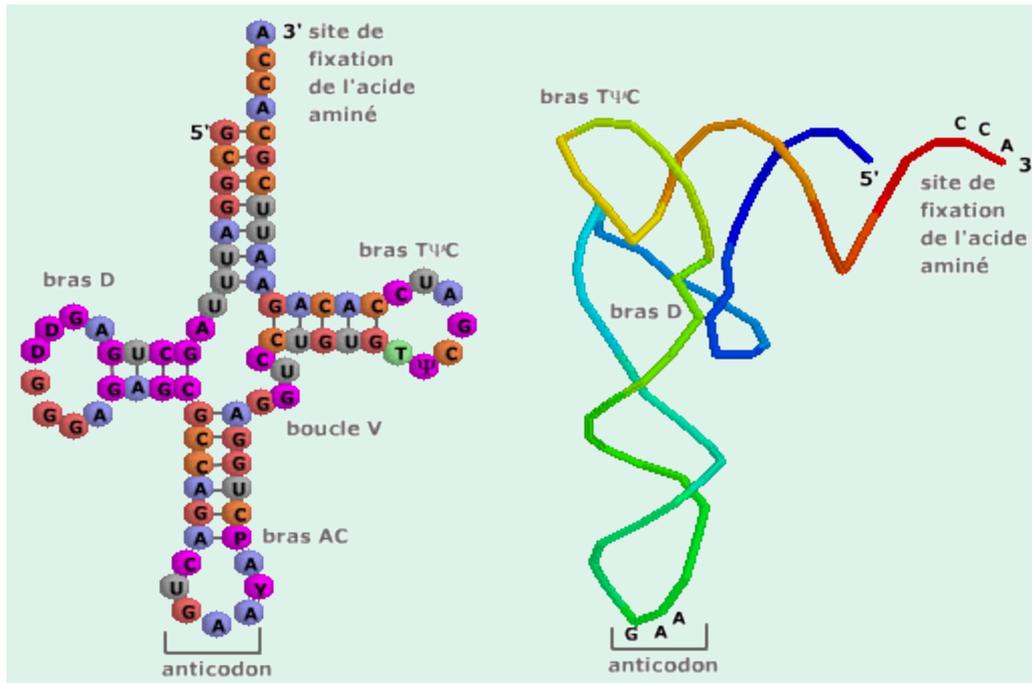
Les acides aminés polymérisent en formant des liaisons peptidiques

4 composants principaux font partie de la machinerie de traduction :

- l'ARN messager (ARNm) contenant la séquence codant la protéine ainsi que des régions de reconnaissance pour l'initiation et la terminaison de la traduction.
- les ARN de transfert (ARNt) qui une fois chargé de leur acide aminé (processus réalisé par les aminocyl-ARNt- synthétases) vont mettre en relation un codon et un acide aminé.
- les aminocyl-ARNt- synthétases permettant donc de charger l'acide aminé sur l'ARNt
- le ribosome : gros complexe formé de deux sous-unités (une grande et une petite) chacune constituées par des protéines ribosomiques et un ou plusieurs ARN ribosomique (ARNr).

La traduction : les ARNt

Ce sont de petites molécules d'ARN, d'une taille variant entre 70 et 100 nucléotides. Ils se replient en une structure secondaire caractéristique appelée feuille de trèfle et présente une structure tridimensionnelle en forme de L.



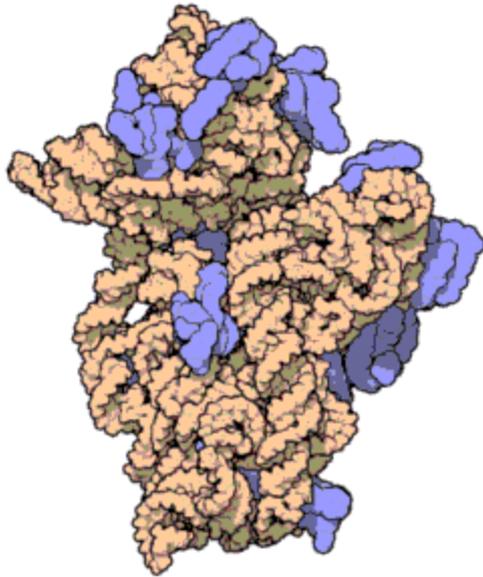
L'anticodon s'apparie avec le codon de l'ARNm. L'ARNt chargé portant à son extrémité 3' l'acide aminé correspondant permet donc de faire la correspondance codon-acide aminé. Après reconnaissance du codon, l'acide aminé est transféré à la chaîne peptidique en croissance. Ce processus est réalisé à l'intérieur du ribosome.

Structures secondaire et tertiaire d'un ARNt

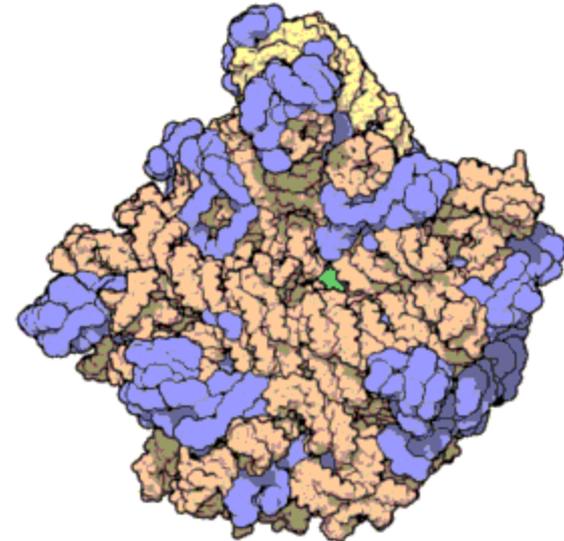
La traduction : le ribosome

Le ribosome est constitué de deux sous-unités : la grande et la petite sous-unité.

Les procaryotes ont des ribosomes 70S constitué d'une petite sous-unité 30S et d'une grande sous-unité 50S. La petite sous-unité est composée de l'ARNr 16S (~1500 nt) et de 21 protéines. La grande sous-unité possède deux molécules d'ARNr, l'ARNr 5S (~120 nt) et l'ARNr 23S (~2900 nt), et 31 protéines.



