HAL
open science

La microbiologie moléculaire au service du diagnostic environnemental. Synthèse de Recherche. Contrat n°1406C0050. Projet de recherche coordonné par : "Observatoire des sols vivants"

Théodore Bouchez, Anne Laure Blieux, Samuel S. Dequiedt, Isabelle Domaizon, Alexis Dufresne, Stéphanie Ferreira, Jean-Jacques Godon, Jennifer Hellal, Catherine Joulian, Achim Quaiser, et al.

To cite this version:

Théodore Bouchez, Anne Laure Blieux, Samuel S. Dequiedt, Isabelle Domaizon, Alexis Dufresne, et al.. La microbiologie moléculaire au service du diagnostic environnemental. Synthèse de Recherche. Contrat n°1406C0050. Projet de recherche coordonné par : "Observatoire des sols vivants". ADEME Agence de l'Environnement et de la Maitrise de l'Energie, 34 p., 2016. hal-01608851

HAL Id: hal-01608851

https://hal.science/hal-01608851v1

Submitted on 5 Jun 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

LA BIOLOGIE MOLECULAIRE AU SERVICE DU DIAGNOSTIC ENVIRONNEMENTAL

Août 2016

Contrat n° : 1406C0050
Projet de recherche coordonné par : *Observatoire des sols vivants*

Coordination technique ADEME : *BISPO Antonio* **– Direction\Service :** *ANGERS DPED SAF*
Coordination technique ADEME : *GRAND Cécile* **– Direction\Service :** *ANGERS DVTD SFUSP*
Coordination technique ADEME : *DEPORTES Isabelle* **– Direction\Service :** *ANGERS DECD SMVD*

ADEME
Agence de l'Environnement et de la Maîtrise de l'Energie

SYNTHESE DE RECHERCHE

REMERCIEMENTS

Les auteurs de cette synthèse remercient les experts scientifiques P Amato (CNRS), T Heulin (CNRS), B Balloy (Chambre d'agriculture de France), S Courtois (SUEZ), JY Richard (SUEZ), P Bonin (Université Aix Marseille), JM Baudoin (ONEMA), AM Pourcher (IRSTEA), A Henry (VEOLIA), pour leurs relectures et commentaires qui ont permis d'améliorer significativement la qualité de cet ouvrage.

CITATION DE CETTE SYNTHÉSE

Bouchez T[1], Blieux AL[2], Dequiedt[3], Domaizon I[4], Dufresne A[5], Ferreira S[6], Godon JJ[7], Hellal J[8], Joulian C[8], Quaiser A[5], Martin-Laurent F[3], Mauffret A[8], Monier JM[9], Peyret P[10], Schmitt-Koplin P[11], Sibourg O[9], E D'oiron [12], A Bispo[13], I Deportes[13], C Grand[13], P Cuny[14]*, PA Maron[3]*, L Ranjard[3]*

2016

La microbiologie moléculaire au service du diagnostic environnemental – Synthèse – 33 p.

Cet ouvrage est disponible en ligne www.ademe.fr,/mediatheque

[1] *Irstea, UR HBAN, 1 rue Pierre-Gilles de Gennes, 92761 Antony cedex, France*

[2] *Welience AgroEnvironnement-SATT Grand-Est, AgrOnov – RD31 – 21110 Bretenière, France*

[3] *Agroécologie, AgroSup Dijon, INRA, Univ. Bourgogne Franche-Comté, F-21000 Dijon, France*

4 INRA UMR CARRTEL Thonon, France

[5] *UMR CNRS, Université Rennes, ECOBIO, Rennes France*

[6] *Genoscreen, GENOSCREEN Campus Pasteur - 1 rue du Professeur Calmette 59000 Lille - FRANCE*

[7] *LBE, INRA, F-11100, Narbonne, France*

[8] *Bureau des Ressources Géologiques et Minières, BP6009 - 45060 Orléans cedex 1 – France*

[9] *ENOVEO, 69002 Lyon, France*

[10] *Université d'Auvergne, EA CIDAM 4678, CRBV, 63001 Clermont-Fd.*

[11] *Helmholtz Zentrum Muenchen, 85764 Neuherberg, Germany*

[12] *Observatoire Français des Sols Vivants, Domaine de Danne, 49500 St Martin du Bois, France*

[13] *ADEME, BP 90406, 49004 Angers Cedex 01*

[14] *Université Aix- Marseille, Institut Phytéas, OCEAMED, Luminy 13009 Marseille*

** Ces auteurs ont contribué de façon équitable à la coordination de l'écriture de cet article de synthèse*

Toute représentation ou reproduction intégrale ou partielle faite sans le consentement de l'auteur ou de ses ayants droit ou ayants cause est illicite selon le Code de la propriété intellectuelle (art. L 122-4) et constitue une contrefaçon réprimée par le Code pénal. Seules sont autorisées (art. 122-5) les copies ou reproductions strictement réservées à l'usage privé de copiste et non destinées à une utilisation collective, ainsi que les analyses et courtes citations justifiées par le caractère critique, pédagogique ou d'information de l'œuvre à laquelle elles sont incorporées, sous réserve, toutefois, du respect des dispositions des articles L 122-10 à L 122-12 du même Code, relatives à la reproduction par reprographie.

Résumé

Dans une société où il devient urgent de réduire l'empreinte environnementale des activités humaines, la première étape est de pouvoir faire un diagnostic de la qualité de notre environnement et des différentes matrices le constituant (eau, sol, atmosphère). Les microorganismes, de par leur petite taille, leur énorme diversité taxonomique et génétique, leur réactivité aux perturbations ainsi que par leur implication forte dans les cycles biogéochimiques gérant les flux de matière et d'énergie sont des candidats incontournables pour élaborer ce diagnostic. Ceci est d'autant plus vrai que depuis une vingtaine d'années, les fortes évolutions méthodologiques dans le domaine de la biologie moléculaire ont permis de sublimer les techniques d'étude de ces microorganismes dans l'environnement.

