

Chapitre III : Classification bactérienne

La **systématique** est la branche de la biologie qui a pour but d'étudier la diversité des êtres vivants, de comprendre l'origine de cette diversité et de les **classer** de manière rationnelle au travers de leurs ressemblances et de leurs différences et des relations qui existent entre eux. La **taxonomie** (ou taxinomie), la **nomenclature** et l'**identification** sont les trois disciplines de la **systématique**.

La **taxonomie** (ou **taxinomie**) est la science qui **classe** les êtres vivants (du grec *taxis*, qui veut dire ordre et *nomie* qui signifie lois), de façon hiérarchisée au sein de groupes appelés **taxons**. Un **taxon** (ou **unité taxonomique** ou **groupe taxonomique**) regroupe différents organismes dans l'ensemble qu'il constitue, au sein duquel, ils partagent un grand nombre de caractéristiques communes. Chaque **taxon** reçoit un nom.

La taxinomie microbienne **est en perpétuelle évolution**. Chaque année, de nouvelles espèces sont découvertes. A l'heure actuelle, à peine 10 à 15% des microorganismes existant sont connus. Le reste est à découvrir. Ce constat est dû à l'incapacité d'isoler et de cultiver certains microorganismes.

La **phylogénie** est l'étude de l'histoire évolutive d'un groupe d'organisme à partir d'un **ancêtre commun**. Elle permet de regrouper les organismes selon leurs liens de parenté.

Les **principaux taxons** par ordre décroissant (**hiérarchie taxinomique**) sont :

- **Domaine**
- **Règne**
- **Phylum (Embranchement)**
- **Classe**
- **Ordre**
- **Famille**
- **Genre (ensemble d'espèces)**
- **Espèce (ensemble de souches)**

TAXON	Exemple 1	Exemple 2	Exemple 3
Domaine	<i>Bacteria</i>	<i>Bacteria</i>	<i>Bacteria</i>
Phylum	<i>Proteobacteria</i>	<i>Proteobacteria</i>	<i>Firmicutes</i>
Classe	<i>Gammaproteobacteria</i>	<i>Gammaproteobacteria</i>	<i>Bacilli</i>
Ordre	<i>Enterobacteriales</i>	<i>Vibrionales</i>	<i>Bacillales</i>
Famille	<i>Enterobacteriaceae</i>	<i>Vibrionaceae</i>	<i>Listeriaceae</i>
Genre	<i>Escherichia</i>	<i>Vibrio</i>	<i>Listeria</i>

Espèce	<i>Escherichia coli</i>	<i>Vibrio cholerae</i>	<i>Listeria monocytogenes</i>
Souche Sérovar	<i>Escherichia coli</i> O157:H7	<i>Vibrio cholerae</i> O1	<i>Listeria monocytogenes</i> 4b

Exemple de classification de *E.coli* selon le *BERGEY's Manual of Systemic Bacteriology* (2^{ème} édition, 2001) :

Domaine *Bactéria* (*Eubactéria*)

Règne non défini

Phylum *Proteobacteria*

Classe *Gammaproteobacteria*

Ordre *Enterobacteriales*

Famille *Enterobacteriaceae*

Genre *Escherichia*

Espèce *Escherichia coli*, *E. coli*.

➤ Intérêts de la Classification

Plusieurs intérêts sont majeurs de la classification des microorganismes. En effet, la **classification** organise une "**banque de données**" pour les microorganismes et donne accès à la **phylogénie**. Cette dernière informe sur les liens de parenté entre les différents organismes vivants. De plus, la **taxonomie** est indispensable pour **identifier** un nouvel isolement.

1. Systèmes de classification, Types de classification

1.1 Taxonomie phénotypique ou classification artificielle

Depuis la classification proposée par Cohn en 1872 et jusqu'au début des années soixante (60), toute la taxonomie bactérienne reposait sur une classification **phénétique (ou phénotypique)** qui est basée sur les caractères observables. La définition de chaque espèce en microbiologie n'utilisait donc, que la prise en compte d'un certain nombre de caractères considérés comme importants tels que: la **morphologie**, l'**habitat**, le **pouvoir pathogène**, la **capacité à sporuler**, l'**existence de caractères biochimiques** divers jugés essentiels... De même toutes les tentatives de constructions taxonomiques hiérarchisées (la création des taxons de rang supérieur à l'espèce : les différents genres, familles, ...) étaient établies à l'aide de ce type de critères. C'est pour cela que cette ancienne classification est qualifiée de **classification phénotypique** ou **phénétique**.

En effet, De nombreuses manipulations au laboratoire mettent en évidence les **différences phénotypiques** entre les bactéries: aspect macroscopique des colonies, morphologie et structure de la cellule (forme, Gram, flagelle, capsule, spore...), conditions de culture (type trophique, type respiratoire, température optimale, pH optimal, concentration en dioxygène, exigences nutritionnelles particulières...), caractères biochimiques (ONPG, mannitol, indole, TDA...), Sérotypie (antigènes O et H des entérobactéries, antigènes des streptocoques...), Lysotypie (sensibilité aux phages), Antibiotypie (sensibilité aux antibiotiques). Ainsi, le placement d'une

souche bactérienne dans la famille des *Enterobacteriaceae* nécessite que cette souche soit constituée de bacilles à Gram négatif, non sporulés, AAF, oxydase négative...

1.2 Taxonomie phylogénique, phylogénétique ou classification naturelle

Il s'agit de comparer des **séquences de gènes** et/ou de **protéines** convenablement choisis chez différents organismes, et, par le résultat de la comparaison, d'en déduire l'**arbre phylogénétique** le plus vraisemblable entre les organismes étudiés. Les différences de séquences entre les gènes ou les protéines à comparer pourront aussi se comporter pour l'étude choisie comme "horloges moléculaires" qui marqueront le temps qui s'est écoulé depuis la version ancestrale commune et les différents nœuds de l'arbre évolutif construit.