Petite sous unité 30S de *Thermus thermophilus*. Bleue protéines, orange ARNr 16S



Grande sous unité 50S de *Haloarcula marismortui*. Bleue protéines, orange et jaune les deux ARNr. En vert au centre de la sous-unité, le site actif

La traduction : le ribosome

Les eucaryotes ont des ribosomes 80S constitué d'une petite sous-unité 40S et d'une grande sous-unité 60S. La petite sous-unité est composée de l'ARNr 18S (~1900 nt) et de 33 protéines. La grande sous-unité possède trois molécules d'ARNr, l'ARNr 5S (~120 nt), l'ARNr 28S (~4700 nt), et l'ARNr 5.8S (~160 nt) et de ~49 protéines.

Les ribosomes trouvés dans les mitochondries et chloroplastes forment aussi une particule 70S. Ces organelles sont supposées descendre de bactéries et auraient été acquises par endosymbiose.

La petite sous-unité contient le centre de décodage dans lequel les ARNt chargés décodent les codons de l'ARNm

La grande sous-unité contient le centre responsable de la formation des liaisons peptidiques entre acides aminés et donc de la croissance de la chaîne peptique.

La traduction : le ribosome

Le ribosome : une usine de synthèse des protéines.

Il contient 3 sites de fixation pour les ARNt :

- le site A lie les ARNt chargés de leur acide aminé (ARNt aminoacylés)
- le site P lie les ARNt liés à la chaîne peptidique en cours de synthèse (peptidyl ARNt)
- le site E lie les ARNt libres (déchargé et décroché la chaîne peptidique) avant leur sortie du ribosome (E pour « exit »).

La traduction comprend également 3 étapes :

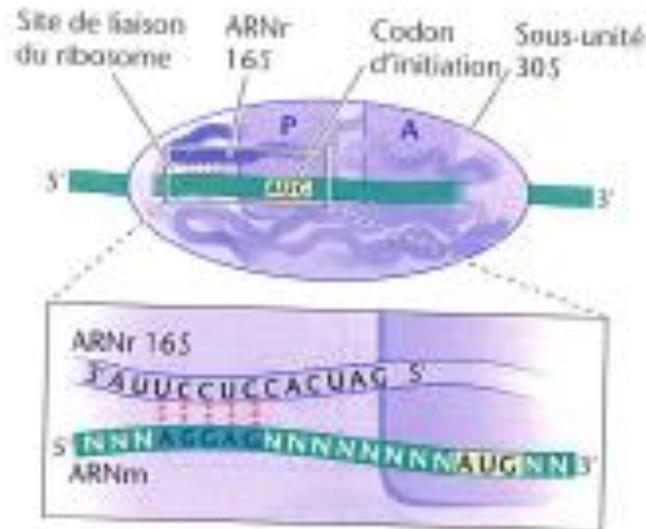
- l'initiation qui charge le ribosome sur l'ARNm
- l'élongation qui conduit à la synthèse de la protéine
- la terminaison conduisant à la dissociation du ribosome de l'ARNm

L'initiation :

➤ chez les bactéries :

La petite sous-unité est chargée en premier sur l'ARNm, ceci grâce à un appariement des bases d'une région de l'ARNr 16S avec le RBS. Pour un RBS idéalement situé, le codon initiateur (AUG, GUG ou UUG) se trouve situé au site P du ribosome (et non au A comme pour l'élongation). Ceci requiert un ARNt particulier, appelé ARNt initiateur. Cet ARNt ne porte ni la méthionine, ni la valine, ni la leucine comme acide aminé, mais une méthionine modifiée (N-formylméthionine) d'où son nom ARNt fMet.

Quand l'ARNt fMet s'apparie au site P, il y a un changement de conformation de la petite sous-unité qui fait que la grande sous-unité peut se lier à elle pour former le ribosome 70S.

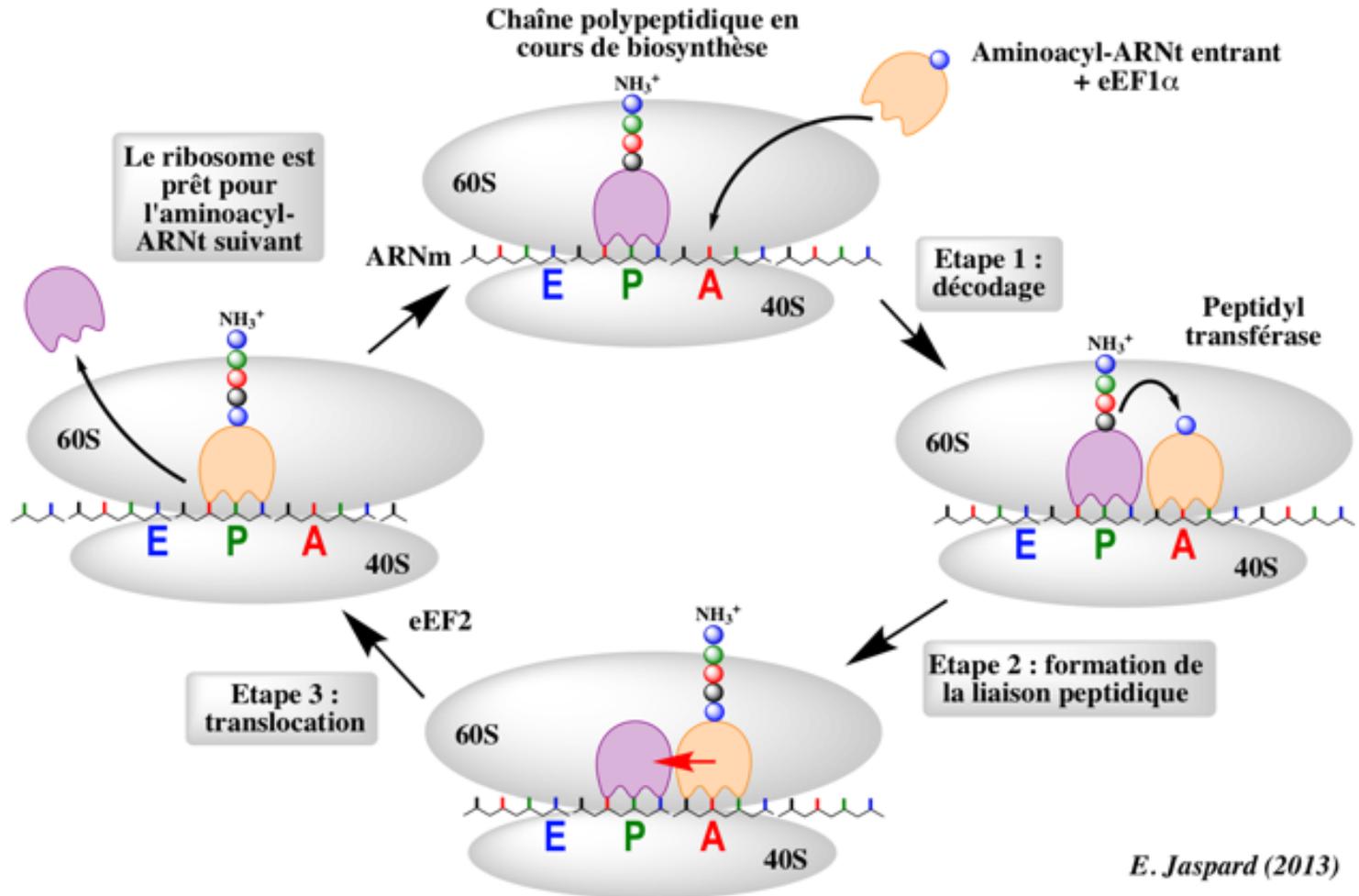


Interaction RBS et ARNr 16S pour positionnement du codon AUG au site P

RBS = Ribosome Binding Site

La traduction

L'élongation :