Toutefois, il devient maintenant important de faire un bilan de ces avancées méthodologiques et de pouvoir identifier, voire hiérarchiser, les techniques les plus à-même de fournir des bioindicateurs sensibles et robustes afin de développer le diagnostic environnemental microbiologique. Dans ce contexte, cette synthèse d'étude a pour objectif premier de décrire les différentes techniques relevant de la microbiologie moléculaire en termes d'état d'avancement, de limites techniques et de sensibilité pour étudier l'abondance, la diversité, l'activité et les potentialités fonctionnelles des communautés microbiennes indigènes de diverses matrices environnementales (eau, sol, air, substrats déchets).
Dans un deuxième temps, l'étude propose une série d'exemples d'applications illustrant des utilisations réelles de ces techniques pour évaluer ou remédier à l'impact des activités humaines (agricoles, industrielles, urbaines) sur les différentes matrices environnementales.
Enfin, la dernière partie est consacrée au référencement des indicateurs microbiens opérationnels pour des prestations de diagnostic de la qualité de différentes matrices environnementales par différents types d'opérateurs (agriculteurs, industriels, gestionnaire de sites, aménageurs).

TABLE DES MATIERE

1- Résumé ……..……..3
2- Introduction..... 5
3- Les techniques de microbiologie moléculaire 7
4- Potentialités d'application des techniques de microbiologie en tant que bioindicateurs pour le diagnostic environnemental..... 14
5-Tableau de bord opérationnel pour le diagnostic environnemental 22
6- Conclusion..... 25
7- Références 26
GLOSSAIRE..... 30

1- Introduction

De tous les organismes vivants sur Terre, c'est dans le monde de l'invisible, de « l'infiniment petit » que l'on trouve la diversité biologique la plus importante. En effet, grâce à une extraordinaire capacité d'adaptation génétique aux variations de leur environnement, les microorganismes ont colonisé, sans exception, l'ensemble des écosystèmes de notre planète. Cette répartition ubiquitaire des microorganismes s'explique par l'extraordinaire plasticité et diversité génétique qui caractérise le monde microbien. Ainsi, on ne compte pas moins de 1 million d'espèces de bactéries et 100 000 espèces de champignons par gramme de sol ; 10 000 et 100 000 espèces bactériennes par mL d'eau, et par m^3 d'air, respectivement. Au-delà de la diversité génétique, ces communautés représentent également une part très importante de la biomasse vivante dans les écosystèmes. On peut citer pour exemple la biomasse microbienne du sol qui peut représenter de 2 à 10 tonnes de carbone par hectare, soit l'équivalent d'une dizaine de vaches pâturant sur la même surface !

Cette richesse formidable confère aux microorganismes une place particulière au niveau de la biosphère en termes de réservoir de ressources génétiques qui doit être considéré comme un véritable patrimoine de l'humanité. L'énorme diversité des microorganismes se traduit également par une implication forte dans les fonctions et les services écosystémiques assurés par les matrices environnementales. Ainsi, les communautés microbiennes contribuent aux services de soutien, notamment à travers leur rôle dans les cycles biogéochimiques d'éléments majeurs tels que le carbone, l'azote, le phosphore, le soufre… La composante microbienne est, par exemple, responsable des transformations du cycle de l'azote comme la fixation de l'azote atmosphérique, l'ammonification, la nitrification et la dénitrification (Hayatsu *et al.*, 2008). De même, la minéralisation de la matière organique, processus central du fonctionnement des écosystèmes terrestres et aquatiques, se réalise en grande partie par les microorganismes qui transforment des molécules organiques complexes en éléments minéraux. Dans les écosystèmes aquatiques de l'ordre de 90% des organismes réalisant la photosynthèse sont des microorganismes. Ils sont responsables d'environ la moitié de la production primaire de matière organique de la biosphère soit de l'ordre de 56 Pg C an^{-1} (Buitenhuis *et al.*, 2013). De par leur plasticité métabolique, les microorganismes interviennent aussi dans la dégradation et l'immobilisation de polluants (ETM, pesticides…) dans l'environnement (Maier et Gentry, 2015; Roane *et al.*, 2015). Certains microorganismes ont également un impact important sur la santé et la croissance des plantes en réalisant par exemple des symbioses (Barrios, 2007) ou en induisant des maladies.

Comme tous les êtres vivants, les communautés microbiennes vivent en interaction avec leur environnement. Elles répondent de façon très sensible aux changements des conditions environnementales ce qui peut se traduire par des modifications de biomasse, de diversité et d'activité (Sharma *et al.*, 2011 ; Pulleman *et al.*, 2012). Dans les années 80, le développement de nouvelles approches moléculaires basées sur l'extraction et la caractérisation de l'ADN extrait de l'environnement (eau, sol, sédiments) a révolutionné l'analyse des communautés microbiennes dans l'environnement, historiquement basée sur des techniques pasteuriennes de mises en culture ou d'observation microscopique. L'aboutissement de ces développements est représenté aujourd'hui par l'ère des approches « OMIQUES », qui offrent la possibilité de caractériser la diversité génétique et fonctionnelle des microorganismes dans leur ensemble, sans *a priori* par l'analyse en haut débit des ADN (génomique), ARN (transcriptomique), protéines (protéomique) ou des métabolites (métabolomique).

Pour contribuer à l'établissement d'un diagnostic pertinent de la qualité de l'environnement les indicateurs biologiques doivent rendre compte de son fonctionnement et être sensibles aux modifications des conditions environnementales (Pulleman *et al.*, 2012 ; Rames *et al.*, 2013). A cet égard, les communautés microbiennes offrent un potentiel important puisque, comme évoqué précédemment, (i) elles sont présentes avec une forte densité et diversité dans tous les environnements, (ii) elles sont fortement impliquées dans le fonctionnement biologique et les services rendus par les écosystèmes, et (iii) elles répondent de façon très sensible aux changements des conditions environnementales en termes de modification de biomasse, de structure/diversité et d'activité.

