

CHAPITRE 3 : PRINCIPAUX TYPES DE LA TAXONOMIE

Deux types de classifications sont généralement utilisés en taxonomie bactérienne : classification artificielle et naturelle.

1. Classifications artificielles

Est une classification phénétique (ou **phénotypique**), donc basée sur la considération des caractères observables. Elle réunit des groupes de bactéries sur la base des propriétés phénétiques communes et aisément reconnaissables par sa présence (+) ou son absence (-) telle que : la forme cellulaire, la réponse à la coloration de **GRAM**, le type respiratoire, la température ou le pH de croissance, la présence d'une enzyme ou d'un pigment particuliers,...

La caractérisation d'une bactérie est donc basée sur la mise en évidence d'un certain nombre de caractères phénétiques considérés comme significatifs (un choix arbitraire des caractères).

Ce type de classification est par nature subjectif car le choix des caractères considéré comme significatifs peut varier d'un auteur à l'autre et leur nombre, forcément limité, ne peut refléter qu'une information réduite sur la souche considérée. Mais elles sont d'un grand intérêt pratique puisqu'elle limite le champ d'investigation aux seules bactéries partageant la propriété spécifique, commune et discriminante. Mais d'un autre côté, elles peuvent réunir des bactéries par ailleurs très hétérogènes et même de phylogénies différentes comme (les bactéries à **GRAM** négatif et les *Mycoplasmes*).

Aussi, 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*.

2. Classifications naturelles

Appelées aussi classifications phylogénétiques (du grec *Phylon* : race et *Genesis* : génération). *Elles se basent sur l'existence d'espèces clairement identifiées et génétiquement homogènes qui, à partir d'un ancêtre commun, ont évolué différemment.* Elles sont basées sur les relations évolutives établies entre les différents taxons.

Les caractères spécifiques des espèces sont aisément discernables chez les animaux et végétaux. La rareté ou même l'inexistence chez les bactéries des éléments de classification utilisés chez les animaux et végétaux *rend inapplicable leur classification naturelle sur les mêmes principes*. Toutefois, il existe maintenant de nombreuses techniques d'étude de la composition chimique des bactéries et surtout de leurs structures moléculaires génétiques qui permettent d'établir entre elles des relations phylogénétiques scientifiquement objectifs et fiables.

Le modèle de la classification phylogénétique proposé pour les bactéries a pour principe de base l'étude comparative des marqueurs moléculaires spécifiquement retenus en raison de leur grande stabilité évolutive ou de leur faible variabilité. Ces marqueurs sont soit directement liés au génome comme l'ADN ou l'ARN et les protéines qui en dérivent, soit des molécules structurales stables, telle que les composants membranaires (acide gras, lipoquinones) ou pariétaux (peptidoglycane). Chacun de ces marqueurs moléculaires détermine un niveau de signification spécifique.

La classification de référence actuelle des bactéries est la classification de **BERGEY's Manual**. C'est une classification naturelle et nouvelle, proposée à partir de l'année 2001, elle est basée sur les données phylogénétiques accumulées depuis les années 1970 sur les bactéries et les *Archaeobactéries* par l'analyse comparative de leur ARN16S.

3. Différentes approches taxonomiques

La taxonomie bactérienne moderne vise l'intégration de toutes les données et les informations phénotypiques, génotypiques et phylogénétiques menant à une taxonomie polyphasique et une classification plus stable.

3.1. Taxonomie phénotypique

L'identification de l'espèce repose sur la comparaison de divers caractères phénotypiques de la souche à étudier vis à vis de ceux d'une souche de référence. La classification phénétique ou phénotypique utilise un faible nombre de caractères considérés comme importants tels que la morphologie, la mise en évidence d'un caractère biochimique, l'habitat. Mais elle ne reflète qu'un nombre réduit d'information, *les caractères considérés comme importants sont subjectifs* et dépendent des conditions environnementales.

3.1-1-Caractéristiques morphologiques

Les traits morphologiques sont importants en taxonomie microbienne pour de nombreuses raisons. La morphologie est facile à étudier et à analyser, en particulier pour les microorganismes eucaryotes et les bactéries ou archées les plus complexes. De plus, les comparaisons morphologiques sont significatives parce que les caractères structuraux dépendent de l'expression de nombreux gènes, qu'ils sont habituellement génétiquement stables et que, normalement (au moins chez les eucaryotes), ils varient peu avec les changements environnementaux. Ainsi, une similarité morphologique est une bonne indication d'une parenté phylogénétique.