L'étude des séquences des gènes des ARN ribosomiaux (en particulier le 16S-18S) permet d'étudier l'ensemble des eucaryotes et des bactéries. Parmi toutes les séquences nucléiques possibles, les gènes d'ARNr sont devenus des **marqueurs phylogénétiques** de choix en bactériologie. En pratique, c'est l'ARNr 16S qui s'est imposé pour les études phylogénétiques de bactéries. En effet, de bonne longueur, il réunit des propriétés assez idéales pour les études phylogénétiques. Notant qu'un seul séquençage des ARNr ne permet pas d'établir l'arbre phylogénétique "vrai" de l'organisme vivant. Toute analyse phylogénétique moléculaire repose sur des hypothèses quant à la nature de l'évolution des gènes séquencés et met en œuvre différents modèles mathématiques. C'est l'étude comparée des différentes analyses qui permet de proposer des résultats phylogénétiques vraiment fondés.

Une des difficultés rencontrée en **phylogénie moléculaire** chez les bactéries est celle des transferts horizontaux de gènes entre organismes. Cette appellation qualifie l'échange de matériel génétique entre individus appartenant à des espèces différentes parfois très éloignées. Un exemple très simple et illustratif est celui d'une souche bactérienne devenue résistante à des antibiotiques en acquérant les gènes de résistance à partir d'une souche d'espèce éloignée phylogénétiquement.

1.3 Classification de *BERGEY*

Le *BERGEY'S Manual* est le plus remarquable travail de classification des bactéries. Sa première édition était en 1923 aux USA sous le nom de *BERGEY's Manual of Determinative Bacteriology*. Son objectif initial était le regroupement exhaustif de l'ensemble des informations phénotypiques disponibles pour l'identification des espèces bactériennes reconnues, afin de permettre l'identification de souches bactériennes inconnues. Il y a eu neuf (9) éditions du *BERGEY's Manual of Determinative Bacteriology* jusqu'au 1994. Avec l'apparition dans la 8^{ème} édition (1984) d'une nouvelle classification des bactéries. Les bactéries sont classées en quatre divisions où une est réservée aux archaebactéries, alors que les trois autres sont définies sur la base de la présence ou non d'une paroi et sur sa nature: les bactéries à GRAM négatif, les bactéries à GRAM positif et les Mycoplasmes (bactéries sans paroi). C'est aussi en 1984 qu'est paru le premier volume de la première édition du *BERGEY's Manual of Systematic Bacteriology* (constitué de quatre volumes) qui est la version exhaustive détaillée du *BERGEY's Manual of Determinative Bacteriology*. Toutes les éditions proposent une classification phénotypique des bactéries, ces données étant alors disponibles. Ce n'est qu'en 2001 que le premier volume de la 2^{ème} édition *BERGEY's Manual of Systematic Bacteriology* est apparu. La classification proposée est pour la première fois phylogénétique, basée sur les données accumulées en ce domaine depuis les années 1970, grâce à la généralisation des études génomiques. C'est une classification naturelle totalement différente des précédentes classifications du *BERGEY's Manual*. Elle s'inscrit dans le cadre de la classification des organismes vivants en trois domaines, selon leur organisation cellulaire et leurs parentés évolutives: *Archaea*, *Bacteria*, *Eucarya*.

➤ **Grandes divisions chez les *Bacteria* selon le *BERGEY's Manual***

Le "*BERGEY's Taxonomic outline*" de Mai 2004 distingue 24 phylums. Les voici par ordre alphabétique avec entre parenthèse la numérotation *BERGEY's* du phylum :

- **Acidobacteria (B XIX)**
- **Actinobacteria (B XIV)** Un des deux très gros ensembles de bactéries monodermes (Gram positives). Caractérisé par des bactéries à GC% élevé. Avec des genres incontournables : *Streptomyces*, *Corynebacterium*, *Mycobacterium*, *Micrococcus* ... Actinobacteria (XIV) et Firmicutes (XIII). Voir les Firmicutes ci-dessous.
- **Aquificae (B I)**
- **Bacteroidetes (BXX)** C'est dans ce phylum qu'on trouve le genre classique *Bacteroides* (bacilles, Gram négatif, chimioorganotrophie anaérobie stricte ...).
- **Chlamydiae (B XVI)**
- **Chlorobi (BXI)** Les bactéries phototrophes vertes sulfureuses, avec le genre *Chlorobium*.
- **Chloroflexi (BVI)** Les bactéries photohétérotrophes dites vertes non sulfureuses avec le genre *Chloroflexus*
- **Chrysiogenetes (B V)**
- **Cyanobacteria (B X)** Les bactéries à phototrophie avec photolyse de l'eau.
- **Deferribacteres (B IX)**
- **Deinococcus-Thermus (B IV)** Des bactéries pour environnement extrêmes. A noter la paroi épaisse mais avec membrane externe. On obtient une coloration positive au Gram alors que la structure est en fait plus proche des "Gram négatives". Avec le fameux *Thermus aquaticus* à l'origine de la Taq polymérase.
- **Dictyoglomi (B XXIII)**
- **Fibrobacteres (B XVIII)**
- **Firmicutes (B XIII)** Regroupe la plupart des bactéries monodermes (Gram positives), celles à GC% faible (mais pas les *Actinobacteria* à GC% élevé notamment). Des classiques de laboratoire : les genres *Clostridium*, *Bacillus*, *Staphylococcus*, *Streptococcus*, *Enterococcus*, *Lactobacillus* ... mais aussi les Mollicutes (= les bactéries dépourvues de paroi, elles ont rang de classe, un genre type = *Mycoplasma* ... C'est avec les Firmicutes Gram positives qu'elles ont trouvé leur placement phylogénétique ! comme quoi une paroi, ça peut se perdre).
- **Fusobacteria (B XXI)**
- **Gemmatimonadetes (B XXIV)**
- **Nitrospirae (B VIII)**
- **Planctomycetes (B XV)**
- **Proteobacteria (B XII)** C'est le gros ensemble des bactéries didermes à membrane externe composée de lipopolysaccharides (LPS) (Gram négatif). La plupart sont chimioorganotrophes hétérotrophes, mais les exceptions sont nombreuses... les bactéries pourpres autotrophes photosynthétiques, des chimioolithotrophes comme *thiobacillus ferrooxidans* et les *Nitrosomonas* ...
Pour information, le "Proteo" de Proteobacteria fait référence au dieu grec Protée dont la caractéristique essentielle était de pouvoir apparaître sous des formes très diverses.

