



# LE SÉQUENÇAGE

## LE SÉQUENÇAGE HAUT DÉBIT POUR L'ÉTUDE DE LA BIODIVERSITÉ MICROBIENNE

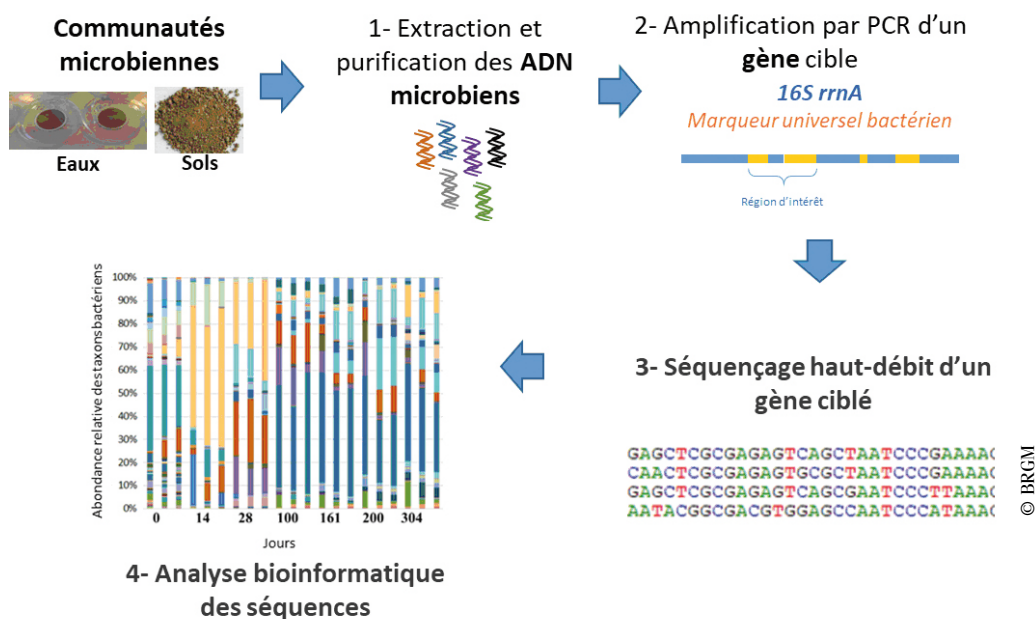
### DESCRIPTION THÉORIQUE DU PRINCIPE DE LA TECHNIQUE

Le séquençage haut-débit est une technique de biologie moléculaire récente qui permet de séquencer un très grand nombre de brins d'ADN. Appliqué par exemple au gène *rrs* codant l'ARNr 16S, marqueur universel des bactéries, cette approche permet l'étude de la biodiversité bactérienne et archéenne. En effet, l'analyse des séquences renseigne sur l'identité et l'abondance relative des microorganismes présents dans un milieu. Les connaissances sur le rôle de certains taxons dans les cycles biogéochimiques connus donnent un aperçu du potentiel métabolique présent dans un milieu. L'impact d'une pollution sur la microflore du site

peut être évaluée ainsi que la présence de taxons impactant les polluants.

En dehors du séquençage de gènes universels tels que le gène *rrs* codant l'ARNr 16S, la méthode peut également être utilisée sur d'autres marqueurs, notamment sur des gènes de dégradation des polluants, pour caractériser la diversité d'une population ciblée.

A l'inverse de la qPCR, le séquençage ciblé fournit une identification des microorganismes présents qui permet de remonter au potentiel fonctionnel, c'est une vision globale.



Principe du séquençage.

### CONTEXTE D'UTILISATION

L'analyse de la biodiversité microbienne par séquençage haut-débit peut s'appliquer sur la plupart des milieux (sol, eau, air, déchets) ainsi que sur la plupart des sites. Elle complète les outils physico-chimiques classiques en identifiant les microorganismes et leur abondance. Elle permet d'étudier les modifications des communautés microbiennes suite à une pollution, une évolution des conditions hydrogéochimiques ou encore en cas de traitement *in situ*. Elle fournit des résultats qualitatifs (diversité bactérienne) ainsi que semi-quantitatifs (abondance relative des microorganismes).

L'analyse de la biodiversité microbienne remplace progressivement les approches dites d'empreinte moléculaire qui fournissent une image de la diversité sans identification des microorganismes présents mais qui restent encore très utilisées pour des suivis de modifications de la diversité.

D'autres approches d'analyse de la diversité existent comme l'approche dite non ciblée ou « shotgun » qui permet, par assemblage des séquences obtenues, de reconstruire des génomes microbiens et des voies métaboliques, cependant ces approches ne sont pas encore accessibles pour un diagnostic de routine. Enfin une troisième génération de séquençage (MSRT) est actuellement au stade de recherche et développement et devrait permettre une miniaturisation des outils et une nouvelle réduction des coûts et du temps d'analyse.