Les caractères morphologiques sont utiles pour orienter le processus d'identification des bactéries. On peut ainsi distinguer les bactéries d'après leur forme, leur groupement et leur taille ; de plus, les différences entre des structures telles que les endospores, les capsules ou des flagelles fournissent des indications précieuses. Beaucoup de caractères morphologiques différents sont utilisés dans la classification et l'identification des microorganismes (**Tab. 01**).

Tableau 01: Quelques caractères morphologiques utilisés pour la classification et l'identification

Caractères	Groupes microbiens
Formes cellulaire	Tous les groupes principaux
Taille cellulaire	Tous les groupes principaux
Morphologie des colonies	Tous les groupes principaux
Caractéristiques ultrastructurales	Tous les groupes principaux
Réaction à la coloration	Bactéries, certains champignons
Cils et flagelles	Tous les groupes principaux
Mécanismes et mobilité	Bactéries mobiles par glissement, spirochètes, protistes
Formes et localisation de l'endospores	Bactéries formatrices d'endospores
Morphologie et localisation des spores	Bactéries, champignons, protistes
Inclusions cellulaire	Tous les groupes principaux
couleur	Tous les groupes principaux

3.1-2-Caractères tinctoriaux : Coloration différentielle

En mettant en évidence les similitudes ou les différences dans la composition de la paroi cellulaire des microorganismes étudiés fournissent des informations à propos des relations phylogénétiques. La coloration différentielle est l'une des premières étapes de processus d'identification d'une bactérie. La majorité des bactéries sont soit Gram positif, soit Gram négatif.

D'autres colorations différentielles, telles que la coloration acido-alcoolrésistante, peuvent servir à identifier des microorganismes appartenant à des groupes plus restreints. Ces colorations

sont élaborées en fonction de la composition chimique de la paroi cellulaire et qu'elles ne permettent pas les bactéries sans paroi ou les archaebactéries dont la paroi présente des caractéristiques inhabituelles.

3.1-3-Caractéristiques physiologiques et métaboliques

Les caractéristiques physiologiques et métaboliques sont très utiles, car elles sont directement en relation avec la nature et l'activité des enzymes microbiennes et des protéines de transport, citons : les sources de carbone et d'azote, les constituants de la paroi cellulaire, les produits de fermentation, le mode générale de nutrition, le domaine de la température optimale de croissance, la luminescence, le mécanismes de conversion de l'énergie, la mobilité, la tolérance osmotique, les pigments photosynthétiques, le besoins et la tolérance en sel, les métabolites secondaires formés, la sensibilité aux inhibiteurs métaboliques et aux antibiotiques et l'inclusions des réserves.

3.1-4-Caractéristiques biochimiques

L'activité enzymatique sert souvent à différencier des bactéries. Il est généralement possible de distinguer des bactéries étroitement apparentées et de les regrouper en des espèces distinctes au moyen d'épreuves biochimiques, comme celle qu'on utilise pour déterminer leur capacité à provoquer la fermentation d'une gamme donnée de glucides.

3.1-5-Caractéristiques écologiques

La capacité d'un microorganisme de coloniser un environnement spécifique a une valeur taxinomique. Certains microorganismes peuvent être très similaires sous beaucoup d'autres aspects, mais habiter des niches écologiques différentes, ce qui suggère qu'ils peuvent n'être pas aussi proches parents que supposé à première vue. Comme exemple de propriétés écologiques taxinomiquement importantes, citons les modèles de cycles biologiques, la nature des relations symbiotiques, la capacité de causer une maladie chez un hôte particulier, les préférences d'habitat telles les exigences de températures, de pH, d'oxygène et de concentration osmotique.

3.1-6-Caractères sérologiques

La sérologie est la science qui étudie le sérum sanguin et les réponses immunitaires mises en évidence par l'examen du sérum. On trouve sur le marché des trousse commerciales contenant des solutions d'anticorps destinées à l'identification de divers microorganismes important d'un point de vue médical. Ce type de solution s'appelle antisérum ou immunosérum. Si on isole une bactérie inconnue d'un patient, on peut souvent l'identifier rapidement à l'aide d'antisérum connus.