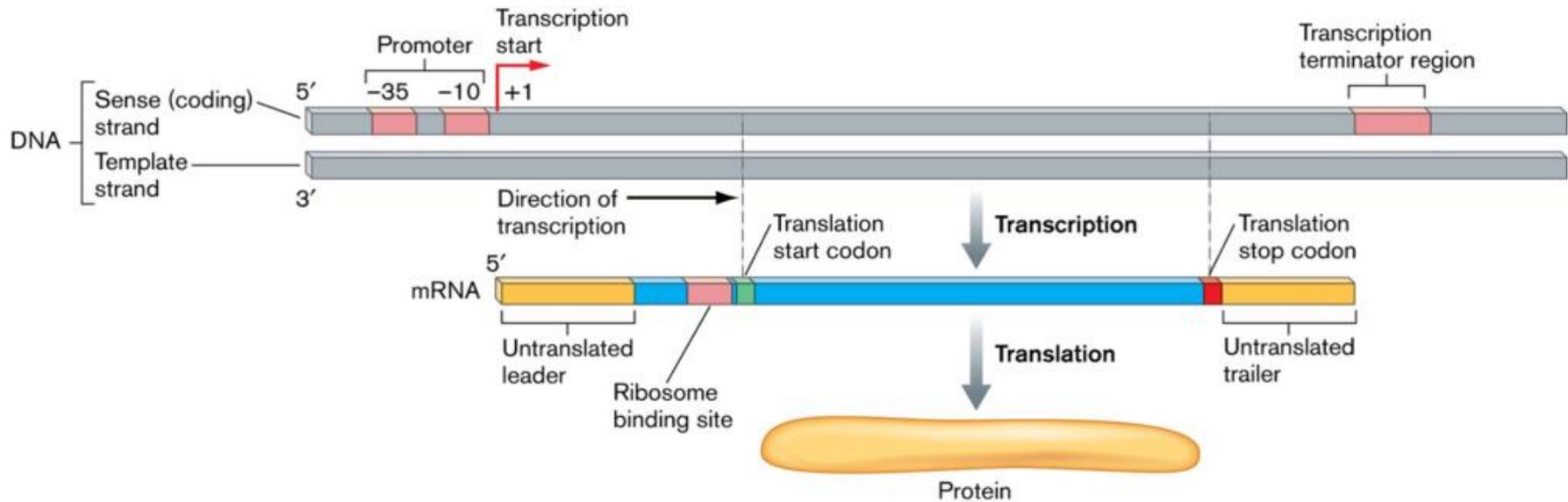


La terminaison :

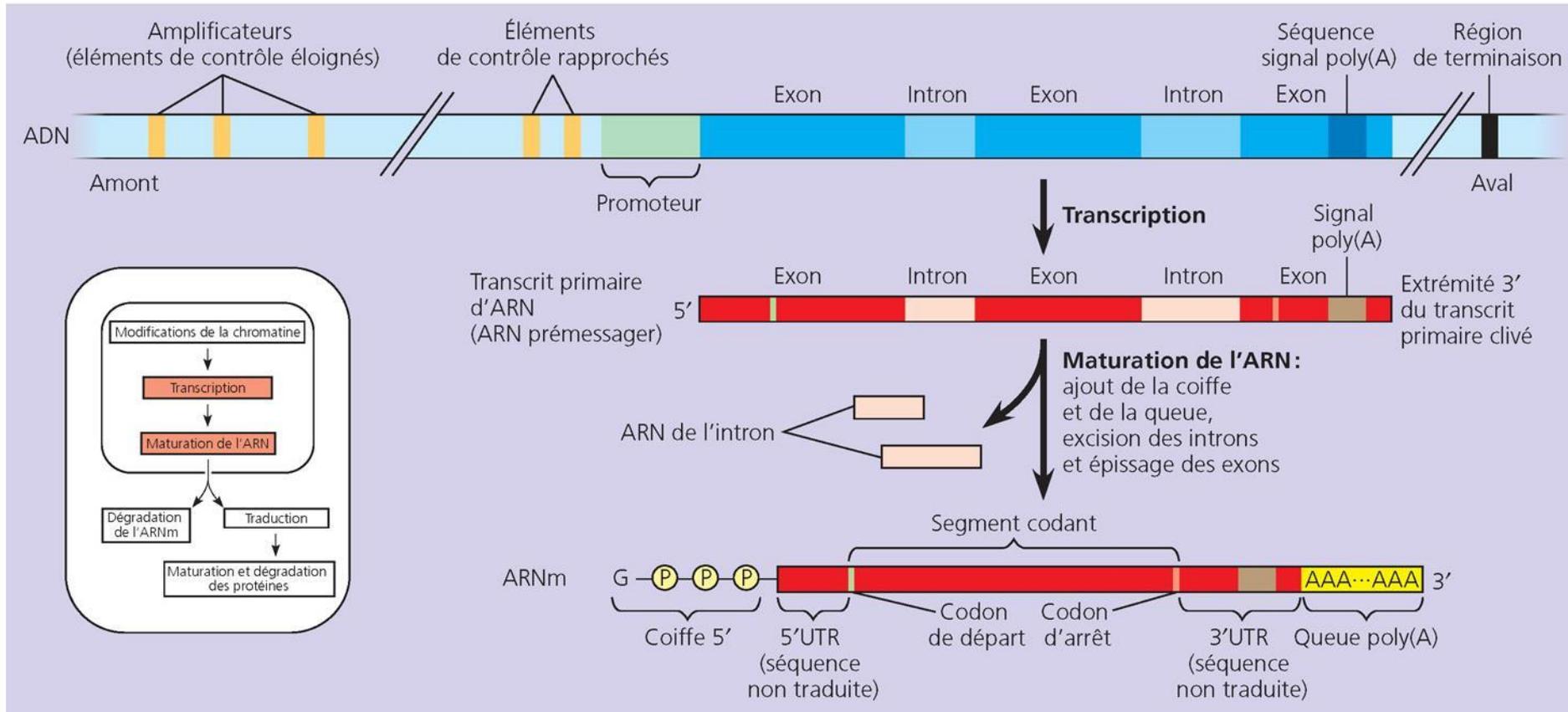
La traduction s'arrête quand le ribosome rencontre un codon stop au site A.