Toutefois, les indicateurs doivent également répondre à des critères pratiques et économiques (*i.e.* rapides et simples à mettre en œuvre et à interpréter, reproductibles, peu onéreuses, accessibles aux utilisateurs…) et être associés à des référentiels permettant de positionner les valeurs mesurées dans une gamme de variabilité opérationnelle (faible/normal/élevé) afin de permettre l'établissement du diagnostic désiré (Ritz *et al.*, 2009 ; Pulleman *et al.*, 2012). Bien que la plupart des méthodes développées au cours des 30 dernières années pour caractériser les communautés microbiennes *in situ* aient été proposées comme des indicateurs de la qualité des matrices environnementales, toutes ne remplissent pas réellement ces différents critères. Le plus limitant étant l'absence de procédures standardisées et de référentiels associés à ces indicateurs du fait, justement, de la grande diversité des microorganismes et des environnements dans lesquels ils se développent. À ce jour, il apparaît donc opportun de faire un bilan sur les indicateurs microbiologiques et surtout sur les potentialités des outils relevant de la microbiologie moléculaire pour fournir de nouveaux indicateurs de qualité des écosystèmes répondant à tous les prérequis décrits ci-dessus.

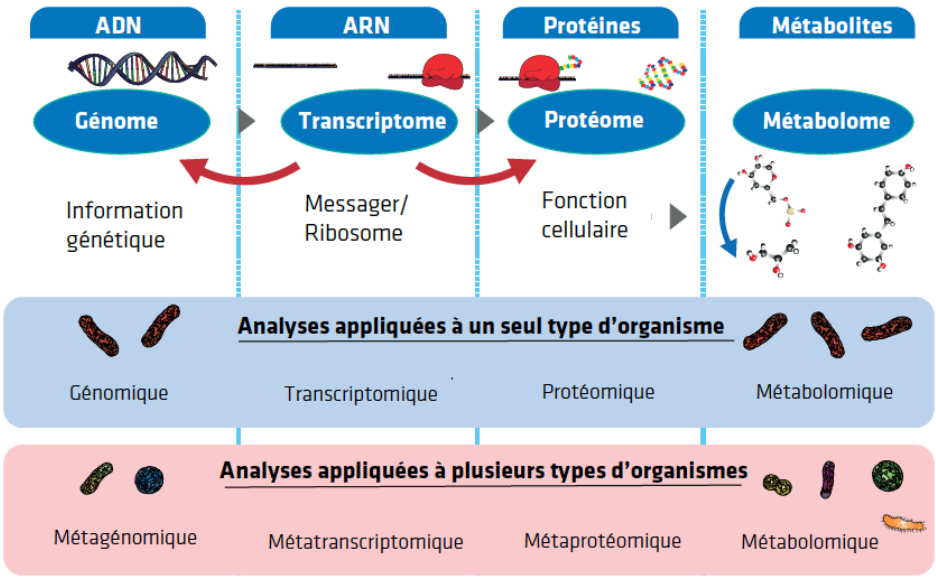
Dans ce contexte, cet article a pour objectif premier de décrire les différentes techniques relevant de la microbiologie moléculaire en termes d'état d'avancement, de limites techniques et de sensibilité pour étudier l'abondance, la structure/diversité, l'activité et les potentialités fonctionnelles des communautés microbiennes indigènes de diverses matrices environnementales (eau, sol, air, substrats déchets). Dans un deuxième temps, l'article proposera une série d'exemples d'applications illustrant des utilisations réelles de ces techniques pour évaluer ou remédier à l'impact des activités humaines (agricoles, industrielles, urbaines) sur les différentes matrices environnementales. Enfin, la dernière partie sera consacrée au référencement des indicateurs microbiens opérationnels pour des prestations de diagnostic de la qualité de différentes matrices environnementales par différents types d'opérateurs (agriculteurs, industriels, gestionnaires de déchets, aménageurs du territoire).

2- Les techniques de microbiologie moléculaire

Les enjeux de la microbiologie environnementale sont nombreux. Il peut s'agir d'identifier des microorganismes présents dans notre environnement et de comprendre leur fonctionnement ou encore d'évaluer l'impact des perturbations environnementales sur la diversité et les activités des communautés microbiennes indigènes. À terme, il faut déchiffrer le lien entre la diversité génétique/fonctionnelle de ces communautés afin de comprendre le fonctionnement intégré des écosystèmes. Ceci implique de pouvoir analyser les différentes composantes (densité, diversité, fonctionnalité, activité, interactions) qui caractérisent la communauté microbienne et leur régulation par les facteurs environnementaux. Actuellement, il existe potentiellement quatre niveaux d'intégration pour étudier les communautés leur microbiennes dans environnement (figure 1) :

- L'analyse de l'ADN des communautés indigènes qui constitue la génomique environnementale. Les techniques utilisées permettent d'accéder à la densité et la diversité génétique et fonctionnelle des communautés microbiennes ;
- L'analyse des ARN des communautés indigènes (ensemble des séquences de gènes exprimés), appelée transcriptomique environnementale, qui peut constituer un moyen d'accéder à l'identification des populations ou des fonctions actives sous certaines contraintes environnementales ;
- L'analyse des protéines : l'étude des protéines totales synthétisées à l'échelle de la communauté microbienne, appelée protéomique environnementale qui peut permettre d'accéder à la fonctionnalité des microorganismes. Par rapport au transcriptome, le protéome présente l'intérêt de cibler les enzymes réellement responsables de l'activité de la communauté dans des conditions déterminées ;
- L'analyse des métabolites : l'étude des métabolites synthétisés par les communautés microbiennes, appelée métabolomique, qui permet l'indentification des produits finaux ou intermédiaires de leur activité.

L'ensemble de ces approches peut être utilisées pour caractériser un organisme ou une communauté. Dans ce dernier cas, on utilise le préfixe « méta » (figure 1).