Le test d'agglutination sur lame avec l'antisérum est une procédure qui consiste à incorporer des échantillons d'une bactérie inconnue dans des gouttes de solution saline (eau physiologique) placées sur différentes lames. On ajoute ensuite un antisérum différent à chaque échantillon. Les bactéries s'agglutinent (ou forment des grumeaux) lorsqu'elles sont mélangées aux anticorps spécifiquement produits en réaction à cette espèce ou souche de bactérie; l'agglutination indique que l'épreuve est positive. Les épreuves sérologiques permettent de différencier non seulement des espèces de microorganismes, mais aussi des souches d'une même espèce.

3.1-7- Caractères physico-chimiques ou Chimiotaxonomiques

L'application des techniques chimiques et physiques dans l'étude de la composition chimique du milieu interne de la cellule bactérienne a contribué à fournir des informations de grande utilité aussi bien dans l'identification que dans la classification des bactéries. *Le terme chimiotaxonomie décrivant la classification des bactéries sur la base de leur constitution chimique a été lentement discuté.*

3.1-7-1-Composition de la paroi cellulaire

Différencier qualitativement les bactéries sur la base de leur paroi cellulaire demande l'analyse microscopique à transmission et l'analyse de la composition de différents polymères : peptidoglycanes, polysaccharides, lipopolysaccharides, acides téichoïques.

3.1-7-2-Composition lipidique

Déjà à partir de la structure de base des lipides, on peut distinguer les Archaeobactéries de tous les autres organismes ; en effet leurs lipides contiennent des chaînes latérales isoprénoides avec des liaisons éther (isoprenyl-glycerol-ether) à la différence des autres organismes (eucaryotes) à chaînes hydrocarbures avec liaison type ester. Les lipides eubactériens comprennent de nombreuses et différentes classes dont quelques-unes ont un potentiel chimiotaxonomique à l'exemple des glycolipides moins répartis que les phospholipides.

Parmi les caractéristiques employées en taxonomie microbienne, il faut citer l'analyse des acides gras par une technique appelée l'analyse FAME (pour Fatty Acids Methyl-Ester). Un profil d'acides gras d'une souche peut révéler des différences spécifiques dans la longueur des chaînes, le degré de saturation, les branchements dans les chaînes et les groupes hydroxyles. Les microbes de la même espèce auront des profils d'acides gras identiques pour autant qu'ils aient été cultivés dans les mêmes conditions.

3.1-7-3-Quinones isoprénoidiques

Les quinones isoprénoides sont une classe de terpènes localisés dans la membrane cytoplasmique de nombreuses bactéries. Elles jouent un rôle important dans le transport des électrons, la phosphorylation oxydative et éventuellement dans le transport actif. Les benzoquinones (Ubiquinone et Plastoquinone) et les naphthoquinones (Menaquinones, Dimethylmenaquinones, phyloquinones) forment deux groupes structuraux.

3.1-7-4-Acides aminés et protéines

La comparaison des séquences d'acides aminés des différents types de protéines entre différentes souches est utilisée comme référence pour les mesures phylogénétiques entre microorganismes. Les séquences en acides aminés des protéines reflètent directement les séquences des ARNm et par conséquent, les gènes qui codent pour leur synthèse.

