

## Introduction

Principe de l'analyse des génomes

### Génomique

Analyse globale du génome d'un organisme

Tous les gènes et régions intergéniques

#### 1. Génomique structurale

- Organisation et position des gènes, taille du génome
- Séquençage de l'ADN et analyse des séquences

#### 2. Génomique fonctionnelle

Analyse du génome pour comprendre la fonction des gènes

- Comparaisons de génomes entre organismes
- Expression du génome
- Variabilité du génome

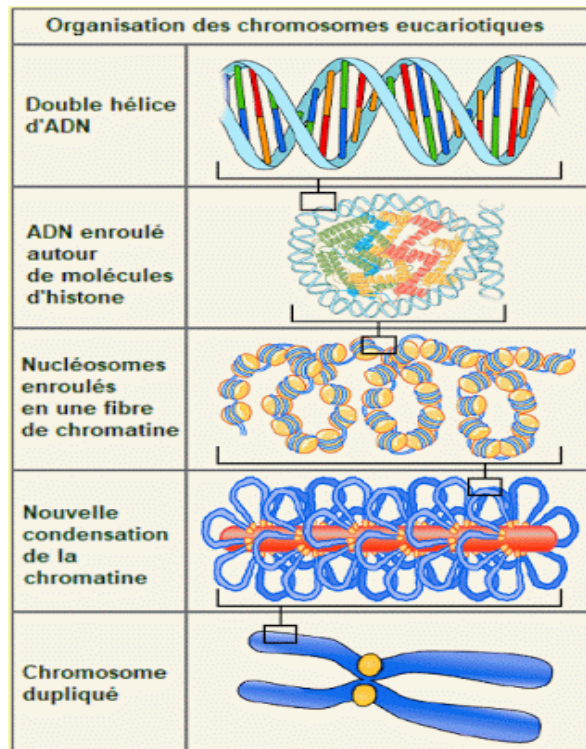
## Le génome

### Définition

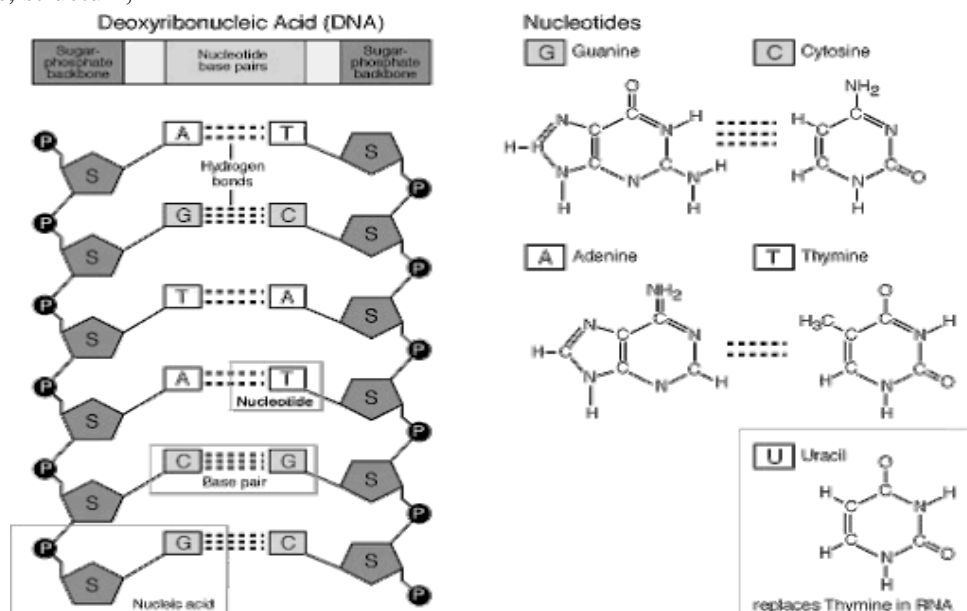
Le génome est l'ensemble des gènes représentant l'information génétique des individus (être vivants). Ces gènes sont répartis au long des brins d'ADN. Au cours des processus physiologiques des cellules, le génome est répliqué, voir transcrit sous forme d'ARN qui est lui-même traduit en protéines (Figure ci-après). Les protéines formant le protéome, interviennent dans la catalyse des réactions ainsi que la régulation du métabolisme cellulaire.