Ce codon stop est reconnu par un des deux facteurs de terminaison de classe 1 chez les procaryotes. Ce facteur stimule l'hydrolyse du polypeptide et du peptidyl-ARNt, libérant ainsi le peptide complet.

Structure d'un gène bactérien



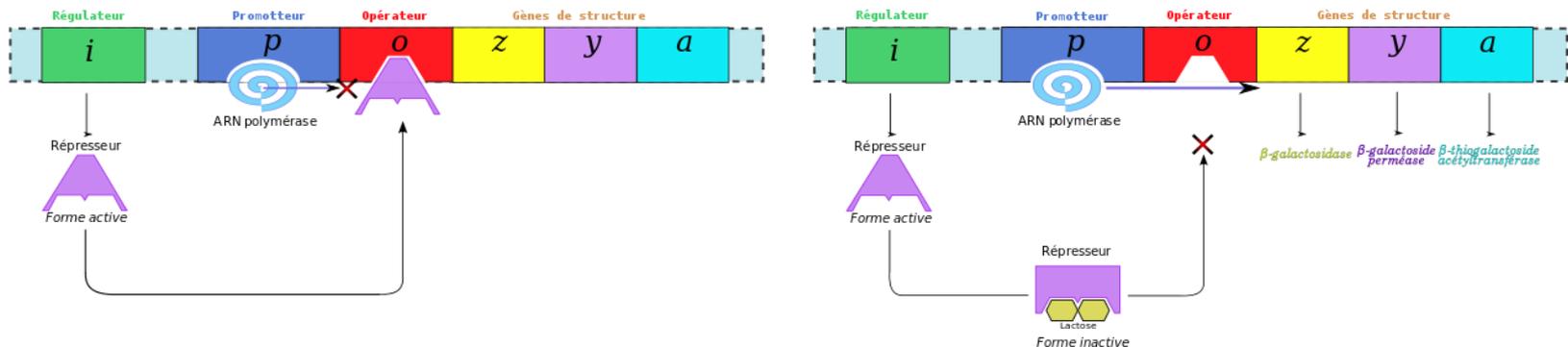
Structure d'un gène eucaryote



Différence entre procaryotes et eucaryotes :

- **localisation** : la transcription a lieu dans le cytoplasme chez les bactéries car pas de compartiment. Comme l'étape suivante de traduction a lieu au même endroit, elle peut commencer avant que la transcription soit terminée. Chez les eucaryotes, la transcription se déroule dans le noyau ainsi que sa maturation. L'ARNm mature est ensuite exporté dans le cytoplasme où il est traduit.
- Dans la majorité des génomes procaryotes, les gènes sont organisés en opérons. L'opéron comprend : i) une région régulatrice (opérateur), région ADN sur laquelle se fixe un activateur ou un répresseur, ii) un promoteur commun et iii) un ou plusieurs gènes. Ces gènes sont donc contrôlés de façon coordonnée et ils sont co-transcrits formant un seul ARNm dit polycistronique.

Exemple de l'opéron lactose



Par NicolasGrandjean — Création personnelle, CC BY-SA 3.0, <https://commons.wikimedia.org/w/index.php?curid=800126>

- Chez les eucaryotes, les gènes sont transcrits de façon individuelle donnant un ARNm appelé monocistronique.

La transcription

Transcription des gènes codant pour des protéines

Chez les eucaryotes réalisée par l'ARN polymérase II :

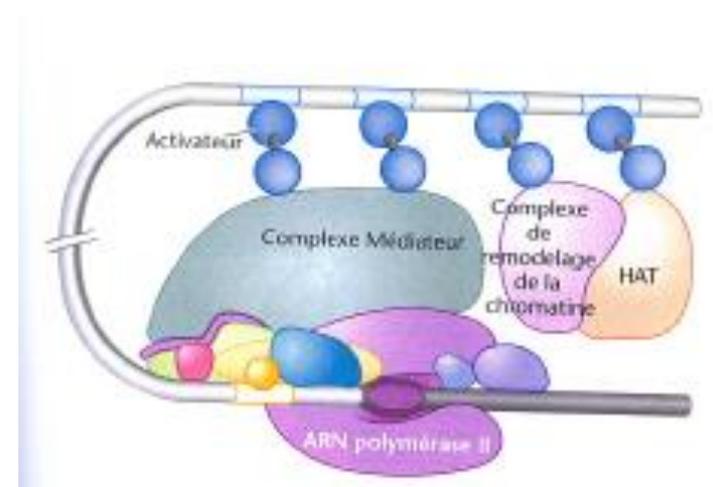
L'initiation :

Elle nécessite un plus grand nombre de facteurs que dans le cas des procaryotes où il n'y en a qu'un seul. Ces facteurs sont appelés **facteurs généraux de transcription**.

Cependant, l'ADN étant empaqueté sous forme de chromatine, des facteurs supplémentaires sont requis pour que les facteurs généraux puissent avoir accès au promoteur. Des protéines régulatrices de la transcription, appelées activateurs, aident au recrutement de l'ARN polymérase au promoteur. Le recrutement est réalisé grâce à des interactions entre ces activateurs, des facteurs modifiant la chromatine et des parties de la machinerie de transcription.

Une de ces interactions implique le **Médiateur**.

C'est un complexe formé de nombreuses protéines (plus de 20). Il est associé par une de ces faces avec la queue CTD de la grande sous-unité de l'ARN polymérase (absente chez les procaryotes) et ces autres faces sont impliquées dans des interactions avec les activateurs.