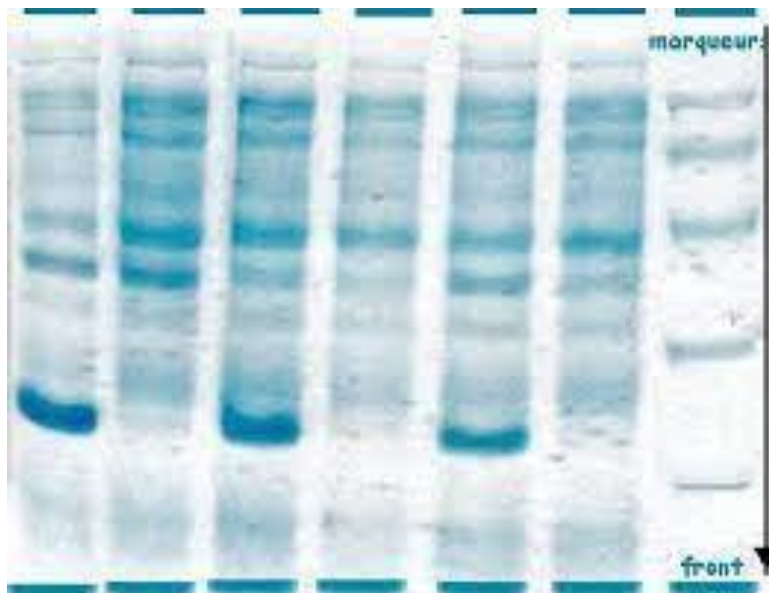


Figure 01: Profils protéiques issus par une électrophorèse SDS-PAGE

(<http://biochimej.univangers.fr/Page2/TexteTD/8TPmethodologie/5Electrophorese/1Electrophorese.htm>)

Il y a plusieurs façons de comparer les protéines. Les méthodes indirectes de comparaison des protéines sont d'usage courant, en particulier la mobilité électrophorétique qui est utilisée pour étudier les relations au niveau de l'espèce et de la sous espèce. Les profils issus par l'électrophorèse SDS-PAGE ont indiqué que les différentes souches possèdent différents profils mais les souches qui appartiennent à la même espèce ont des profils semblables.

3.1-8-Taxonomie numérique

La *taxinomie numérique* ou *taxométrie* ou *taxonomie adansonienne* (du nom d'ADANSON son auteur) est une approche quantitative rendue possible grâce à la puissance des ordinateurs car elle implique un volume de calcul considérable. Elle est basée sur la comparaison de caractères de différentes natures : morphologiques, physiologique, génétiques appartenant à des souches prises deux à deux.

Les caractères retenus sont considérés d'égale valeur et sont quantifiés numériquement, de manière à établir des distances taxonomiques (*indice de distance, d*) qui traduisent à la fois la similarité (ressemblance) et les rapports d'ascendance évolutive (parenté, filiation) entre les organismes confrontés. La quantification binaire (0 ou 1, c'est-à-dire absence ou présence) des similitudes et des différences permet alors de caractériser les taxons par un *coefficient de similitude (s)* ($d = 1-s$), calculé de diverses manières, selon le choix des caractères sélectionnés et le codage et le traitement appliqués aux données recueillies. Parmi les indices les plus couramment utilisés, il y a l'indice de Jaccard Sneath donné par la relation suivante :

$$S_{AB} = nS^+ / nS^+ + nd$$

S_{AB} : coefficient de similitude entre la souche A et la souche B

nS^+ : nombre de caractère similaires

nd : nombre de caractères différents pour donner une valeur significative à l'étude.

La taxonomie numérique implique le choix de caractères taxonomiques indépendants les uns des autres et non hiérarchisés, ainsi que la standardisation rigoureuse de la méthode. Le nombre de caractère retenus doit être significatif. En pratique, ce nombre se situe entre 30 et 300 caractères qui peuvent être de diverses natures : paramètres génomiques, caractères phénétiques classiques (morphologie, physiologie, structure, métabolisme), composants chimiques (peptidoglycane, LPS...).

Les résultats de l'analyse taxonomique numérique sont souvent exprimés sous la forme d'un diagramme ramifié analogue à un arbre. C'est un **dendrogramme** où les organismes ayant la plus grande similitude sont groupés en ensembles appelés **phénons**. La signification taxonomique de ces phénons est cependant controversée, surtout au niveau du pourcentage de similitude à partir duquel un phénon peut être considéré comme espèce, un genre... mais en général, on estime qu'au-delà de 80% de similitude des phénons peuvent être assimilés à une même espèce.

Autrement dit :

On calcule cet indice pour chacun des couples de souches à classer. (si on veut classer, on doit calculer les indices de distances $S_1S_2, S_1S_3 \dots S_1S_{(n-1)}, S_1S_n ; S_2S_3 \dots S_2S_{n-1}, S_2S_n ; S_3S_{n-1}, S_3S_n ; S_{n-1} S_n$).