- **Spirochaetes (B XVII)** Et le fameux filament axial qui court entre la membrane et la paroi. Où l'on trouvera les célèbres genres *Treponema* et *Treponema pallidum* de la Syphilis.
- **Thermodesulfobacteria (B III)**
- **Thermomicrobia (B VII)**
- **Thermotogae (B II)**
- **Verrucomicrobia (B XXII)**

1.4 Classification mixte et consensuelle

Actuellement, la classification des bactéries se fonde sur la prise en compte d'un maximum de données : données **génétiques** et constructions **phylogénétiques** mais aussi et toujours données **phénotypiques** et données **écologiques** et autres. Ainsi, par le biais des consensus scientifiques entre spécialistes, elle tente de recevoir l'agrément d'un maximum de bactériologistes. Les auteurs la qualifient de "**polyphasictaxonomy**" que l'on peut traduire par **taxonomie polyphasique** ou par **taxonomie mixte et consensuelle**.

2. Critères de classification et d'identification

L'analyse préliminaire dans la classification et l'identification microbienne se fait souvent avec une ou plusieurs **méthodes phénotypiques**. Les méthodes phénotypiques conviennent aux microorganismes qui sont cultivables et dont les paramètres de croissance, ainsi que les profils physiologique et biochimique sont bien établis. L'expression des **phénotypes microbiens** est hautement dépendante des variables environnementales (pH, température, l'appauvrissement des nutriments, la présence de facteurs de stress, etc.). Par ailleurs, les résultats obtenus des méthodes phénotypiques requièrent l'appui de données provenant d'autres méthodes afin d'identifier avec précision un microorganisme. Les critères de classification et d'identification bactérienne les plus importants sont :

2.1 Caractères morphologiques

La morphologie de la colonie et des cellules bactériennes permet d'obtenir une identification initiale relativement à une bactérie. Cela s'effectue par des méthodes simples d'isolement et de culture des microorganismes et surtout par l'observation visuelle subséquente au moyen du microscope optique. Les propriétés morphologiques comprennent: la **forme**, la **taille**, les **caractéristiques de la surface** et la **pigmentation**, les **caractéristiques de la paroi cellulaire** (procédé de **coloration de Gram**), les **caractéristiques de sporulation**, les mécanismes de motilité (**mobilité**) et les autres **inclusions cellulaires** et caractéristiques ultrastructurales (**granules de réserves, plasmides...**), **nature des antigènes (Ag)**, **sensibilité aux antibiotiques**, etc.

2.2 Caractères métaboliques

L'évaluation phénotypique comprend l'étude du profil **biochimique** et des propriétés **métaboliques** d'une bactérie par l'application des tests sur ses **conditions de croissance**, sur ses **activités enzymatiques**, sur la **composition de ses cellules en acide gras**, et autres. Les tests biochimiques utilisent des milieux de croissance, des nutriments, des produits chimiques ou des conditions de croissance particuliers afin d'obtenir une réponse biochimique observable et mesurable du microorganisme, permettant ainsi son identification et sa caractérisation. Ces tests comprennent : l'utilisation de sources de carbone et d'azote, différentes conditions de croissance (anaérobie ou aérobie; température optimale et intervalle de température, pH optimal et intervalle de pH), des conditions osmotiques préférentielles, la production de produits de fermentation, la production d'enzymes, la production de composés antimicrobiens, de même que la sensibilité aux

inhibiteurs métaboliques et antibiotiques. Parmi les tests reconnus, on retrouve : les tests au rouge de phénol et carbohydrate, les tests de catalase et d'oxydase, les tests oxydation-fermentation, les tests au rouge de méthyle, les tests de Voges-Proskauer, la réduction des nitrates, l'hydrolyse de l'amidon, l'hydrolyse du tryptophane, la production de sulfure d'hydrogène, l'utilisation du citrate, etc.

Actuellement, plusieurs systèmes commerciaux miniaturisés et automatisés sont actuellement offerts et sont assortis de procédures de contrôle de la qualité bien définies qui permettent l'identification rapide des microorganismes.

2.3 Caractères génomiques

L'identification microbienne par des méthodes **génotypiques** est une alternative ou un complément aux méthodes phénotypiques établies. Cependant, la biologie moléculaire apporte un ensemble de nouveaux outils puissants : des méthodes moléculaires qui ont aidé à détecter les plus petites variations parmi les espèces microbiennes et les souches. Ces méthodes ont beaucoup amélioré la capacité de détecter, d'**identifier** et de **classifier rapidement** les microorganismes. En plus, d'établir la **relation taxonomique parmi les genres et les espèces apparentés**. L'identification, par les méthodes moléculaires, s'appuie sur la **comparaison de séquences d'acides nucléiques (ADN, ARN)** ou de **profils protéiques** d'un microorganisme avec les données documentées d'organismes connus.

3. Unités de classification

3.1 Notion d'espèce bactérienne

Dans la classification, le taxon de base est l'**espèce (species)**. C'est l'unité fondamentale de la classification en microbiologie. L'espèce se définit comme un ensemble d'organismes partageant de nombreux caractères. Le **genre** étant un groupe d'espèces semblables, la **famille** étant un groupe de genres semblables, l'**ordre** est un groupe de familles semblables, jusqu'au **règne** qui est un ensemble d'**embranchements (phylum)** semblables. Notant que l'espèce biologique se définit comme un ensemble d'organismes reconnaissables par leurs caractères, partageant une même niche écologique, et capables de se reproduire sexuellement entre eux en donnant naissance à une progéniture fertile. Mais l'espèce bactérienne ne peut pas être définie de la même façon que l'espèce biologique puisque, chez les bactéries, la reproduction dans leur majorité est asexuée. Cependant, toute espèce convenablement définie en taxonomie phylogénétique devrait respecter les deux conditions suivantes : tous les individus d'une espèce descendent d'un ancêtre particulier. Ainsi, tous les descendants de cet ancêtre particulier sont dans l'espèce.