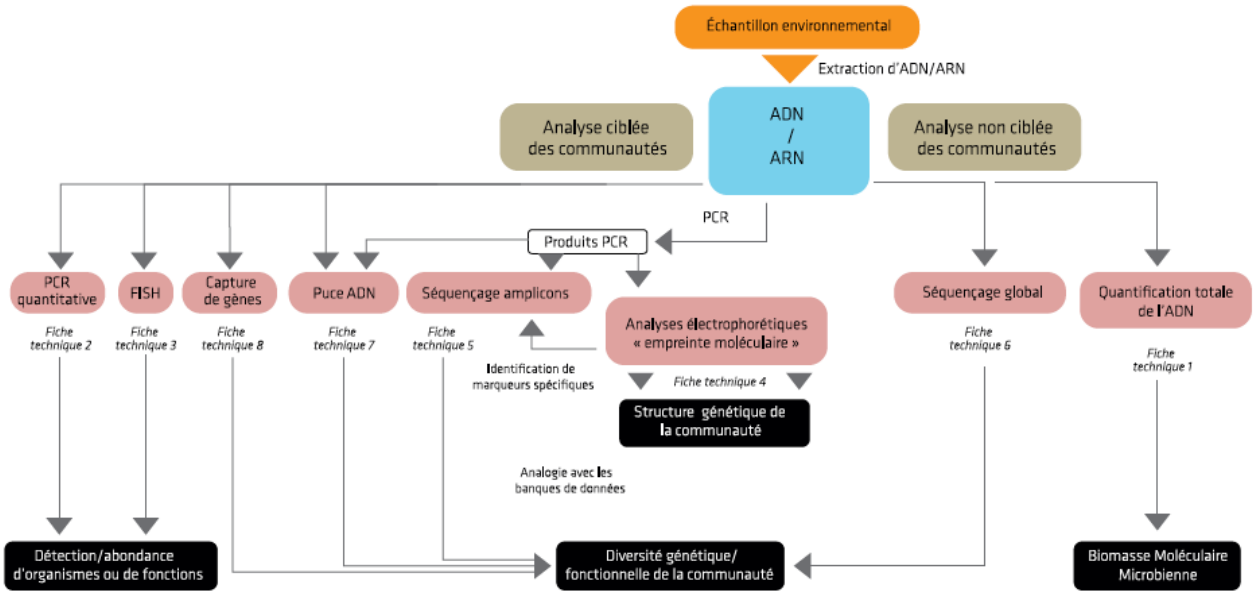
- **Figure 1 -** Différents niveaux d'intégration (ADN, ARN, protéines et métabolites) des techniques de microbiologie moléculaire pour étudier les microorganismes ou les communautés de microorganismes (approches « méta ») (messager = ARN messager, Ribosome = ARN ribosomique).

2.1 La génomique et la transcriptomique environnementale

L'ADN (Acide Désoxyribonucléique) constitue le support physique de l'information génétique pour tous les organismes vivants. L'ensemble de cette information génétique constitue le génome. Celui-ci regroupe l'ensemble des gènes (gènes codant pour les ARN ribosomaux, pour les protéines). L'ADN est une molécule de très grande taille formée par l'association de nombreux nucléotides de 4 types différents (A, T, C et G) et chaque gène est caractérisé par une séquence nucléotidique spécifique. Lire la séquence d'un gène ou d'un génome permet de décoder cette information.

L'ARN (Acide RiboNucléique) est présent chez l'ensemble des organismes et représente le produit de l'expression des gènes. Tout comme l'ADN, l'ARN est formé d'une suite de 4 nucléotides (A, U, C et G). Cependant, les différences dans la structure de la molécule d'ARN font que celui-ci est beaucoup moins stable que l'ADN et sa dégradation survient très facilement. Les molécules d'ARN sont synthétisées au cours de la transcription en utilisant l'ADN comme modèle.

L'analyse de l'ADN ou de l'ARN représente aujourd'hui le moyen privilégié pour décrire le répertoire de gènes et de fonctions associées aux microorganismes (figure 2).



- **Figure 2 -** Outils moléculaires pour l'analyse de l'ADN et de l'ARN et types d'informations obtenues.

L'analyse de l'ADN permet d'identifier les micro-organismes présents dans un échantillon, leurs proportions dans la communauté ou encore leurs potentiels physiologiques et métaboliques. L'étude des gènes et des génomes renseigne également sur les relations de parenté entre les micro-organismes et donne la possibilité de reconstruire leur histoire évolutive. Cependant, elle ne permet pas d'identifier les gènes qui sont exprimés à un instant donné. Il n'est donc pas possible de déterminer les fonctions qui sont utilisées en rapport avec l'état ou les fluctuations de l'environnement. Par ailleurs, les micro-organismes les plus actifs dans une communauté ne sont pas forcément les plus abondants. Il a aussi été montré qu'une fraction importante de l'ADN extrait correspond à de l'ADN extracellulaire, excrété activement par les microorganismes ou libéré passivement à la mort des cellules (Nielsen *et al*, 2007). Ainsi, la détection d'ADN dans un échantillon ne signifie pas que les organismes d'où proviennent ces ADN sont viables ou que les gènes associés sont fonctionnels.

L'analyse des transcrits (ARN) des gènes permet d'établir un lien plus direct entre la présence d'un organisme et son activité. Cependant, différents mécanismes de régulation intervenant sur la dégradation des ARN ou sur la synthèse des protéines peuvent fortement modifier les liens entre l'expression des gènes et l'activité des protéines. Pour cette raison, la relation entre l'expression d'un gène, la quantité de protéines correspondantes et l'activité réelle constitue une question de recherche non résolue à l'heure actuelle. L'analyse de l'ARN permet de caractériser la diversité et la composition taxonomique (ARN ribosomique) ainsi que les fonctions réalisées (ARM messager) par les micro-organismes actifs ou tout au moins viables. L'ensemble de ces informations apporte des réponses à de nombreuses questions d'écologie microbienne telles que : quels sont les micro-organismes impliqués dans la transformation d'un élément chimique ou dans la dégradation d'un polluant ? Comment les micro-organismes perçoivent et s'adaptent aux perturbations de leur environnement ? Comment interagissent-ils entre eux et avec les autres organismes du milieu ?

- ***Comment extraire l'ADN et l'ARN des matrices environnementales***

Depuis une trentaine d'années de nombreuses approches méthodologiques ont été développées pour extraire les ADN et ARN de différentes matrices environnementales. Ces approches sont basées sur la lyse dans les échantillons expérimentaux des cellules (thermique et/ou chimique et/ou mécanique et/ou enzymatique) suivi d'une séparation/purification de l'ADN du reste de la matrice environnementale (minérale et organique). Devant l'extrême diversité microbienne et les différentes matrices explorées, il n'est pas possible actuellement d'extraire l'intégralité de l'ADN présent dans un échantillon. L'application de plusieurs protocoles d'extraction pourrait permettre d'accéder à une information relativement exhaustive de la composition de la communauté (Delmont, *et al.*, 2011). Cependant, cette approche est difficilement envisageable pour un grand nombre d'études compte tenu du nombre d'échantillons à traiter voire les trop faibles quantités d'échantillon à disposition. Aussi, des approches de standardisation de cette étape d'extraction ont été initiées pouvant permettre une comparaison efficace des échantillons (Terrat, *et al.*, 2015). Pour la matrice sol, cet effort de standardisation a débouché sur une norme ISO (ISO 11063:2012) pour extraire directement l'ADN d'échantillons.