Transcription des gènes codant pour des protéines

Chez les eucaryotes réalisée par l'ARN polymérase II :

Terminaison:

Liée à l'étape de polyadénylation. Quand la polymérase atteint la fin d'un gène, elle rencontre des séquences spécifiques qui, une fois traduites en ARN, sont reconnues par les facteurs de polyadénylation. Ceci engendre les 4 évènements suivants :

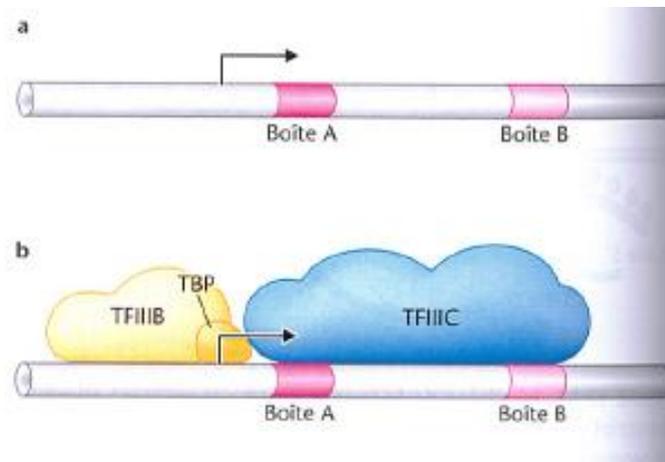
- le clivage de l'ARNm
- l'ajout d'un grand nombre de résidus adénine à l'extrémité 3'
- la dégradation de l'ARN encore associé à la polymérase
- la terminaison de la transcription.

La transcription

Les promoteurs de l'ARN Pol III présentent différentes structures mais ont comme caractéristique d'être localisés après le site d'initiation de la transcription.

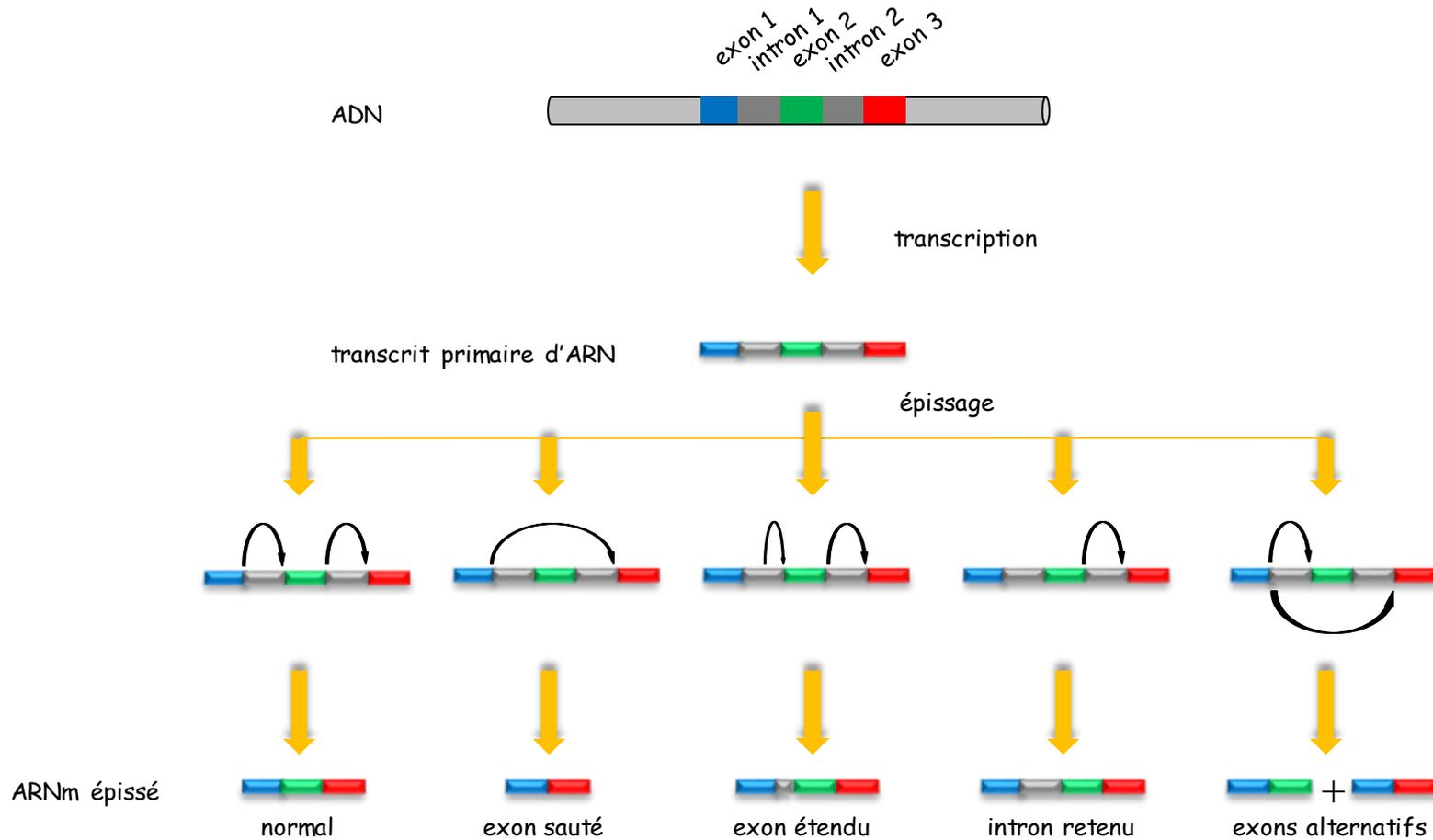
Certains promoteurs comme ceux des gènes codant les ARNt sont composés de deux boîtes A et B ; d'autres contiennent la boîte A et une boîte C (ex, le gène codant l'ARNr 5S); d'autres contiennent une boîte TATA comme ceux de la Pol II.

Ces promoteurs sont reconnus par des facteurs spécifiques : TFIIB et TFIIC pour les gènes d'ARNt plus le facteur TFIIA pour le gène d'ARNr 5S.



Promoteur schématisé d'un gène de levure codant un ARNt

Cinq manières d'épisser un ARN



La génomique

La génomique a pour sujet d'étude les génomes. Elle travaille au séquençage des génomes et à l'étude des fonctions associées.