### Rappels et définitions

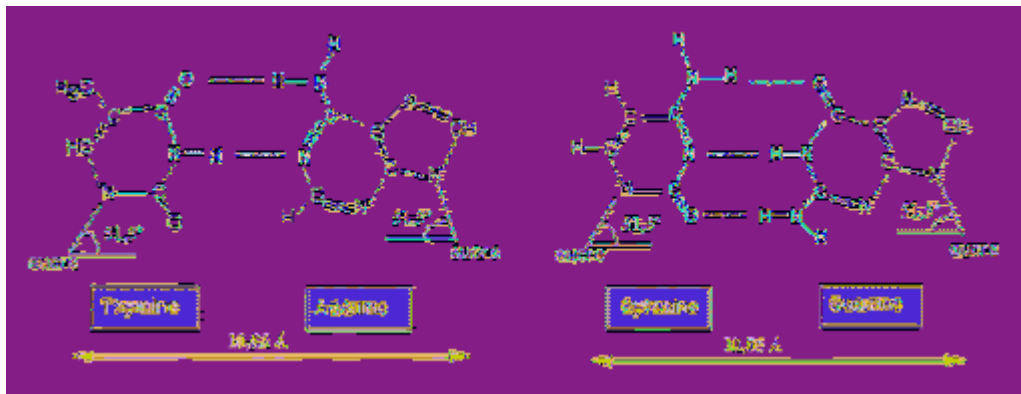
**ADN : Acide désoxyribonucléique :** enroulé en hélices et stabilisé par des protéines basiques appelées Histones. L'ADN est composé de Désoxyribose + base azotée + acide phosphorique. Il est généralement en double brins : un brin codant orienté 5'.....3' et portant l'ensemble des gènes contenus dans le noyau cellulaire chez les eucaryotes et dans le cytoplasme chez les procaryotes ; un brin orienté inversement, 3'.....5' qui s'apparie au premier par l'intermédiaire de liaisons hydrogène.



**ARN : Acide ribonucléique** : est le produit de transcription d'une partie de l'ADN (pour l'ARNm). L'ARN existe sous forme d'un simple brin, il est traduit par les ribosomes, en protéines, produit final intervenant dans des mécanismes biochimiques dans les cellules : défense, transport, catalyse, récepteurs, structure,...etc.



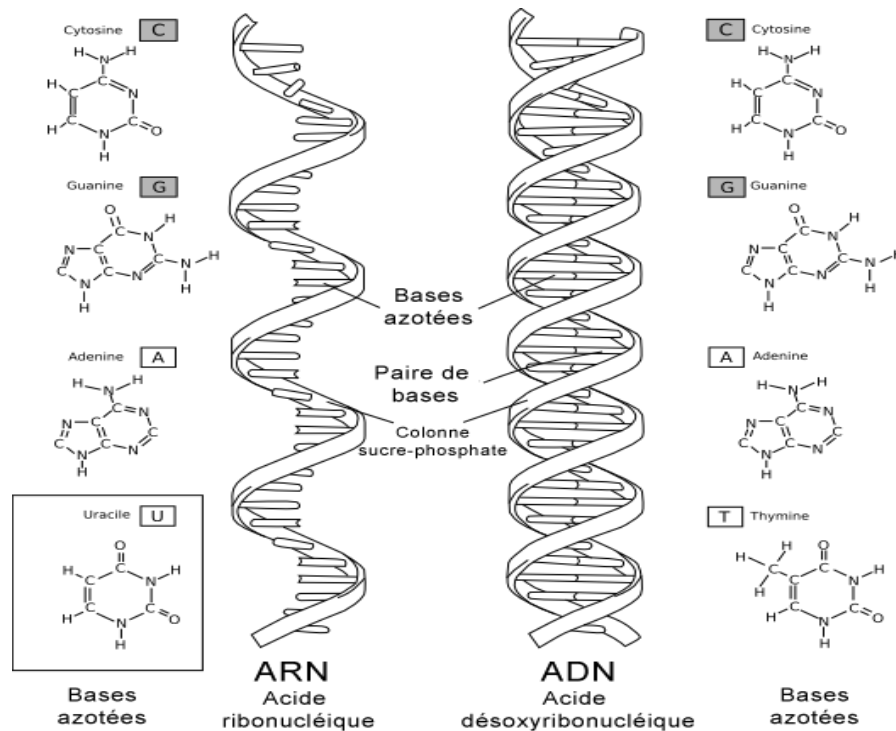
L'image ci-après illustre bien que les nucléotides (Ribose + base azotée + acide phosphorique) sont stabilisés à l'intérieur de la double hélice d'ADN en se positionnant à des distances bien précises :



