

# Les bases de la Taxonomie Bactériennes

La taxonomie comprend trois domaines différents. D'une part la classification qui consiste à regrouper les organismes au sein d'ensembles en fonction de leurs similitudes: ensembles = « taxons ».

Ensuite la « nomenclature » qui consiste à attribuer un nom à chaque taxon.

Enfin la démarche d'identification qui utilise les 2 domaines précédents afin de reconnaître et donner un nom.

Grâce à la biochimie on a pu classer le monde des microorganismes, le génome permet d'aller encore plus loin.

## **1. Classifier: pourquoi et comment?**

### **• Intérêts et Structure des Classifications**

Il y a trois intérêts majeurs pour classer les microorganismes:

- le classement organise une « banque de données » sur les microorganismes,
- indispensable pour identifier un nouvel isolement,
- donne accès à la phylogénie (cad le liens de parenté entre les différents organismes).

Il existe plusieurs hiérarchie au sein d'une classification, ici classification phylogénétique.

D'abord il y a les « Domaines » (Bactéries, Archéobactéries, Eucaryotes).

(...)

### **• Différents types de classification**

Il existe deux grands types: artificielle ou naturelle.

#### **Classification artificielle:**

Elle est basée sur la prévalence de quelques caractères choisis arbitrairement. Ce type de classification vient du fait que les bactériologistes n'ont pas tous la même approche de la microbiologie (médicale, agroalimentaire, biotech, fondamental, ...). chaque domaine est tenté d'utiliser ses propres critères pour faire de la classification ce qui implique que certaines bactéries peuvent avoir différents noms.

Ex: *Erwinia hubicola* (saprophyte des plantes) = *Enterobacter agglomerans* (bactérie intestinale) ou encore *Bacillus cereus* = *Bacillus thurengiensis*.

#### **Classification naturelle:**

On classe ici les microorganismes avec un maximum de critères sans les hiérarchiser les uns par rapport aux autres.

Il existe deux approches différentes:

- Classification numérique phénotypique,
- Classification phylogénétique.

## **2. Les méthodes de la taxonomie numérique.**

### **• Choix des tests**

Cela entraîne un codage des résultats.

Ce système de codage donne des problèmes d'importance (utilisation de plus ou moins de chiffres). Or l'implication génétique n'est pas fortement liée à l'importance des ces caractères.

Parfois quelques gènes donnent un pigment alors que cela correspond à 4 chiffres différents. Et des centaines de gènes donnent des spores et ce critère ne correspond qu'à un chiffre. Ce n'est donc pas équitable.

### **• Traitements des Résultats**

On calcul des indices de ressemblances entre chaque microorganismes de la base de données. Il y a deux indices différents:

Coefficient de Similitudes:

$$\text{nbr de caractères communs} / \text{nbr de caractères étudiés}$$

Coefficient de Jacquard:

$$\text{nbr de caractères (+) communs} / \text{nbr de caractères étudiés} - \text{caractères (-) communs}$$

Car un caractère (-) n'est jamais certains, alors qu'un caractère (+) est sur. Le caractère peut être (-) car souche dans de mauvaises conditions

On considère que des souches appartiennent à la même espèce à partir de 85% de similitudes (= limite de l'espèce). Pour appartenir au même genre, il faut 65% de similitudes.

### 3. Apports de la biologie moléculaire à la classification phylogénétique.

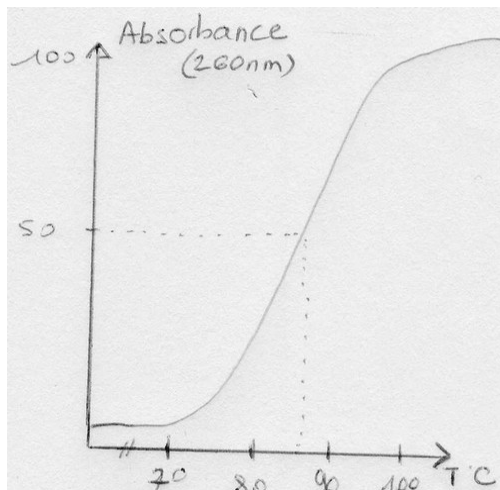
- Étude de la teneur en (G+C) de l'ADN

#### Justification

Si les microorganismes sont proches les uns les autres alors ils ont une teneur en GC proche.

Chez les vertébrés ce n'est pas possible car la teneur en GC varie de 35 à 45%. Or chez les bactéries cela est intéressant car cette variation se fait de l'ordre de 25 à 75%.

#### Technique



$$(G+C)\% = 2,44 T_m - 169$$

$T_m$  = température à laquelle 50% de l'ADN est dénaturé.

On étudie ici l'effet hyperchrome de l'ADN (ce qui donne le  $T_m$  et donc le %GC).

D'autre part on peut aussi faire une centrifugation, sédimentation en gradient de chlorure de Césium.

Plus le %GC est important, plus la migration se fait loin (cad plus ça tombe au fond du tube).

Courbe de dénaturation thermique de l'ADN

#### Résultats

Il faut moins de 3% d'écart pour que deux souches appartiennent à la même espèce. Il faut moins de 10% d'écart pour que deux souches appartiennent au même genre.

Attention : condition nécessaire mais pas suffisante!

Ex: spirilla --> 62% de GC, pseudomonas --> 62% de GC également, mais ne sont pas de la même espèce car physiologiquement, biologiquement et morphologiquement très différentes.