3.2 Souches bactérienne, Clone bactérien

Les **cultures pures** obtenues à partir d'une **même bactérie mère** peuvent se différencier par certains **caractères métaboliques**. C'est pourquoi dans certains cas, les espèces sont, elles-mêmes **subdivisées en sous espèces** qui désignent des **souches** différenciées par quelques critères stables. La **souche** étant un **clone** qui veut dire la descendance exclusive d'une bactérie mère unique. Une souche pure est constituée à l'aide des descendants (par succession de cultures) d'une colonie préalablement parfaitement isolée (sur un milieu d'isolement). L'espèce est à la base de la classification. Pour des raisons pratiques dans le domaine médical, des subdivisions de l'espèce de l'espèce sont proposées par rapport à différents caractères spécifiques :

- **Biovars (biotypes)** : souches appartenant à la même espèce mais présentant des marqueurs biologiques chimiques différents.
- **Sérovars (sérotypes)** : souches appartenant à la même espèce mais présentant des antigènes différents.
- **Phagovars (lysovars)**: souches appartenant à la même espèce mais présentant des sensibilités différentes aux phages.
- **Pathovars (pathotypes)** : souches appartenant à la même espèce mais exprimant des pouvoirs pathogènes différents.
- **Zymovars (zymotypes)** : souches appartenant à la même espèce mais présentant des capacités enzymatiques différentes.
- **Antibiovars (antibiotypes)** : souches appartenant à la même espèce mais exprimant des sensibilités différentes aux antibiotiques.

La **souche type** est d'une importance fondamentale en systématique bactérienne puisqu'une espèce est constituée par la souche type et par l'ensemble des souches qui sont considérées comme suffisamment proches de la souche type pour être incluses dans cette espèce.

4. Nomenclature

La **nomenclature** est l'ensemble des règles qui permettent de donner un nom stable à chaque taxon avec les niveaux hiérarchiques de la classification en cours. Il existe un **Code International de Nomenclature des bactéries** (CIN). C'est Le **système binomial** du botaniste suédois **Carl Von Linné** qui est utilisé en nomenclature bactérienne. Ce système désigne l'espèce bactérienne par deux noms en latin: en premier le nom du genre dont seule la première lettre est écrite en majuscule, suivi du deuxième qui est l'**épithète** qui caractérise l'espèce. Le tout **souligné ou écrit en caractères italiques** exemple : *Escherichia coli* ou Escherichia coli. Habituellement et après avoir mentionné une première fois le nom scientifique de l'espèce en entier dans le texte, il est abrégé par la suite, exemple: *E.coli*.

Les noms **informels** sont des noms non spécialisés, par exemple : **Colibacille** c'est le nom informel d'*E.coli* O157.

5. Méthodes de Taxonomie phylogénique

La méthode de classification phénotypique a ses limites : la forme peut varier en fonction du milieu de culture et peut être parfois difficile à définir (bacilles coccoïdes et coques cocco-bacillaires...), une activité enzymatique donnée peut ne pas être détectée lors de l'utilisation d'un substrat synthétique, deux bactéries éloignées peuvent présenter des antigènes en commun, etc.

Dans cette classification, les caractères phénotypiques utilisés sont peu nombreux (une centaine au maximum) par rapport au nombre de gènes habituellement présents chez les bactéries (5 000 environ). De plus, ces caractères sont hiérarchisés les uns par rapport aux d'autres. Par exemple, le résultat de la coloration de Gram est considéré comme beaucoup plus significatif que l'utilisation du sorbitol. Une telle conception facilite l'identification bactérienne, mais elle est néanmoins subjective (car elle attribue une valeur différente aux caractères utilisés).

D'une façon générale, les **méthodes génotypiques** pour l'identification et de classification microbienne peuvent se diviser en deux catégories comme suit :

a) Techniques faisant appel au modèle ou à l'empreinte génétique

Dans les techniques faisant appel au **modèle**, une série de fragments provenant de l'ADN chromosomique d'un organisme sont générés au moyen d'une **méthode systémique**. Les fragments générés sont alors séparés en fonction de leurs tailles pour créer un **profil**, ou une **empreinte génétique unique** à cet organisme et aux organismes très apparentés. On peut créer un répertoire ou une **base de données** avec un nombre suffisant de renseignements. Le profil des tests du micro-organisme peut alors être comparé au répertoire ou aux bases de données. Selon la similarité entre les profils de deux organismes, ils seront considérés comme très apparentés ou pas du tout. Les exemples de quelques techniques faisant appel au modèle ou à l'empreinte génétique sont :

- ✓ Hybridation des acides nucléiques (analyse de transfert de Southern ou hybridation en phase solution);
- ✓ Polymorphisme de longueur des fragments amplifiés, ou AFLP;
- ✓ Electrophorèse sur gel en champ pulsé;
- ✓ Amplification aléatoire de l'ADN polymorphe;
- ✓ Réaction en chaîne de la polymérase (PCR) multiplex

b) Techniques faisant appel au séquençage

Dans les techniques faisant appel au **séquençage**, la séquence d'un segment particulier d'ADN est déterminée. Comme pour le génotypage, une base de données de séquences précises d'ADN est générée, ce qui permet alors de **comparer la séquence test obtenue aux séquences de la base de données**. Le **degré de similarité**, ou la concordance, entre les deux séquences indique dans quelle mesure ces deux organismes sont apparentés. Jusqu'à présent, un certain nombre d'algorithmes au moyen d'ordinateur ont été créés afin de pouvoir comparer les nombreuses séquences les unes aux autres et d'en établir un **arbre phylogénétique** à partir des résultats. Voici quelques exemples de techniques faisant appel au séquençage :

- ✓ Analyse séquentielle du gène d'ARNr 16S
- ✓ Typage génomique multi-locus ou MLST
- ✓ Analyse génomique multi-locus ou MLSA.