Par comparaison avec l'ADN, l'ARN est moins stable et plus sensible aux perturbations engendrées par le prélèvement des échantillons et une dégradation partielle de l'ARN provenant d'échantillons environnementaux est souvent observée. L'extraction et la manipulation de l'ARN se font donc dans des conditions beaucoup plus drastiques que l'ADN (salle blanche, traitement à la ribonucléase du matériel). Après extraction, l'ARN peut être converti en ADN complémentaire (ADNc) grâce à une étape de rétrotranscription. L'ADNc, qui possède le même degré de stabilité que l'ADN, peut alors être analysé avec les mêmes outils de biologie moléculaire (figure 2).

- ***Les différentes techniques de caractérisation de l'ADN et de l'ARN***

Les outils d'analyse des acides nucléiques des micro-organismes peuvent être regroupés en deux grandes approches, ciblés et non ciblés selon qu'ils sont basés sur l'analyse de gènes particuliers choisis *a priori* (i.e. marqueurs génétiques) ou sur l'ensemble des gènes sans *a priori* (figure 2).

Pour l'approche ciblée, deux types de marqueurs peuvent être utilisés: des marqueurs d'identité (gènes phylogénétiques) ou des marqueurs de fonctions (gènes fonctionnels). Le marqueur phylogénétique le plus utilisé en écologie microbienne est le gène (ADN) exprimant la petite sous-unité de l'ARN ribosomique (ARNr 16S chez les procaryotes et ARNr 18S chez les eucaryotes). Les marqueurs fonctionnels correspondent aux gènes codant pour les protéines. Leurs séquences sont en général plus variables que celles des ARN ribosomaux. De plus, ces gènes ont une distribution beaucoup moins ubiquiste et fortement influencée par les transferts horizontaux d'information génétique qui s'effectuent entre certains micro-organismes. Ils renseignent le plus généralement sur la présence potentielle d'une fonction dans un échantillon (présence du gène dans l'échantillon d'ADN) ou sur l'expression de cette fonction (détection du transcrit du gène dans l'échantillon d'ARN). Quelle que soit la technique utilisée, l'analyse de ces marqueurs repose au départ sur l'utilisation d'oligonucléotides (petits fragments d'ADN simple brin) synthétisés chimiquement. La complémentarité entre les bases des acides nucléiques des oligonucléotides et celles des marqueurs génétiques ciblés permet d'isoler et de caractériser ces derniers.

La PCR ("Polymerase Chain Reaction", ou "amplification génique") est utilisée pour amplifier spécifiquement une sous-partie de la séquence d'un marqueur.

Les fragments d'ADN amplifiés par PCR seront ensuite caractérisés par leurs empreintes génétiques ou leurs séquences. Sur le même principe d'appariement des bases, les approches biopuces ADN et capture de gènes par hybridation donnent la possibilité d'analyser spécifiquement un ensemble plus ou moins grands de marqueurs phylogénétiques ou fonctionnels. Toutes ces approches nécessitent d'extraire les acides nucléiques des cellules. À l'inverse, l'hybridation *in situ* (FISH : Fluorescent In Situ Hybridization) permet de détecter la présence de gènes directement à l'intérieur des cellules.

En complément de ces approches ciblées et face à l'immense diversité des micro-organismes au sein des matrices environnementales, des approches non ciblées dites « globales » ont été développées. Le principe repose sur une analyse directe et sans *a priori* de l'ADN et de l'ARN extrait des matrices environnementales. Du point de vue de la technique, l'approche globale repose sur le séquençage haut-débit de métagénomes ou de métatranscriptomes de communautés microbiennes. L'analyse bioinformatique des séquences permet d'établir « qui est là ? », « dans quelle proportion? » et « qui fait quoi ? ». Elle offre de plus l'immense avantage de pouvoir identifier de nouveaux taxons ou de nouvelles fonctions. Ces questions constituent un enjeu majeur de l'écologie microbienne pour déterminer le lien entre la structure de la communauté (diversité, abondance) et les fonctions réalisées. Cette approche est aujourd'hui la plus utilisée pour étudier la diversité et l'activité des communautés microbiennes. Toutefois, le coût du séquençage et la complexité de l'analyse des données produites avec cette approche font que son utilisation en tant qu'outil de diagnostic est encore très peu développée. Le séquençage haut-débit de métagénomes ou de métatranscriptomes génère des masses de données très complexes qui nécessitent le plus souvent de disposer de matériel informatique avec des capacités de stockage et de calculs importantes. Il faut souligner que l'analyse informatique des données de séquençage haut-débit est particulièrement chronophage. En effet, cette étape demande souvent beaucoup plus de temps que les étapes amont qui vont de la collecte des échantillons au séquençage.

Grâce à ces approches de métagénomiques différents organismes et mécanismes microbiens impliqués dans des fonctions de dépollution, (Bertin, *et al.*, 2011, Pelletier, *et al.*, 2008), dans des applications industrielles (Mao, *et al.*, 2014), dans l'adaptation à des conditions environnantes changeantes dans différentes matrices environnementales (Mondav, *et al.*, 2014 ; Tseng and Tang, 2014) ou dans les cycles biogéochimiques (Ghai, *et al.*, 2014) ont pu être mis en évidence.

Au cours de la dernière décennie, le séquençage du métatranscriptome a permis de caractériser la diversité microbienne et les fonctions exprimées par les micro-organismes dans un grand nombre d'environnements tels que : le sol (Bailly *et al*, 2007 ; Tveit *et al.*, 2013 ; *Geisen et al.*, 2015) ; les sédiments (Dumont *et al*., 2013 ; Militon *et al.*, 2016) ; la couche euphotique des océans (Frias-Lopez *et al*., 2008 ; Gilbert *et al.*, 2008 ; Gifford *et al*, 2011 ; Poretsky *et al.*, 2005 ; Poretsky *et al*., 2009) et l'océan profond (Lesniewski *et al.*, 2012) ; les lacs (Vila-Costa *et al.*, 2013) ; les tapis microbiens (Quaiser *et al.*, 2014) ou encore la rhizosphère (Turner *et al.*, 2013 ; Chapparo *et al*., 2014). Les résultats obtenus ont contribué notamment à l'identification de nouveaux acteurs clés du cycle de l'azote dans les sols (Urich *et al.*, 2008) et les océans profonds (Baker *et al.*, 2013) ; à la découverte de mécanismes de coordination de l'expression des gènes de différentes espèces d'une communauté en réponse aux variation temporelles des conditions environnementales (Ottesen *et al.,* 2013) ; à la description de la dynamique temporelle des communautés impliquées dans le recyclage de la matière organique (McCarren *et al.*, 2010).

