

PCR et séquençage du gène ARN 16S : applications en Bactériologie

Depuis 1985, l'utilisation des techniques de biologie moléculaire dans le diagnostic des maladies infectieuses est croissante. Elles permettent d'identifier une bactérie dans un produit pathologique ou une culture. Le principe sur lequel elles reposent consiste à amplifier un gène entier ou non, avec des amorces spécifiques, qui peut être ultérieurement révélé par électrophorèse sur gel ou capillaire, par hybridation, ou encore séquencé et comparé avec des séquences déposées dans des banques de données (par exemple, EMBL, NCBI, BiBi, Genbank).

Ainsi, il existe deux principales approches de la *Polymerase Chain Reaction* (PCR) :

- une approche ciblée sur un microorganisme, utilisant des amorces spécifiques ("on sait ce que l'on cherche") ;
- une approche large spectre, en ciblant un gène commun à toutes les bactéries, le gène ARNr 16S ("on ne sait pas ce que l'on cherche"), puis séquençage et comparaison aux banques de données pour identifier la bactérie en cause.

Les avantages des techniques de biologie moléculaire par rapport aux techniques bactériologiques classiques sont un gain de sensibilité (x 2 par exemple pour la recherche des *Chlamydia* génitaux), un gain de temps important (divisé par 2 pour l'identification des mycobactéries à partir de la culture) et/ou un gain de spécificité, pour l'identification des germes inhabituels à partir des cultures.

Le gène 16 S : l'intérêt

Ce gène code la sous-unité 16S de l'ARN ribosomal (ARNr) et est essentiellement utilisé en raison de sa structure, très conservée dans toutes les bactéries. En effet, il est constitué d'une succession de domaines conservés, sites de complémentarité pour les amorces universelles utilisées pour le séquençage de ce gène, et d'autres portions de séquences propres à un groupe de bactéries, nommées séquences signatures (espèce, genre, famille).

Le choix de l'ARN 16S plutôt que 23S ou 5S est d'ordre technique (taille du gène, nombre d'informations) et, surtout, les banques

de données de séquences du gène 16S sont aujourd'hui très développées. Enfin, l'identification est fiable : les résultats obtenus par séquençage du gène 16S sont similaires à ceux obtenus avec le génome entier.

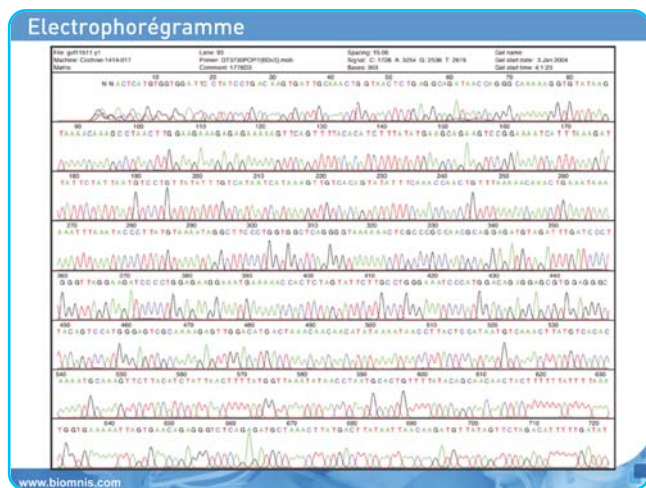
Le gène 16S : les limites

Il n'est pas possible de faire une quantification bactérienne, car les gènes codant l'ARNr peuvent être en plusieurs copies, le nombre de copies variant en fonction de la taille du génome et de la vitesse de croissance de la bactérie. Pour quantifier, d'autres gènes sont utilisés, présents en même nombre de copies chez toutes les bactéries (une copie généralement). D'autres gènes sont utilisés pour différencier des espèces bactériennes proches, pour lesquelles l'ARN16S n'est pas discriminant, par exemple MLSA pour *Neisseria meningitidis*.

Principe de la technique PCR-séquençage

Une amorce universelle est choisie, courte chaîne nucléotidique de 15 à 25 nucléotides, complémentaire de la séquence d'ADN connue à amplifier et située juste en amont de la zone à séquencer. Cette amorce est nécessaire à l'accrochage de l'ADN polymérase. Le séquençage de l'ADN consiste à déterminer la succession de nucléotides composant l'ADN. La première phase est la synthèse d'ADN complémentaire au brin matrice par l'ADN polymérase (à partir de l'amorce), par l'ajout de désoxyribonucléotides : dNTP (dATP, dCTP, dGTP, dTTP). Ce qui est particulier, pour le séquençage, est l'utilisation de nucléotides légèrement différents, les didésoxyribonucléotides ddNTP (ddATP, ddCTP, ddGTP, ddTTP). Dans le milieu de réaction, les dNTP sont en grande quantité et les ddNTP en faible nombre. Lorsque l'ADN polymérase utilise un ddNTP au lieu d'un dNTP, la synthèse du brin s'arrête. L'incorporation aléatoire de ddNTP permet d'obtenir des fragments d'ADN de taille variable. Les ddNTP sont incorporés, marqués par 4 fluorochromes différents. Les fragments d'ADN obtenus sont de longueurs différentes ; le séquenceur sépare les fragments d'ADN selon leur taille par chromatographie. Les systèmes actuels détectent la fluorescence des 4 nucléotides à partir de la même colonne. Le résultat est présenté sous forme d'électrophorégramme. Ensuite, est