CONNAITRE LE GENOME

La génomique

Structure physique

Structure génétique

Les espèces modèles

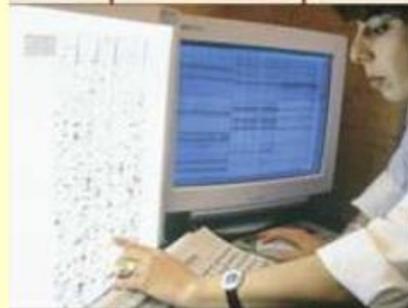
Plantes
Champignons
Procaryotes
Animaux

↓
Un ancêtre commun ?
LUCA

Last Universal Common Ancestor

Les objectifs

- Localiser et séquencer les gènes



- Étudier la fonction des gènes

Analyse fonctionnelle

Les applications

Construction de génotypes élités

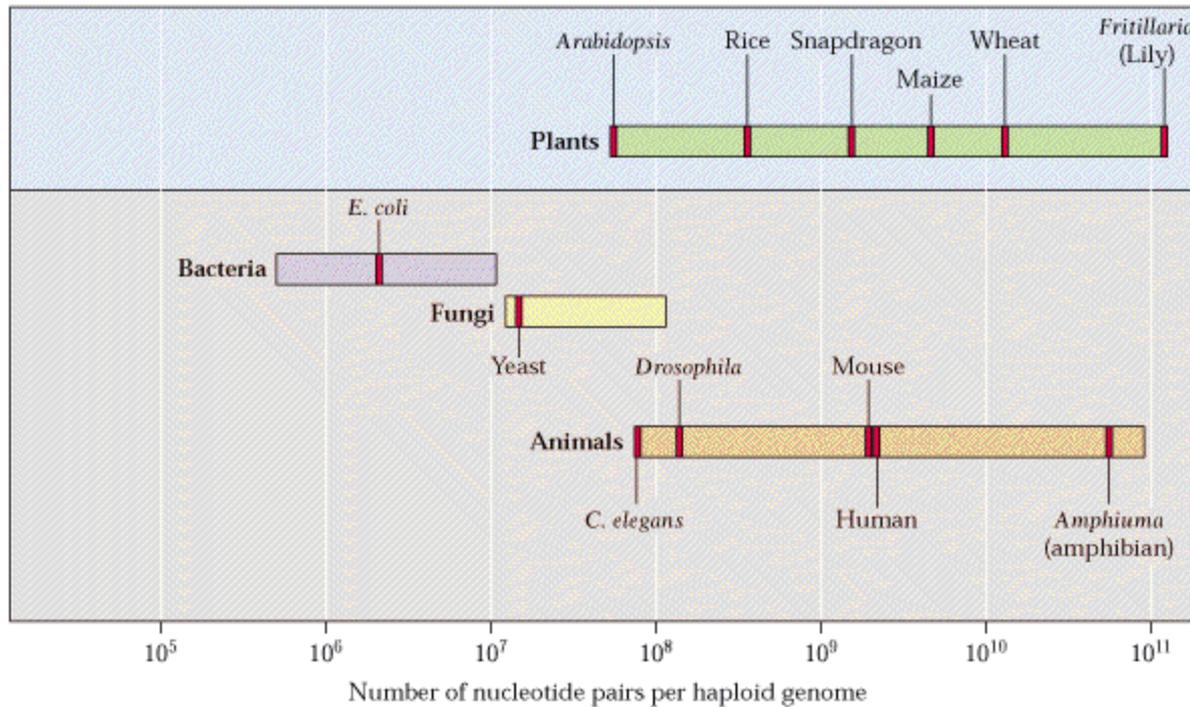
Par :

- la fourniture d'une grande quantité de marqueurs moléculaires
- un meilleur contrôle de la régulation des gènes
- l'identification de gènes candidats pour l'analyse des QTL de caractères majeurs
- l'identification d'allèles favorables

Organismes modèles

Le choix des espèces modèles se fera sur la taille des génomes et sur leur facilité de «culture» en laboratoire.

Ci-dessous, quelques exemples d'organismes modèles



Métagénomique

La **métagénomique** est une méthode d'étude du contenu génétique d'échantillons obtenus à partir de prélèvements réalisés dans des environnements naturels complexes (ex : intestin, océan, sols, air, etc.) par opposition à des échantillons cultivés en laboratoire).

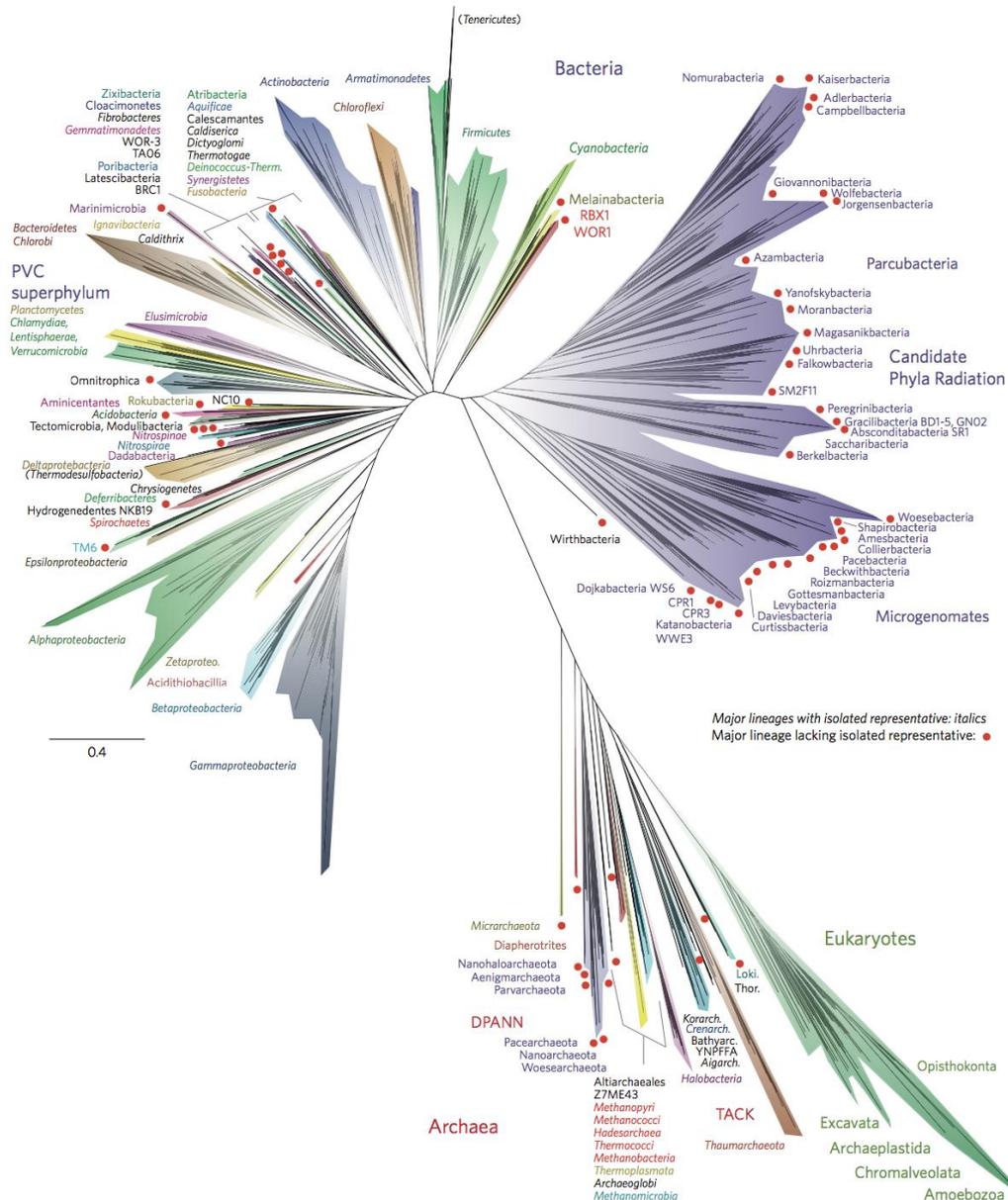
Cette approche, via le séquençage direct de l'ADN présent dans l'échantillon, permet une description génomique du contenu de l'échantillon mais offre aussi un aperçu du potentiel fonctionnel d'un environnement.