Les techniques d'empreinte génétique et les méthodes faisant appel au séquençage ont toutes les deux des avantages et des inconvénients. Les méthodes faisant appel au séquençage, tel que l'analyse du gène d'ARNr 16S, se sont avérées efficaces quant à l'établissement des relations phylogénétiques globales parmi les bactéries au niveau du genre, de la famille, de l'ordre et du phylum. Tandis que, les méthodes ayant recours à l'empreinte génétique sont bonnes pour distinguer les relations au niveau de la souche ou de l'espèce, mais sont moins fiables pour établir l'interdépendance au-dessus du niveau de l'espèce ou du genre. D'autres limites liées aux méthodes génotypiques comprennent les difficultés relatives à la différenciation entre les espèces qui partagent des séquences de régions conservées identiques ou similaires, les informations sont limitées sur la qualité des données séquentielles disponibles dans les bases de données publiques. Ainsi, que la complexité de l'ensemble de la nomenclature taxonomique. Les bases de données de séquences génétiques qui sont utilisées couramment sont : Gen Bank, Projet sur la base de données ribosomique ou RDP et autres.

5.1 Taxonomie phylogénétique ou moléculaire

Les progrès de la **biologie moléculaire** ont permis de comparer les bactéries entre elles d'une manière plus rigoureuse. La classification est obtenue par la **comparaison des molécules d'ADN** des bactéries.

5.1.1 Le GC% ou coefficient de Chargaff

En 1940, il a été remarqué que quelque soit l'espèce d'origine, l'ADN contient toujours autant de purine que de pyrimidine soit donc $(A + G) = (C + T)$ ou $(A+G) / (C+T) = 1$.

De plus, il y a autant de thymine que d'adénine $A/T = 1$, autant de guanine que de cytosine $G/C = 1$

Par contre, le **rapport $(A+T)/(C+G)$ varie beaucoup, il est caractéristique de l'espèce.**

Ce coefficient appelé **coefficient de Chargaff**, peut être calculé suite à un séquençage par la formule:

$$\text{GC\%} = \frac{[G] + [C]}{[G] + [C] + [A] + [T]} \times 100$$

Ce rapport est mesurable grâce à une caractéristique particulière de l'ADN bicaténaire. En effet, les deux brins d'ADN peuvent se séparer par rupture des liaisons hydrogènes qui lient les bases azotées, en chauffant la solution d'ADN (ou milieu acide ou basique qui ionise les bases). Ce phénomène est appelé **fusion**. En effet, l'ADN double-brin absorbe faiblement à **260 nm**, car les bases se font face dans la double hélice. La rupture des **liaisons hydrogènes** entre les bases, par action d'un **agent dénaturant** (la **chaleur** par exemple), entraîne la séparation des deux brins (**dénaturation**) et une **forte augmentation de l'absorbance à 260 nm**, appelée **effet hyperchromique**. Lorsque l'agent dénaturant est la **chaleur**, la température permettant d'obtenir une augmentation de l'absorbance de 50 % par rapport à l'absorbance maximale est appelée **température de transition T_m** (pour *temperature melting*) ou **point de fusion**. C'est la **température de la demi-dénaturation d'ADN**. Cette température est **d'autant plus élevée que le nombre de paires GC est élevé**, car il y a **trois liaisons hydrogènes entre G et C**, contre deux entre A et T (les forces de cohésion des deux brins sont donc d'autant plus importantes que les paires GC sont fréquentes).

Exemple : *E. coli* $T_m = 72^\circ\text{C}$ (GC% = 50) • *P. aeruginosa* $T_m = 79^\circ\text{C}$ (GTC% = 66).

Le **GC% varie selon les espèces**, il est donc intéressant du point de vue taxonomique où:

- Deux bactéries appartenant à la même espèce possèdent des GC% identiques (à 2,5 % près).
- Deux bactéries ayant des GC% différents n'appartiennent pas à la même communauté génétique.
- Deux bactéries qui ont un GC% identique ne présentent pas obligatoirement les mêmes séquences nucléotidiques, et peuvent donc être éloignées génétiquement.
- Deux espèces microbiennes ayant les mêmes séquences nucléotidiques ont nécessairement le même GC%.

Le GC% ne peut être qu'un **critère d'exclusion**. Il permet seulement d'affirmer que deux individus sont éloignés du point de vue génétique. Il permet, par exemple, d'affirmer que deux souches n'appartiennent pas au même genre. Ainsi, l'espèce *Proteus morgani*, dont le GC% est égal à 50 %, est devenue *Morganella morgani*, car les autres bactéries du genre *Proteus* ont un GC% compris entre 38 et 42 %. Autre exemple, le GC% de *Staphylococcus aureus* est d'environ 30

%, tandis qu'il est compris entre 62 et 70 % pour le genre *Micrococcus*. Il est donc inconcevable que les staphylocoques appartiennent à la famille *Micrococcaceae*...(Tableau 3).

Tableau 3 : Représentation de la taille du génome de quelques espèces bactérienne

Totalement séquencé.

Espèce	Taille du génome (pb)	Taux de [G+C]	Fraction codante
<i>Bacillus subtilis</i>	4.214.810	43,5%	97%
<i>Borrelia burgdorferi</i>	910.725	29%	93%
<i>Campylobacter jejuni</i>	1.641.481	31%	94%
<i>Chlamydia trachomatis</i>	1.042.519	41%	nd
<i>Escherichia coli</i>	4.639.221	51%	88%
<i>Fusobacterium nucleatum</i>	2.174.500	27%	90%
<i>Haemophilus influenzae</i>	1.830.127	38%	87%
<i>Helicobacter pylori</i>	1.667.867	39%	91%
<i>Lactococcus lactis</i>	2.365.589	35%	87%
<i>Mycobacterium tuberculosis</i>	4.411.529	66%	91%
<i>Mycoplasma pneumoniae</i>	816.394	40%	89%
<i>Neisseria meningitidis</i>	2.272.351	51,5%	83%
<i>Pseudomonas aeruginosa</i>	6.264.403	67%	89%
<i>Rickettsia prowazekii</i>	1.111.523	29%	76%
<i>Shigella flexneri</i>	4.607.203	51%	80%
<i>Streptomyces coelicolor</i>	8.667.507	72%	89%
<i>Streptococcus pneumoniae</i>	2.160.837	40%	nd
<i>Treponema pallidum</i>	1.138.006	53%	93%

<i>Vibrio cholerae</i>	4.033.460	47,5%	88%
<i>Xanthomonas campestris</i>	5.076.187	65%	84%
<i>Yersinia pestis</i>	4.653.728	48%	84%