#### À quelle étape ?

Lors de l'évaluation de l'état d'un sol ou du diagnostic d'un site potentiellement pollué, le séquençage haut-débit permet de confirmer la présence et l'abondance relative

# LE SÉQUENÇAGE

de microorganismes pouvant contribuer à une atténuation naturelle des polluants et donc d'aider dans le choix de scénarios de gestion.

Lors de la phase d'évaluation de la faisabilité et de la performance du biotraitement, le séquençage haut-débit permet de caractériser la présence de communautés microbiennes impactant (par dégradation ou transformation) les polluants. Il est ainsi possible de déterminer si un

traitement par biodégradation est envisageable mais aussi d'anticiper l'efficacité et d'identifier les causes d'un mauvais fonctionnement du traitement biologique grâce à une meilleure compréhension des mécanismes microbiens en jeu.

Enfin, lors du suivi de traitement de bioremédiation ou de la surveillance environnementale, le séquençage haut-débit permet de réaliser un suivi de la dynamique des populations microbiennes d'intérêt.

## POLLUANTS CONCERNÉS

L'étude de la biodiversité microbienne par séquençage haut-débit est applicable indépendamment du polluant lorsqu'elle cible un marqueur universel comme le gène rrs. Par contre,

elle nécessite de connaître les différents taxons liés à un polluant ou à une famille de polluants (dégradation de molécules organiques ou biotransformation de métalloïdes).

## MATÉRIEL NÉCESSAIRE

Le matériel nécessaire au prélèvement des micro-organismes est identique à celui des prélèvements physico-chimiques, à la différence que les contenants doivent être stériles. Pour une eau, selon la quantité de biomasse, prévoir des volumes de un à plusieurs litres qui seront filtrés stérilement. Pour un sol, l'extraction des ADN se fait sur des faibles quantités (1 à 10 g).

D'autres approches se développent avec des cultures ou échantillonnages *in situ* de communautés microbiennes sur des supports spécifiques (échantillonneurs passifs) qui peuvent nécessiter plusieurs déplacements (installation et récupération un mois plus tard).

Du matériel de laboratoire courant est requis (pipettes, centrifugeuses, hottes, électrophorèse) ainsi que du matériel spécifique à la technique (séquenceur à haut débit).



Séquenceur haut débit Illumina.

## MÉTHODOLOGIE

L'étape de prélèvement des échantillons de sol ou d'eau diffère peu des prélèvements physico-chimiques. Une fois l'échantillon réceptionné au laboratoire, l'ADN est extrait de l'échantillon et purifié. Les échantillons peuvent être conservés congelés à -20°C ou -80°C avant et après extraction de l'ADN dans l'attente d'analyses. Les phases de prélèvement et conservation de l'échantillon et d'extraction de l'ADN sont des phases critiques dont la qualité influe directement sur les résultats.

L'ADN est ensuite amplifié par PCR et séquencé dans un séquenceur à haut débit. Un run permet de séquencer en parallèle entre  $10^6$  et  $10^8$  fragments. Les séquences obtenues sont ensuite filtrées pour éliminer celles qui contiennent des erreurs. Enfin, une analyse bio-informatique est réalisée pour fournir un résultat d'analyse. Les résultats d'analyse dépendent des bibliothèques disponibles (bases de données de séquences de référence).

## AVANTAGES – INCONVÉNIENTS – MATURITÉ DE LA TECHNIQUE

### AVANTAGES

#### Echantillonnage

- Simplicité du prélèvement comparable à celui pour des analyses physico-chimiques,
- Réalisé sur la plupart des matrices.

#### Polluant

- Technique disponible pour un grand nombre de groupes ou taxons microbiens connus et de gènes fonctionnels de biodégradation ou de biotransformation

### INCONVÉNIENTS

#### Polluant

- Besoin de connaître les voies métaboliques et les gènes cibles.

#### Laboratoires, matériel d'analyse

- Matériel très spécialisé et coûteux (séquenceur et matériel informatique),

#### Résultats d'analyse

- Complexité dans l'analyse des résultats, nécessité de personnel spécialisé en biologie moléculaire et en bio-informatique,
- Peu de bases de données existantes pour les comparaisons aux séquences de référence,
- L'identification s'arrête au niveau du genre microbien et ne permet pas une identification au niveau de l'espèce.