Préfixe « méta » → « *ce qui vient après* » : la métagénomique vient après la génomique.



Exemple : études des communautés microbiennes présentes dans ce cours d'eau recevant le drainage acide de mines de charbon en surface.

Métagénomique : nouvel arbre du vivant



Les phyla marqués par des points rouges ont été identifiés par métagénomique et ne possède pas de représentant qui ont été isolé.

La génomique

Le séquençage génomique : il donne accès d'un seul coup à toute l'information génétique d'un organisme. Mais comment la gérer, la déchiffrer et l'utiliser

Nouveaux besoins pour la bioinformatique.
Explosion de données à exploiter

Besoins :

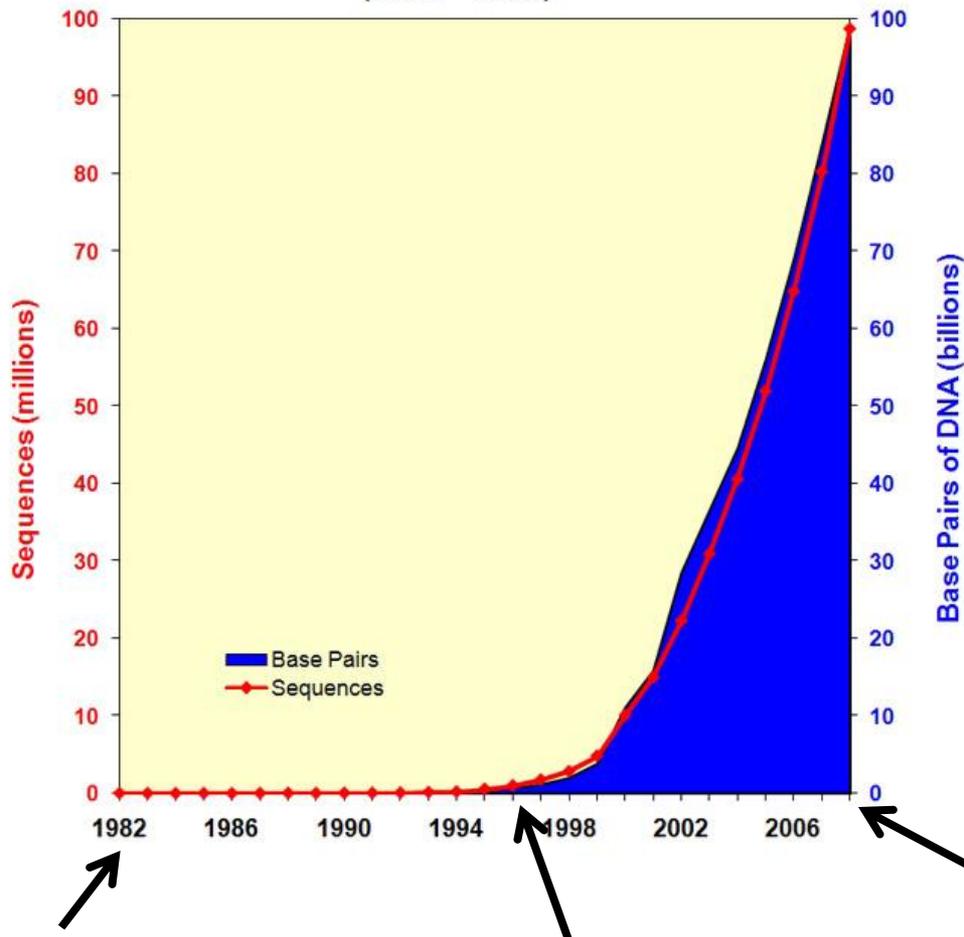
- Gérer les programmes de séquençage
- **Annoter les génomes : où sont les gènes, quelle est leur fonction**
- Enrichir les ontologies (comme la Gene Ontology)
- **Pouvoir comparer les génomes entre eux : la génomique comparative**



La Biologie Moléculaire a besoin de la Bioinformatique pour traiter les données.
La bioinformatique a aussi besoin des données générées par la Biologie Moléculaire.

Séquences disponibles : quelques chiffres

Growth of GenBank
(1982 - 2008)