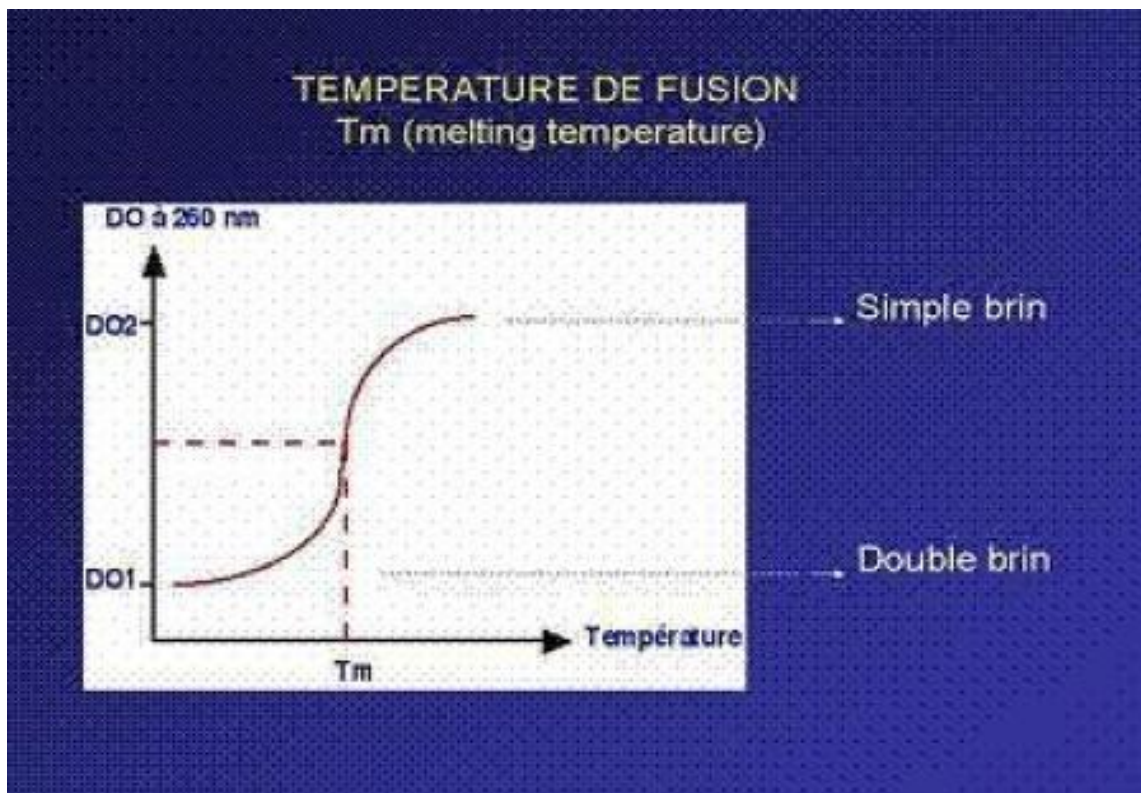


Figure 35 : Courbe de la mesure de la température de la demi-dénaturation d'ADN (T_m).

5.1.2 Hybridation ADN / ADN

Cette méthode permet la comparaison de la totalité du génome de deux bactéries par la mesure du **degré d'homologie** des deux ADN. Les différentes techniques reposent sur le même principe : la **renaturation *in vitro*** de deux brins d'ADN **hétérologues** (chaque brin provenant de deux bactéries comparées) conduit à la formation d'un **hétéroduplex**. Le degré d'homologie est le pourcentage de séquences complémentaires par rapport aux séquences totales. Ces techniques d'hybridation ont bouleversé la classification bactérienne.

- **Technique**

L'ADN chromosomique de chaque souche est extrait, purifié et fragmenté avant d'être **dénaturé** puis **renaturé (hybridé)** en présence d'ADN marqué. La technique la plus utilisée est le **marquage** radioactif : le **marqueur** permet de repérer les hybrides formés entre l'ADN de la souche **cible (C)** et celui de la souche de **référence (R)** :

➤ **Homologie ADN/ADN à T_{or}**

L'homologie ADN/ADN est le rapport (en pourcentage) de bases appariées sur les bases totales. Deux souches n'appartiennent à la même espèce que si le taux d'**hybridation (renaturation)** entre les deux ADN est supérieur ou égal à 70 %. Il varie de 0 à 60% pour des souches d'espèces différentes.

La renaturation (hybridation) de l'ADN est maximale à une température définie appelée: **température optimale de renaturation (T_{or})** qui est inférieure de 25° à 30°C à la température de dénaturation. La renaturation ou l'hybridation, est dans son maximum à T_{or} .