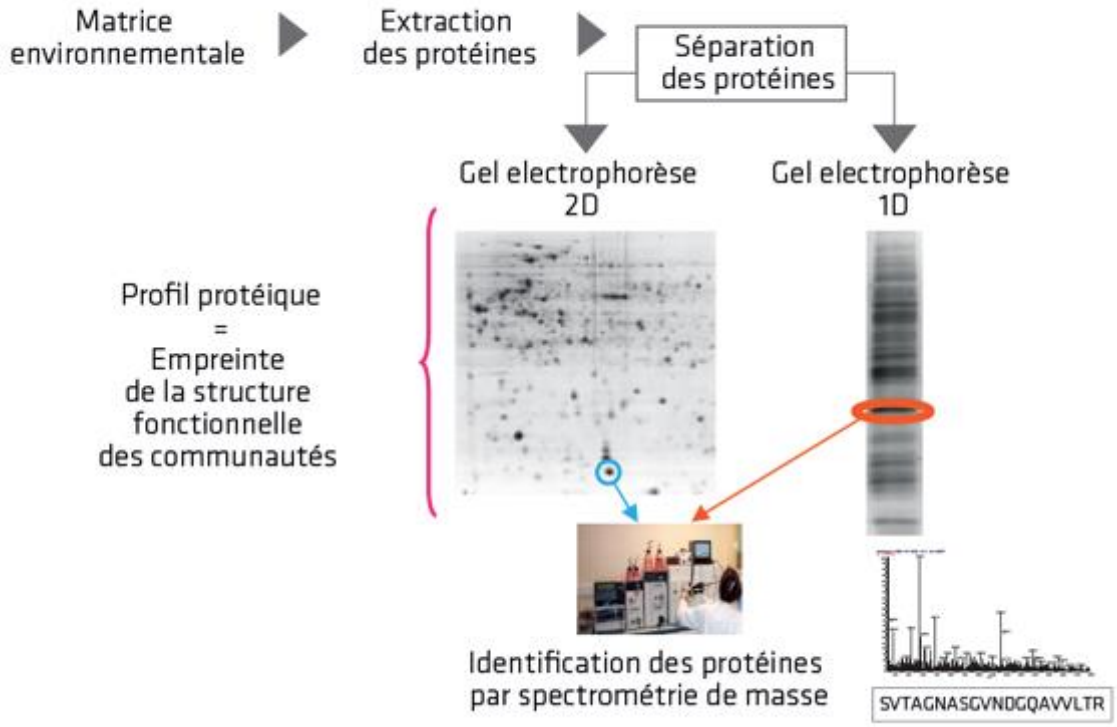
2.2 La protéomique environnementale

Cette approche cible l'analyse des protéines à l'échelle de l'organisme microbien ou de la communauté et apporte le complément fonctionnel aux travaux basés sur l'analyse de l'ADN, ou même de l'ARN. En effet, ce sont les protéines fonctionnelles et non les gènes qui agissent : elles catalysent les réactions chimiques, contrôlent les mécanismes à la base de l'activité des organismes et les interactions qu'ils ont avec leur environnement. Elles sont des indicateurs instantanés de l'état fonctionnel des organismes, tandis que les gènes informent plutôt sur leurs prédispositions métaboliques.

Le protéome désigne l'ensemble des produits des gènes fonctionnels (*i.e.* les protéines) d'un organisme à un instant donné. La protéomique consiste en une analyse extensive des protéines. Le métaprotéome, par analogie au métagénome, se définit comme l'ensemble des protéines synthétisées à l'échelle de la communauté microbienne à un temps donné. L'analyse du métaprotéome constitue ce que l'on appelle la métaprotéomique.

En raison de sa grande sensibilité, la protéomique permet de détecter des modifications phénotypiques au niveau cellulaire, avant même leur apparition au niveau macroscopique. Cette propriété en fait un outil particulièrement intéressant pour évaluer l'impact de perturbations sur les microorganismes, mais aussi pour suivre leur fonctionnement dans l'environnement.

- Le principal défi de la protéomique environnementale est de cartographier les protéines extraites de communautés microbiennes colonisant les différentes matrices environnementales et d'identifier de manière non ciblée de nouvelles voies métaboliques et leurs gènes codant associés.
- L'analyse du métaprotéome des communautés microbiennes dans l'environnement implique différentes étapes techniques, de l'extraction des protéines microbiennes de la matrice environnementale à leur séparation et leur identification (figure 3). L'analyse des protéines est le plus fréquemment basée sur la combinaison de trois méthodologies : l'électrophorèse bidimensionnelle (E-2D) qui permet de séparer les protéines selon leur point isoélectrique (pI) et leur poids moléculaire (PM), la spectrométrie de masse (SM) et la bioinformatique pour leur identification.



- **Figure 3 –** Stratégie pour l'analyse du métaprotéome des communautés microbiennes dans leur environnement.

Comme dans l'analyse des acides nucléiques, l'étape la plus cruciale pour l'étude du métaprotéome est l'extraction. Elle doit permettre une récupération d'un ensemble de protéines (i) représentatif de l'échantillon, et (ii) être de qualité et quantité suffisantes pour pouvoir être analysé avec les outils moléculaires disponibles (Leary *et al.*, 2013). La stratégie d'extraction varie en fonction de la fraction de protéines ciblée (par exemple procaryote/eucaryote; extracellulaire/cellulaire) et des méthodes mobilisées pour l'analyse (analyse comparative de cartes protéiques, mesure/détection de protéines spécifiques ou des activités enzymatiques).