### MATURITÉ DE LA TECHNIQUE



R&D aboutie, indicateurs non développés, technique rarement utilisée sur le terrain

## DÉLAIS DE MISE EN ŒUVRE

La phase d'échantillonnage sur site est très rapide et identique à un prélèvement pour analyse chimique. La phase de séquençage et de nettoyage des séquences a une durée comprise entre 4 et 6 semaines. L'analyse globale des résultats est réalisée par traitement bio-informatique. Des

délais supplémentaires, entre 1 jour à 1 semaine, sont à prendre en compte pour l'analyse spécifique au regard des données physicochimiques et de la problématique du site et réalisée par un expert (recherche d'un gène spécifique, polluant peu étudié, ...).

| PHASE         | PRÉLÈVEMENT | ANALYSE | TRAITEMENT |
|---------------|-------------|---------|------------|
| Délai associé | ⌚ ⌚ ⌚       | ⌚ ⌚ ⌚   | ⌚ ⌚ ⌚      |

⌚ : jour / ⌚⌚ : semaine / ⌚⌚⌚ : mois

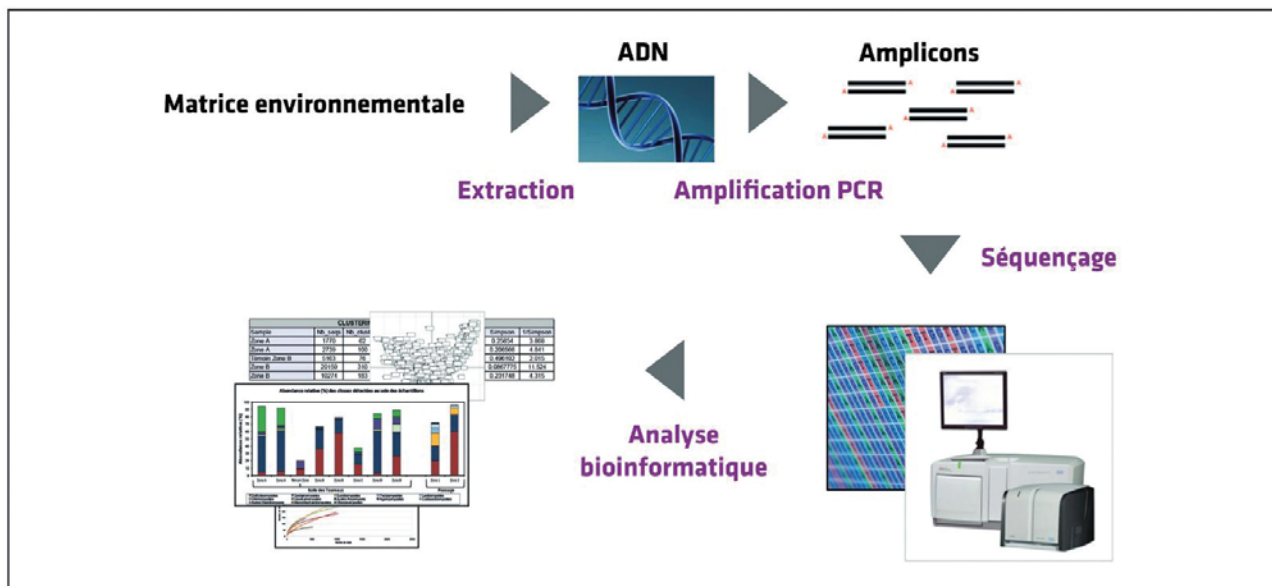
## ÉLÉMENTS DE COÛTS

Les coûts de l'échantillonnage sont faibles ; ils correspondent au coût de l'opérateur sur site. Le coût d'un séquençage en laboratoire est compris entre 150 et 300 € pour un échantillon, coût comprenant la préparation, l'analyse, la vérification de

la qualité. Une analyse globale des résultats ou un traitement spécifique par un expert des résultats doit également être pris en compte dans le cas d'une étude adaptée (polluant peu connu, outil de découverte).

| PHASE        | PRÉLÈVEMENT | ANALYSE | TRAITEMENT |
|--------------|-------------|---------|------------|
| Coût associé | €€€         | €€€     | €€€        |

€ < 100 € / €€ < 1000 € / €€€ > 1000 €



Principales étapes techniques nécessaires pour le séquençage d'amplicons.

## POUR EN SAVOIR PLUS - RÉFÉRENCES BIBLIOGRAPHIQUES

[1] *La microbiologie moléculaire au service du diagnostic environnemental*. Expertises Ademe, octobre 2017 - Fiche technique 6

[2] *Etude Ademe, Janvier 2015 : Les outils de biologie moléculaire et leur utilisation dans le domaine de la gestion des sites pollués*, J.M. Monier, S. Cécillon - Fiche F4

[3] *Fiche bioindicateur 3 Description des communautés microbiennes sur site effectuée par analyses métagénomiques* Charlotte Urien & Stéphanie Ferreira - Société GenoScreen

[4] *Norme NF EN ISO 11063 Qualité du sol - Méthode pour extraire directement l'ADN d'échantillons de sol*