effectué l'alignement, soit graphiquement, soit par texte, c'est-à-dire que les séquences écrites en ligne sont envoyées sur Internet et comparées à des séquences déposées dans des banques de données. On peut ainsi obtenir des BLAST, c'est-à-dire des comparaisons avec les séquences les plus proches et l'identification qui correspond, rendue avec un pourcentage d'homologie.



Intérêt en microbiologie clinique : l'identification bactérienne

La PCR séquençage de l'ARN 16S est intéressante pour l'identification précise des souches peu décrites, rarement isolées, ou phénotypiquement difficiles à identifier (peu de caractères biochimiques connus, ou profil biochimique aberrant), pour l'identification de bacilles à Gram négatif peu décrits car atypiques dans des prélèvements cliniques, et pour celle des bactéries corynéformes et des bacilles à Gram positif, pour lesquelles les galeries d'identification sont peu fiables ou ne donnent pas de résultats satisfaisants. Dans ces situations, l'identification par séquençage du gène 16S est plus rapide et sans ambiguïté.

Un autre intérêt important est l'identification de bactéries de croissance lente telles les mycobactéries. Les résultats sont rendus plus rapidement après la culture, évitant les tests biochimiques souvent longs et peu performants.

Cette technique a aussi permis la description de nouveaux agents pathogènes et a contribué à la (re)classification des bactéries.

Enfin, la technique est utile pour l'identification de bactéries non cultivables ou difficiles à cultiver, telles que *Bartonella quintana* ou *Coxiella burnetii*, avec un impact clinique important dans les endocardites à hémocultures négatives.

Identification bactérienne par séquençage du gène 16S : les limites

Le séquençage de l'ARN 16S a un faible pouvoir résolutif à l'espèce pour certains genres bactériens : ainsi, ne peuvent être distingués *Neisseria cinerea* et *N. meningitidis*, ou *Streptococcus pneumoniae*, *S. mitis* et *S. oralis* ou encore le groupe des

Campylobacter non jejuni non coli (même ARN 16S).

Certaines espèces présentent > 99,5 % de similarité entre elles, mais des différences phénotypiques claires (ex. : 3 espèces d'*Edwardsiella*). Enfin, des séquences ambiguës ont été déposées dans certaines banques de données publiques par le passé.

Recommandations pour l'utilisation du séquençage de l'ARN 16S pour l'identification bactérienne

[Drancourt et al. 2000 J. Clin. Microbiol]

Il convient de séquencer au moins 500 à 525 bp, idéalement 1300 à 1500 bp. Les résultats sont exprimés en % d'homologie : les critères pour l'identification à l'espèce sont plus de 99 % d'homologie de séquence, idéalement > 99,5 %, et une comparaison à une séquence de souche de référence identifiée par hybridation DNA-DNA (technique de référence).

Intérêts en microbiologie clinique pour la détection bactérienne directement dans les prélèvements

- En cas d'antibiothérapie préalable inhibant la culture. Ainsi, ont été décrits, dans des endocardites, *Streptococcus salivarius* et *Capnocytophaga canimorsus* par PCR-séquençage 16S (sur valve de patient déjà traités et dans des hémocultures) ;
- En présence de biofilm, dans les ostéites et ostéomyélites, le biofilm inhibant la culture ;
- Dans les endocardites à culture négative, en cas d'antibiothérapie préalable, de microorganismes fastidieux, ou non cultivables en routine (*C. burnetii*, *B. quintana*). La PCR 16S est alors directement réalisée sur l'hémoculture ;
- Détection de *Mycobacterium* sp. dans des prélèvements respiratoires et dans le LCR ;
- Détection directement dans des biopsies de *Tropheryma whipplei*, *Bartonella henselae* ;
- Détection directe dans des prélèvements ostéo-articulaires ;
- Détection des bactéries dans l'environnement.

Détection bactérienne (non ciblée) directe : les limites

La PCR 16S est possible sur des prélèvements de sites normalement stériles (biopsies, hémocultures, ponctions) ou paucimicrobiens. Mais l'une des principales limites est le risque d'amplification de contaminants : réactifs contaminés (Taq polymérase), détection d'agents microbiens commensaux, contamination lors des manipulations (extraction ADN, ou PCR). En outre, le délai de résultat est plus long qu'avec les PCR spécifiques d'espèce et la technique PCR 16S est plus lourde : 1,5 jours de travail en moyenne pour obtenir une séquence à partir d'une colonie ou directement de l'échantillon.

Carole Emile, d'après une communication de Véronique Jacomo, Biomnis Paris.