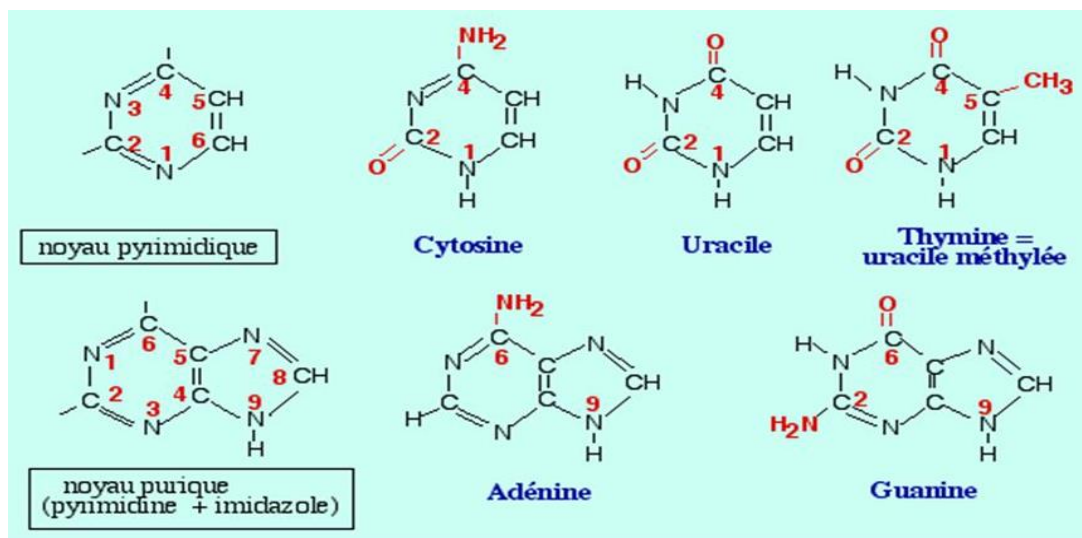
Le tableau ci-après illustre une composition comparative, entre les éléments composant le génome :

| COMPOSITION DES ACIDES NUCLEIQUES |   |   |
|-----------------------------------|---|---|
| Groupes                           | Molécules   |   |
|                                   | ADN   | ARN   |
| ○ Oses ou sucres = S              | ● Désoxyribose                                    | ● Ribose  |
| □ Bases azotées = B               | □ Adénine<br>□ Cytosine<br>□ Guanine<br>□ Thymine | □ Adénine<br>□ Cytosine<br>□ Guanine<br>□ Uracile |
| ◁ Phosphates acides = P           | Présents  | Présents  |