Une fois les protéines extraites, selon le type d'information et le niveau de résolution voulus, différentes méthodes biochimiques peuvent être appliquées pour l'analyse du métaprotéome. Pour obtenir une "carte protéique " de la communauté de façon non ciblée, les protéines peuvent être séparées par électrophorèse sur gel uni (E-1D) ou bidimensionnel (E-2D). L'électrophorèse sur gel 2-D est préférable pour obtenir une meilleure séparation et faciliter ensuite l'identification des protéines. Les protéines intéressantes peuvent ensuite être extraites des gels pour être identifiées par spectrométrie de masse (SM). Si le génome de l'organisme ou de la communauté sur lequel on travaille est séquencé, des analyses par SM de type MALDI-TOF ("matrix-assisted laser desorption/ionisation–time of flight") permettent le plus souvent d'obtenir des identifications valables des protéines par confrontation des spectres de masses expérimentaux à ceux qui sont obtenus de façon théorique à partir des bases de données protéiques (technique du "Peptide Mass Fingerprinting"). En l'absence d'identification ou si les organismes ne sont pas séquencés, ce qui est le cas dans la grande majorité des études de protéomique environnementale, on utilisera de préférence des stratégies permettant de générer par spectrométrie de masse tandem (SM/SM) des spectres de fragmentation qui permettront un séquençage *de novo* d'un ou plusieurs peptides de la protéine. L'interprétation *in silico* de ces spectres permettra ensuite de faire des recherches par homologies avec les données disponibles dans les bases de données génomiques ou protéomiques en ligne.

La protéomique environnementale a des applications très vastes pour l'étude *in situ* et sans *a priori* de la fonctionnalité des communautés microbiennes. L'analyse du métaprotéome a été appliquée à l'heure actuelle dans une grande variété de matrices environnementales telles que le sol, l'eau (douce ou marine), les sédiments, les aquifères profonds, les effluents miniers ou encore les boues de stations d'épuration. Quel que soit l'environnement considéré, tous ces travaux montrent que le métaprotéome est très complexe, dynamique et sensible aux variations environnementales. Ils révèlent les potentialités de la protéomique environnementale pour rendre compte de l'état biologique de l'environnement.

Toutefois, même si elles ont beaucoup évolué les technologies mises en œuvre pour caractériser le métaprotéome sont encore limitées par rapport à celles utilisées pour les acides nucléiques (ADN et ARN). Ainsi, l'extraction des protéines pose encore des problèmes méthodologiques importants (Leary *et al.*, 2013 ; Keiblinger *et al.*, 2012) et représente encore un frein à l'application en routine de cette approche. De même, la séparation des protéines par E-2D qui est encore la méthode la plus utilisée reste limitée du fait de problèmes importants tels que la difficulté d'analyser les protéines faiblement abondantes, très hydrophobes, très acides ou très basiques. Pour ces raisons, l'électrophorèse sur gel 2-D a également été désignée comme "le talon d'Achille de la protéomique" (Figeys, 2000). En raison des limites liées à l'extraction et la séparation des protéines, il a été estimé que seul 1% du métaprotéome total peut être caractérisé par ces méthodes (Wilmes and Bond, 2006). Dans ce contexte, des efforts considérables ont déjà été consacrés pour développer des procédés de séparation alternatives à l'E-2D, avec la mise au point de différentes technologies telles que la séparation par chromatographique / électrophorèse capillaire (Yates, 2004), les puces à protéines (Ramachandran, 2004). Ces développements devraient à terme permettre l'analyse à haut débit de mélanges complexes de protéines et d'augmenter les applications possibles de la protéomique dans les études environnementales.

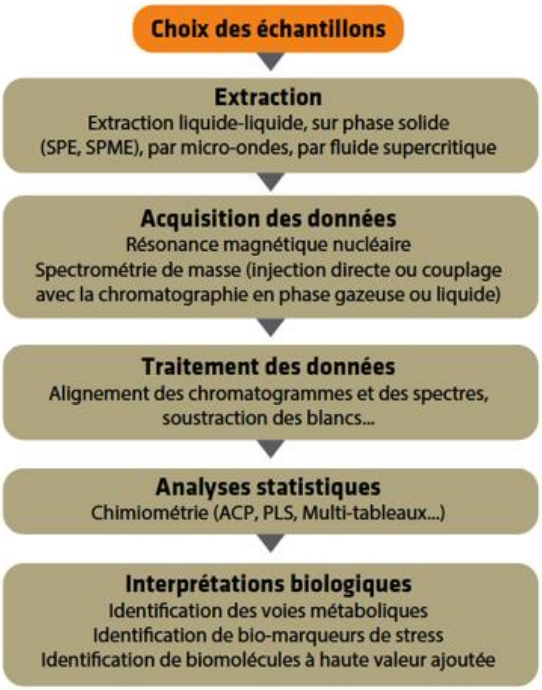
2.3 La métabolomique environnementale

La métabolomique est le plus récent représentant des technologies dites "omiques". Cette approche, ne fait pas appel *sensu stricto* aux techniques de microbiologie moléculaires et ne débouche pas encore sur le développement de techniques utilisables en routine sur des matrices environnementales, mais elle présente à terme de fortes potentialités d'application dans le diagnostic environnemental par le développement de bioindicateurs de la qualité microbienne ou de potentialités fonctionnelles de différents écosystèmes. La métabolomique consiste à analyser l'ensemble des molécules de faible poids moléculaire synthétisées par un organisme (intermédiaires métaboliques, hormones...).

À l'échelle d'un organisme, l'analyse du métabolome permet d'avoir une vision réelle des voies métaboliques activées et de leur importance (Fiehn, *et al.*, 2000 ; London et Houck, 2004, Nicholson, *et al.*, 1999) et donc d'étudier les mécanismes de réponse à différents stress (température, stress hydrique, UV, limitation nutritionnelles, pollutions chimiques...). Ce concept a été introduit dans la science du vivant par Holmes, Lindon et Nicholson à la fin des années 1990 pour être rapidement développée par la suite (Nicholson, *et al.*, 1999). De nos jours, la métabolomique est appliquée dans différents domaines, de la recherche biologique fondamentale, médicale jusqu'aux sciences environnementales.