On dresse alors un tableau en carré avec en abscisses et en ordonnées les souches S_1 à S_n et dans chaque case les indices correspondants. Si l'indice est égal à 0, les deux souches sont identiques, s'il est égal à 1, les deux souches sont totalement dissemblables pour les tests effectués. Dans la diagonale, les indices sont nuls puisqu'on y compare chaque à elle-même. Les souches ayant les plus faibles indices sont les moins dissemblables et peuvent former un groupe G_1 dont on calculera, pour chacune des souches non incluses dans le groupe, un nouvel indice qui sera la moyenne des indices des constituants de G_1 . On forme ensuite un deuxième groupe G_2 constitué des souches ou groupe les plus semblables. On calcule les indices de ce nouveau groupe et de proche en proche on pourra constituer un groupe unique qui contiendra toutes les souches soumises au classement. On peut ensuite dessiner un graphe appelé dendrogramme qui représente les relations de ressemblance entre les sujets et les groupes.

Cette méthode, qui nécessite de nombreux calculs, a grandement bénéficié des outils informatiques qui en rendu l'utilisation plus accessible.

3.2-Taxonomie moléculaire

Il est difficile de surestimer combien l'étude de l'ADN, de l'ARN et des protéines a fait progresser notre compréhension de l'évolution et de la taxinomie microbiennes.

3.2-1-Composition en bases des acides nucléiques

Les génomes microbiens peuvent être directement comparés et leur similarité taxinomique estimée de nombreuses façons. La première et probablement la plus simple des techniques consiste à déterminer la composition en bases de l'ADN. Dans l'ADN double brins, A s'apparie avec T et G avec C. Donc le rapport $(G+C)/(A+T)$, ou teneur en G+C le pourcentage de G+C dans l'ADN, est un reflet de la séquence en bases et varie avec les modifications de séquences comme suit :

$$G+C\% = [(G+C)/G+C+A+T] \times 100$$

Le contenu en G+C est souvent déterminé à partir de la température de fusion (température de demi-dénaturation) (T_m : *Melting Temperature* ») de l'ADN.

Le contenu en G+C de nombreux microorganismes a été déterminé. L'ADN des microorganismes varie fortement dans son contenu en G+C. Celui des bactéries et des archées s'étale d'environ 25 à presque 80% et est plus variable que celui des protistes et des fungi. En dépit d'une telle échelle de variation, le contenu en G+C des souches à l'intérieur d'une espèce particulière est constant et varie très peu à l'intérieur d'un genre. Si deux organismes diffèrent de plus de 10% dans leur contenu en G+C, leurs génomes ont des séquences en bases très différentes.

3.2-2-Hybridation des acides nucléiques

La similarité entre génomes peut se comparer plus directement par l'hybridation des acides nucléiques, appelée aussi hybridation ADN-ADN. Deux souches dont les ADN accusent une parenté d'au moins 70 % dans les conditions optimales d'hybridation et moins de 5% de différence de T_m , sont souvent, mais pas toujours, considérées comme des membres de la même espèce. Si des molécules d'ADN ont des séquences très différentes, elles ne formeront pas d'hybrides stables détectables. Par conséquent, l'hybridation ADN-ADN n'est utilisée que pour l'étude de microorganismes étroitement apparentés. On peut comparer des organismes plus distants, en réalisant des hybridations ADN-ARN avec des ARN ribosomiques ou de transfert, radioactif.

3.2-3-Séquençage des acides nucléiques

Les ARNr de la petite sous unité sont un outil presque idéal pour l'étude de l'évolution et de parentés microbiennes, car ils jouent le même rôle chez tous les microorganismes.

Les ARNr ont été choisis en taxonomie pour plusieurs raisons évidentes :

- Molécule ubiquiste ;
- Structure bien conservée car toute modification pourrait nuire à la synthèse protéique ;
- Séquences d'ARNr identiques chez tous les êtres vivants ;
- Abondants dans la cellule et donc facilement purifiables

La stabilité des ARNr est mise à profit pour analyser les relations des bactéries au niveau de l'espèce et à des niveaux hiérarchiques plus élevés. L'ARNr 16S est le plus utilisé. L'ARNr 16S est utile à la classification phylogénétique et à l'identification bactérienne puisqu'il est présent dans toutes les bactéries. Il comporte des séquences conservées (stables) communes à des unités de taxons élevés et des séquences variables spécifiques d'espèces. La séquence nucléotidique de l'ARNr 16S peut-être comparé via internet à celles de souches déposées dans des banques de données internationales.

La tendance actuelle est de travailler sur le gène correspondant. La séquence du gène codant l'ARNr 16S est connu pour environ 4000 souches et est accessible par interrogation de bases de données (EMBL, GenBank). Les programmes FASTA et BLAST permettent de comparer une séquence nucléotidique d'une souche inconnue avec les banques de séquence et retiennent les séquences les plus proches.