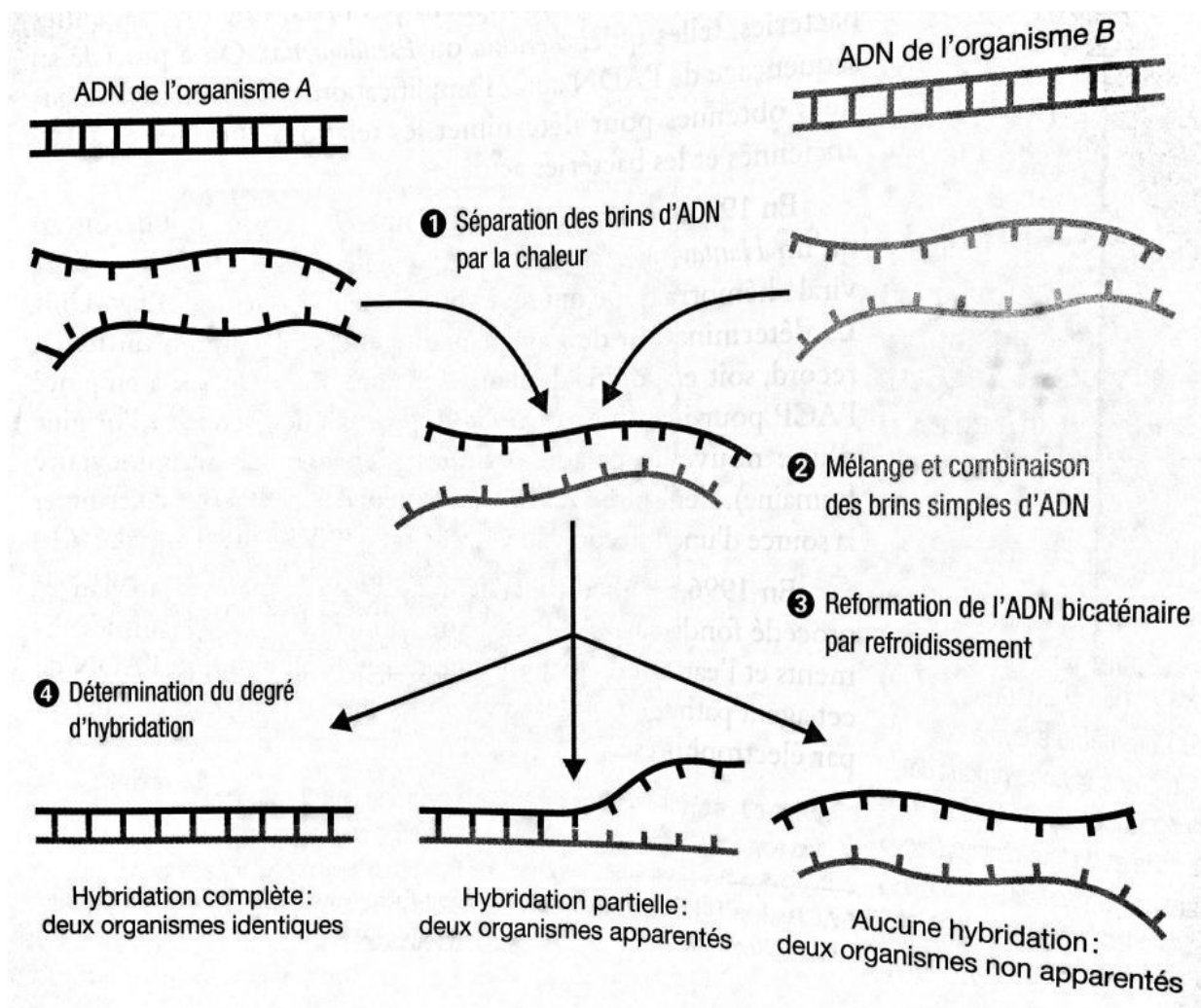


Figure 36 : Représentation schématique de l'hybridation de l'ADN.

➤ **Stabilité thermique des hybrides (T_m (e))**

Elle est donnée par le paramètre $T_m(e)$ (thermal elution mid point) qui est la température de dénaturation de 50 % de l'hybride. Ce paramètre est obtenu par l'élévation progressive de la température de dénaturation dans la solution d'ADN hybridé, au-dessus de la température optimum de renaturation (hybridation). La comparaison des $T_m(e)$, respectivement dans des réactions d'hybridations homologues ($A \times A'$) prises comme témoins et des réactions d'hybridations hétérologues ($A \times B$) permet d'obtenir une estimation précise de la stabilité thermique des hybrides hétérologues. Cette dernière est directement corrélée à la stabilité de la structure chimique des brins d'ADN hybridés, résultant elle-même de la complémentarité de leurs séquences de bases.

Les souches appartenant à une même espèce bactérienne ont des $\Delta T_m(e)$ comprises entre 1 et 5 °C, soit aux environ de 5% de bases non appariées. Alors que des souches d'espèces différentes présentent des ΔT_m (e) entre 8 à 20 °C.

Plusieurs méthodes sont utilisées pour l'analyse par hybridation ADN/ADN :

- ✓ Hybridation sur filtre de nitrocellulose (utilisation de radioactivité)
- ✓ Technique de l'hydroxyapatite (utilisation de radioactivité)
- ✓ Technique de l'endonucléases (utilisation de radioactivité)
- ✓ Technique optique (pas de radioactivité)

5.1.3 Etude des ARN ribosomaux (ARNr)

Les ARNr sont considérés comme des **chronomètres moléculaires**. Ils sont présents dans les cellules procaryotes et eucaryotes et ils ont une structure bien conservée chez tous les êtres vivants. Des portions d'ARNr ont une séquence identique chez tous les êtres vivants. Ils sont abondants dans la cellule, faciles à purifier et à séquencer (utilisation d'une reverse transcriptase).

Les ARNr s'associent à des protéines pour former les ribosomes. La sous-unité 30S d'un ribosome contient de l'ARNr 16S et la sous-unité 50S contient de l'ARNr 5S et de l'ARNr 23S. L'ARNr 5S est formé d'environ 120 nucléotides, l'ARNr 16S de 1 500 nucléotides et l'ARNr 23S comprend environ 2900 nucléotides. Le plus utilisé pour les études taxonomiques est l'**ARNr 16S**. Les travaux sur ces ARNr 16S ont permis de distinguer les Eubactéries (Bactéries vraies) des Archéobactéries (ou Archéobacteries) (Figure 37).