## Structures en hélices



### Les différents types de bases azotées constituant les acides nucléiques

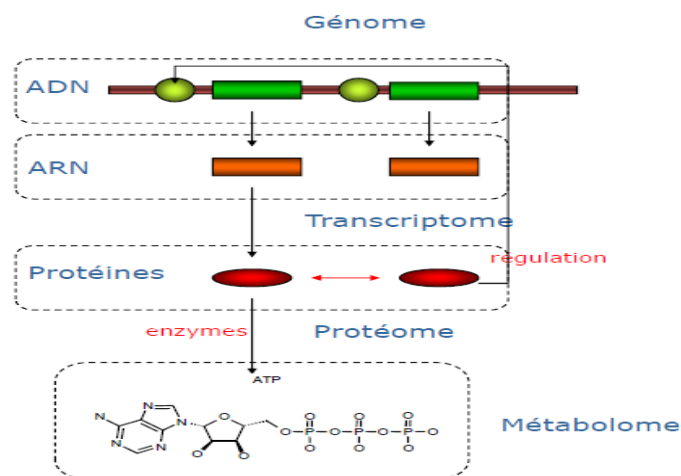
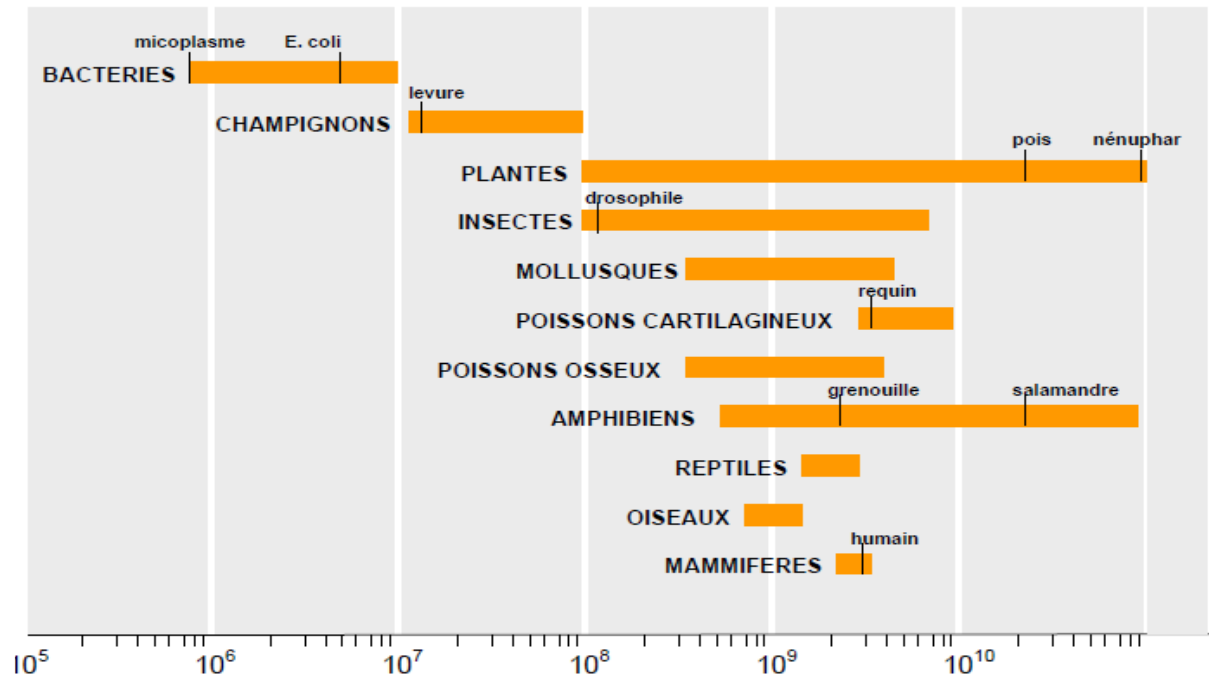


### Notion des Omiques dans le développement de l'analyse génétique et les biotechnologies

- Génomique : analyse des gènes (ADN) : stabilité, variabilité (mutations), fonctionnalité.
- Transcriptomique : analyse des ARN(s) (surtout ARNm) : expression des gènes dans un type cellulaire ou tissu.
- Protéomique : analyse des protéines cellulaires (relation structure/fonction), détection de protéines modifiées ou tronquées (signe de mutations : maladies génétiques), synthèse de protéines recombinantes d'intérêt industriel.