L'amplification par PCR présente un intérêt pour le diagnostic de bactéries non cultivées.

Il est admis qu'en dessous de 97% d'homologie deux bactéries ne peuvent appartenir à la même espèce. Ainsi, il n'est donc pas utile de faire des hybridations ADN/ADN en dessous de ce seuil. Si le pourcentage d'homologie est > 97%, le placement de 2 souches dans une même espèce ou pas repose sur les résultats de l'hybridation ADN/ADN.

Limites : deux espèces peuvent avoir des séquences ARNr 16S très proches et être cependant très différentes par hybridation ADN/ADN.

Ex : *Aeromonas trota* et *A. caviae* (99.9% de similitude pour ARNr 16S et 30% de similitude pour l'hybridation ADN/ADN).

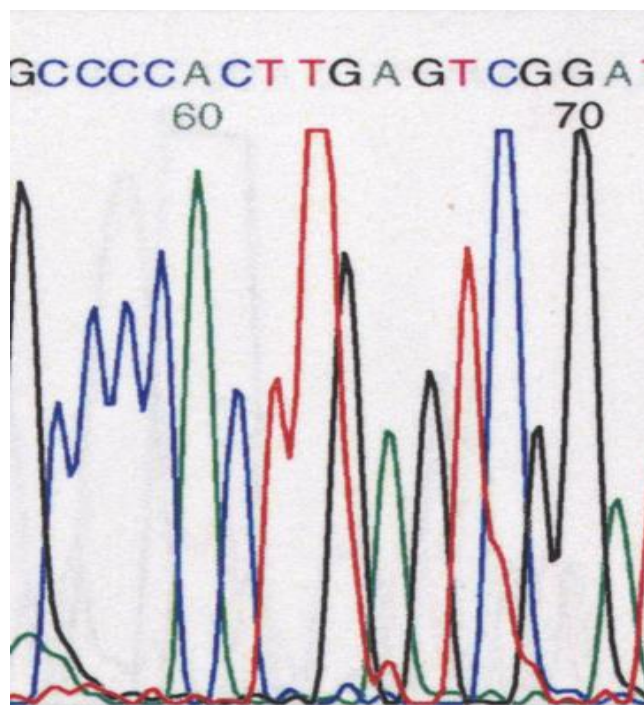


Figure 02: Exemple de résultat du séquençage

(http://mpronovost.profweb.ca/BIONP1/bionp1_biotechnologie.html)

3.2-4-Prise d'empreinte génomiques

Pour classer les microorganismes au niveau de l'espèce et aider à déterminer les relations phylogénétiques, on peut aussi utiliser un ensemble de techniques, appelé prise d'empreintes génomiques. «fingerprinting». Souvent, on séquence et on compare de cinq à sept gènes domestiques, via une technique appelée analyse de séquence multilocus (MLSA).

Une autre forme de prise d'empreinte génomiques appelée l'analyse du polymorphisme de longueur de fragments de restriction (RFLP) est basée sur le pouvoir qu'ont les endonucléases de restriction de reconnaître des séquences de nucléotides spécifiques. L'analyse RFLP détecte des changements dans la taille des fragments de restriction (polymorphisme) comme moyen de détecter des différences dans les séquences d'ADN des souches microbiennes.

Un autre test est basé sur les séquences d'ADN hautement conservées et répétitives présentes chez la plupart des bactéries Gram négatives et chez quelques Gram positives. Il y a trois familles de séquences répétitives : les éléments BOX de 145pb, la séquence ERIC (enterobacterial repetitive intergenic consensus) de 124-127pb et REP (repetitive extragenic palindromic) de 35-40pb. Une autre technique alternative, le polymorphisme au niveau du nucléotide ou SNP, « single nucleotide polymorphism » permet d'échantillonner une fraction nettement plus grande du génome que ne le permettent l'analyse de l'ARNr 16S ou la MLSA. L'analyse SNP se concentre sur des changements mononucléotidiques, ou polymorphismes à ce niveau, dans des gènes spécifiques, des régions intergéniques ou dans d'autres régions non codantes. Ces régions attirent l'attention parce qu'elles sont normalement conservées et que par conséquent, le changement d'une paire de bases peut avoir une signification évolutive.