Ainsi on ne compare le %GC que pour des souches ayant le même phénotype. Cela permet de revoir la classification.

- Hybridation entre acides nucléiques

#### Justification

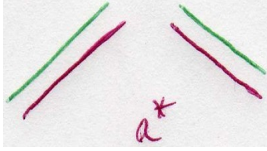
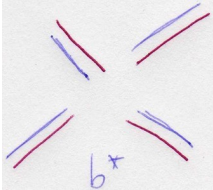
A travers le pourcentage de réassociation entre 2 ADN monocaténaux provenant de 2 souches différentes, on peut voir le pourcentage de ressemblance entre les deux séquences.

On étudie le pourcentage d'homologie (d'hybridation), on étudie aussi le  $\Delta T_m$  = (le  $T_m$  de l'ADN hybride = « hétérologue ») - ( $T_m$  de l'ADN souche référence = « homologue »).

#### Technique

On travaille avec l'ADN des deux souches, l'ADN de référence étant marqué par radioactivité. On mélange les deux types d'ADN en situant le  $T_m$  à  $T_m - 15^\circ\text{C}$  ou  $T_m - 30^\circ\text{C}$  de la souche référence. Le mélange est alors passé sur une colonne de chromatographie pour séparer les ADN monocaténaux des doubles brins; les monocaténaux étant chargés et donc liés avec la colonne.

Sinon on utilise des nucléases qui agissent sur le monocaténaire et qui dégradent ce qui n'a pas été hybridé ce qui permet de garder que les doubles brins.

	
Rouge: Souche de référence radioactive hybridée avec la souche inconnue (verte) non radioactive en excès.	Souche de référence radioactive hybridée avec la même souche de référence non radioactive: 100% d'hybridation théorique
$(a^* / b^*) \times 100 = \text{pourcentage d'hybridation}$	

### **Résultats**

Pour qu'il y ait appartenance à la même espèce il faut plus de 70% d'hybridation et un  $\Delta T_m < 5^\circ\text{C}$ .

### • **Séquençage des acides nucléiques**

On s'intéresse d'abord aux ARN ribosomiques.

#### **Justification:**

L'ARN reflète une partie de l'ADN.

On s'intéresse à l'ARN16s car il n'est pas trop gros (en taille l'ARN23s est trop gros et l'ARN5s est trop petit) et il possède suffisamment d'informations.

L'ARN est une structure universelle qui se retrouve dans tout le monde vivant, on peut donc acquérir une phylogénie de tous les être vivants.

L'ARN a une fonction identique chez tous les êtres vivants, ce qui implique que certaines séquences soient conservées.

Il y a de l'ARN en grande quantité dans les cellules donc il s'agit d'un matériel biologique accessible.

Les gènes qui codent les ARN ribosomiaux ne sont pas soumis à des « transferts horizontaux » (transfert verticaux = transferts de gènes de mères à filles; a contrario transferts horizontaux = transferts de gènes entre même génération, comme les plasmides!).

#### **Technique:**

On a d'abord fait du séquençage partiel après digestion par une RNase T1 qui coupe de façon systématique des 3'G. Ce sont ainsi ces fragments qui ont été séquencés.

**Méthode de SANGER:** « méthode des didéoxynucléotides ». On dispose de brins d'ARN, une ARN polymérase qui permet de faire une copie = matrice dans 4 tubes différents dans lesquels on rajoute des nucléotides dd:

ddATP, ddUTP, ddCTP et ddGTP.

L'incorporation par la matrice d'un nucléotide dd entraîne le STOP de la polymérisation. On peut faire alors une électrophorèse et on lit le résultat de la séquence en partant du nucléotide qui a le plus migré jusqu'au bout du gel.



Séquence lue: UCA GCC UGU AA...

Séquençage totaux de 200 à 300 nucléotides d'un coup d'où l'ARN16s séquencé en 3 ou 4 fois!

#### **Résultats:**

On possède des catalogues d'oligonucléotides dont certains sont spécifiques de certains groupes bactériens= se sont des « séquences signatures ».

C'est notamment à cause de celles-ci qu'on a individualiser le groupe des Archéobactéries (éclatement du groupe des procaryotes) d'après les travaux de Woese.

Cela a amener à l'hypothèse endosymbiotique de la formation de la cellule Eucaryote. Ainsi on pense que l'ancêtre de la sous division alpha des protéobactéries a pu être à l'origine des mitochondries, tandis que les cyanobactéries seraient à l'origine des chloroplastes. Ainsi les mitochondries et les chloroplastes possèdent du matériel génétique.

On a donc pu alors établir une nouvelle classification phylogénétique.

Ici le critère retenu est la ressemblance de l'ARN!

Ainsi certains groupes éclatent avec des liens de parenté entre eux, tandis que d'autres se rapprochent alors qu'ils étaient éloignés dans la classification bactérienne.

**Conclusion:**

***A la base, on note deux approches de la classification qui s'opposent:***

***L'approche Taxonomie Numérique Phénotypique qui s'intéresse aux caractères visibles / étudiables.***

***L'approche de la classification génétique qui s'intéresse à l'ADN et à la phylogénétique entre les êtres vivants.***

**Problème: laquelle faut-il conserver?**

***La classification génétique est intéressante intellectuellement mais en pratique en elle est difficile car pas adaptée. Donc l'approche phénotypique est tout à fait valable car plus abordable. Aujourd'hui on mélange phénotype et génotype pour déterminer les « génomospécies ».***