-Métabolomique : analyse des voies métaboliques, catalysées par les enzymes clé du métabolisme cellulaire, permettant de déceler des anomalies liées à la fonction cellulaire.

## I-2-Taille des génomes



### Les outils des analyses génomiques :

-Spectrométrie de masse.

-Séquençage : ADN, ARN.

-Réalisation de banques de données génomiques.

-Electrophorèse : d'ADN (contrôle de synthèse d'un gène, vérification de structure génétique nucléotidiques) ; de protéines (contrôle de fonctionnalité des gènes, localisation de protéines d'intérêt).

-Clonage : synthèse et étude de fonctions génétiques.

-PCR.

-Miniaturisation et puces à ADN.

### **Les applications :**

-Mise au point de nouvelles méthodes de diagnostic des pathologies héréditaires, métaboliques et tumorales.

-Synthèse de biomolécules d'intérêt industriel, thérapeutique, nutritionnel.

-Identification de germes pathogènes par PCR : bactéries, virus, champignons.

-Identification moléculaire des germes d'intérêt biotechnologiques :

Exemple de germes d'intérêt biotechnologiques : les Actinomycètes.

L'identification morphologique et chimiotaxonomique des actinomycètes sont remplacées par l'identification moléculaire, rapide et pointue grâce au séquençage du génome bactérien. De nouvelles espèces ont été découvertes par l'utilisation de l'identification moléculaire, notamment par les techniques du séquençage de l'ADNr 16S et de l'hybridation ADN/ADN (Anderson et Wellington, 2001).

Ainsi les séquences présentant des homologies supérieures à 97% sont considérées de la même espèce. D'autre part, la construction de dendrogrammes est réalisée par les algorithmes des programmes DNAPARS et PHYLIP DNAML.

L'analyse des séquences du gène codant pour l'ARNr 16S a été corrélée aux pourcentages de réassociations ADN/ADN. Car des chercheurs ont démontré que des séquences du gène codant pour l'ARNr 16S présentant des similarités inférieures à 97% ne correspondaient jamais à des pourcentages de réassociation ADN/ADN supérieurs à 70% (Kim *et al.* (2014). De plus des séquences du même gène, partageant moins de 97% de similarité correspondent à des espèces différentes.

Cependant, plusieurs travaux récents ont émis des propositions pour augmenter le seuil de similarité des séquences du gène codant pour l'ARNr 16S : Meier-Kolthoff *et al.* (2013) proposent 98,2% puis Kim *et al.*, (2014) indiquent, à travers leurs travaux publiés dans la revue *International Journal of Systematic and Evolutionary Microbiology* et basés sur une analyse statistique pointue, qu'il est admis que des séquences du gène ADNr 16S sont considérées identiques, si elles présentent un taux de similarité dépassant 98,65%. Ce qui veut dire que les séquences qui partagent moins de 98,65% de similarité du gène codant pour l'ARNr 16S correspondent à des espèces différentes.

<https://www.youtube.com/watch?v=fCd6B5HRaZ8> (MiSeq : Illumina).

<https://www.youtube.com/watch?v=8vDTQh1HAh0> (Séquençage haut débit – applications).

<https://www.youtube.com/watch?v=N5wcTllhjCo> (Illumina : équipements et bioinformatique).

<https://www.youtube.com/watch?v=pDLtfyigPM0> (NGS et traitement du cancer).

<https://www.youtube.com/watch?v=jFCD8Q6qSTM> (1- les 03 techniques NGS).

<https://www.youtube.com/watch?v=-kTcFZxP6kM> (2- Préparation et composition des milieux réactionnels pour les techniques de NGS).

<https://www.youtube.com/watch?v=PGAfWSRYv1g> (3- Réactions et contrôle de qualité des résultats en NGS).

<https://www.youtube.com/watch?v=l4BAfRekohk> (4- Analyse des données en NGS).

<https://www.youtube.com/watch?v=bDsfd6V-U1c> (5- NGS – Applications).