Août 2023:
246 119 175 seq
2 112 058 517 945 bp

2019:
215 333 020 seq
388 417 258 009 bp

2016:
198 565 475 seq
224 973 060 433 bp

1982: 606 seq
680 338 bp

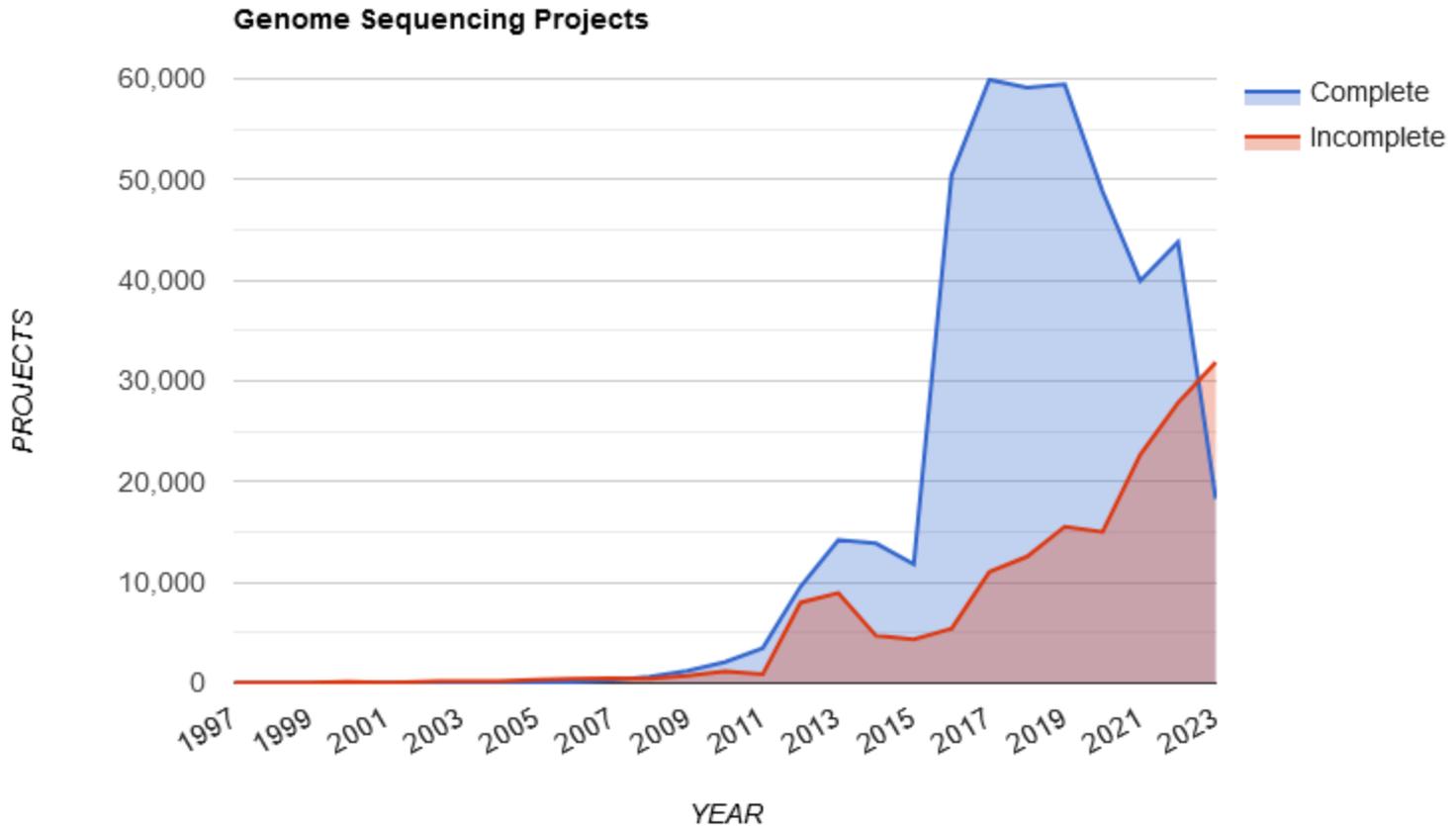
1996: 1 021 211 seq
651 972 984 bp

2008: 98 868 465 seq
99 116 431 942 bp

Génomes : quelques chiffres

GOLD : Genomes Online Database

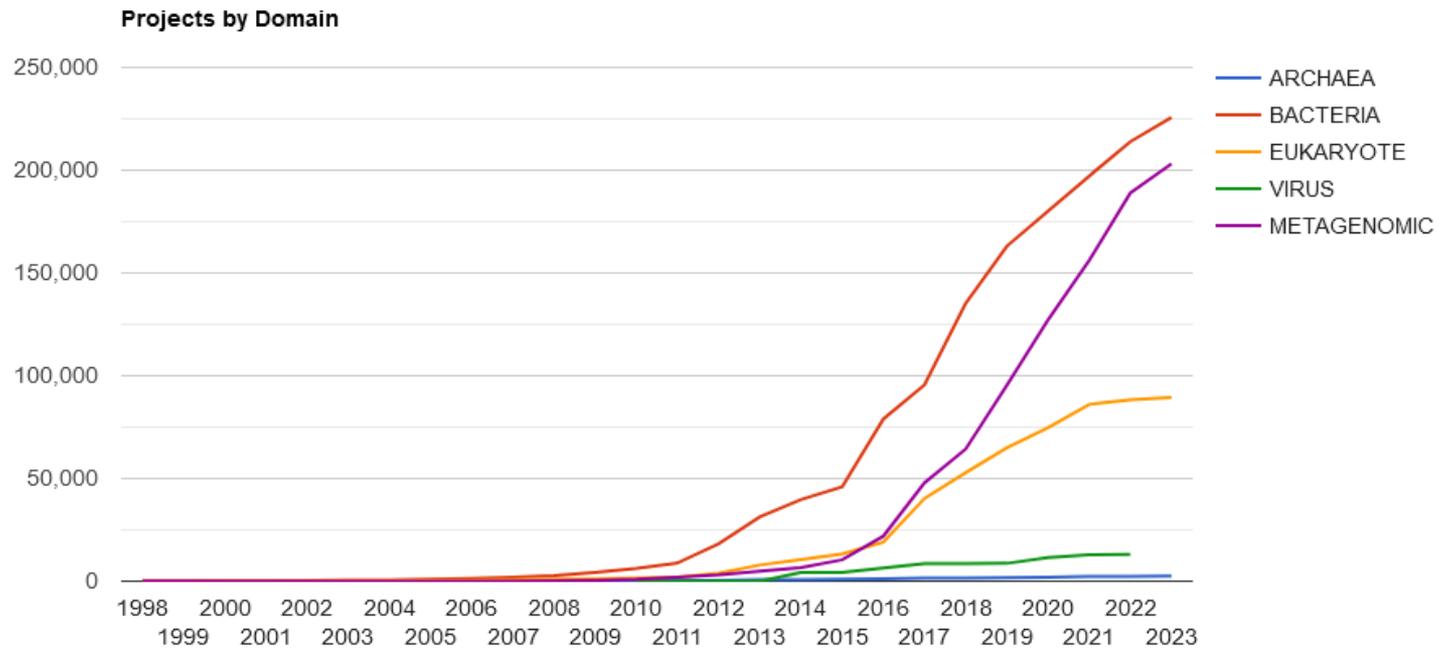
Genome Totals in GOLD (by year and status)



Génomes : quelques chiffres

GOLD : Genomes Online Database

Project Totals in GOLD (by year and Domain Group)



Génomes : quelques chiffres

GOLD : Genomes Online Database

Project Relevance of Bacterial Projects