Arbre phylogénétique de la vie

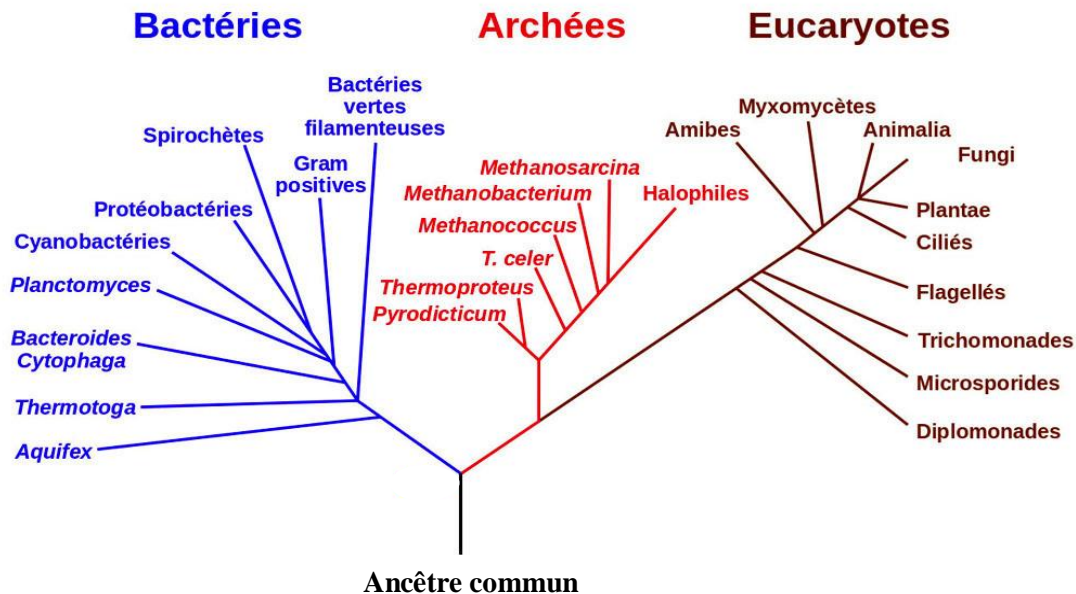


Figure 37 : Schéma de l'arbre phylogénétique universel des organismes vivants, établis par l'analyse de L'ARNr16S et 18S.

6. Taxonomie numérique

En 1763, le botaniste français ADANSON proposait une méthode de classification qui tenait compte de l'ensemble des caractères d'un organisme, chaque caractère ayant la même valeur. En 1957, SNEATH applique une méthode similaire aux bactéries et développe une taxonomie qualifiée de **numérique** ou d'**adansonienne**. De manière schématique, la méthode consiste à étudier, pour chaque souche, plus d'une centaine (100) de caractères morphologiques, biochimiques, culturels, structuraux... et à attribuer le même poids à chacun des caractères qui sont codés **1 (présence du caractère)** ou **0 (absence du caractère)**. Le but recherché est de rassembler dans une **classe de similitude** les individus les plus semblables. L'analyse mathématique de quantités importantes de données concernant les caractères phénotypiques et moléculaires permet le **calcul des distances** existant entre différents groupes taxonomiques. Les résultats obtenus peuvent être représentés graphiquement sous forme de **dendrogramme** (Figure 38).

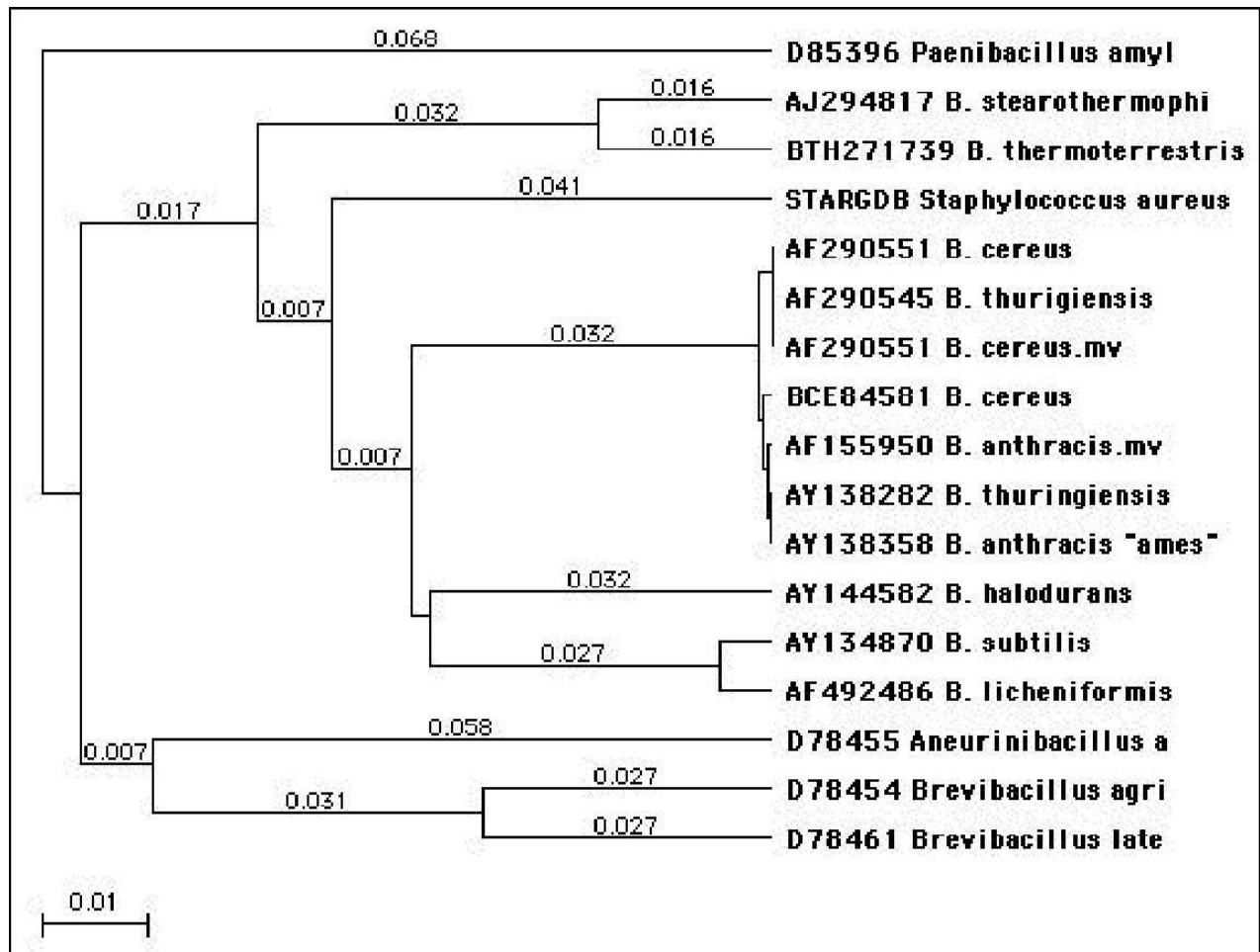


Figure 38 : Dendrogramme de quelques espèces du genre *Bacillus* et de genres apparentés.