La métabolomique environnementale porte sur l'application de cette approche à l'étude des interactions entre les organismes et leur environnement (Lankadurai, *et al.*, 2013). Il est ainsi possible, par exemple, (i) d'étudier la régulation et l'évolution des voies métaboliques en fonction des conditions environnementales, (ii) d'identifier des produits à haute valeur ajoutée bio-synthétisés et des marqueurs d'exposition à un stress, (iii) d'étudier les mécanismes d'adaptation à des perturbations environnementales telles qu'une pollution chimique ou le changement climatique. Cette approche peut s'appliquer non seulement à un organisme mais il est également possible d'étudier le métabolome d'une communauté d'organismes dans son ensemble.

Les différentes étapes techniques nécessaires à l'étude du métabolome d'un organisme ou d'un groupe d'organismes sont présentées dans la figure 4. Le choix des échantillons est particulièrement critique en métabolomique environnementale car l'identification de marqueurs de stress découle souvent de la comparaison des métabolomes exprimés par des organismes soumis et non soumis au stress environnemental étudié. Ceci est à l'image de ce qui est fait en recherche médicale ou l'on compare les métabolomes d'individus sains et malades. Les échantillons étant sélectionnés il est primordial de les rendre compatibles avec la ou les méthodes analytiques utilisées. Il est ainsi nécessaire de procéder à l'extraction des métabolites (Mushtag *et al.*, 2013) via différentes procédures : extraction liquide-liquide, extraction assistée par micro-ondes, extraction en phase supercritique, extraction sur phase solide (SPE), micro-extraction sur phase solide (SPME). Lors de cette étape, le choix des conditions d'extraction (nature des solvants, des phases solides utilisées,…) dépend de la stratégie analytique envisagée.



- **Figure 4 -** Les différentes étapes techniques d'une approche métabolomique.

Ainsi, si on cherche à être le plus exhaustif possible, ce qui est souvent le cas en métabolomique, on sélectionnera les solvants ou phases solides permettant d'extraire une large gamme de composés. À l'inverse, si on souhaite se focaliser sur des familles de composés spécifiques on choisira des conditions plus sélectives. Les extraits obtenus sont ensuite analysés par résonnance magnétique nucléaire (RMN) et/ou spectrométrie de masse, qui sont les deux techniques analytiques les plus adaptées.
Les données générées sont souvent représentées sous forme de spectre (profil de pics plus ou moins importants) et doivent être prétraitées (alignement des spectres, soustraction des blancs et du bruit de fond…) afin de s'assurer que les métabolomes obtenus vont pouvoir être comparés entre eux sans artéfacts analytiques. Vient ensuite l'étape critique de l'analyse statistique des données par des méthodes relevant de la chimiométrie (Analyse en Composante Principale (ACP), Partial Least Squares (PLS), Analyse multi-tableaux…). Ces approches vont permettre d'identifier les métabolites ayant des niveaux d'expression différents entre les échantillons (Lucio, 2009). Cette approche appelée métabolomique non ciblée permet l´élucidation du métabolisme sans hypothèse de départ. L´objectif étant de détecter le plus grand nombre de composés possible afin d'identifier les voies métaboliques mises en jeu et d'identifier parmi ceux-ci les composés représentant des marqueurs d'exposition aux stress environnementaux.

À ce jour, de plus en plus d'applications de cette technique dans des matrices environnementales (sols, eaux, déchets) sont décrites (Lankadurai, *et al.*, 2013). Toutefois, il faut constater qu'aucune stratégie de métabolomique n'est capable de couvrir l'intégralité du métabolome d'un organisme ou d'une communauté d'organisme. Par conséquent, explorer le métabolome d'un écosystème donné nécessite l'utilisation combinée de plusieurs techniques d'extraction et l'utilisation d'approches analytiques de très haute résolution.

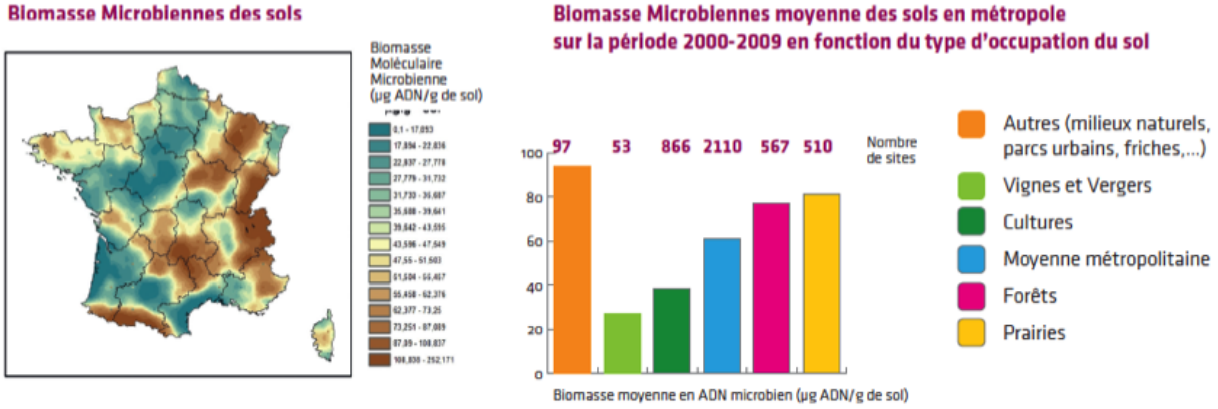
3- Potentialités d'application des techniques de microbiologie en tant que bioindicateurs pour le diagnostic environnemental

Les différentes techniques de microbiologie moléculaire décrites précédemment ont permis non seulement de mieux explorer, comprendre et même prévoir les écosystèmes microbiens en termes d'abondance, de diversité et d'activité mais aussi de développer de nouveaux outils de bio-indication sensibles et robustes permettant d'évaluer l'impact des perturbations naturelles ou dues aux activités humaines. Ces différents bioindicateurs sont :

- la biomasse moléculaire microbienne ;
- la détection et le dénombrement *in situ* de certains organismes ou gènes de fonction microbiens ;
- La diversité et composition taxonomique et fonctionnelle des communautés microbiennes.

3.1 La biomasse moléculaire microbienne

L'ADN étant présent dans tous les microorganismes (bactéries, champignons, virus, etc.), la quantité d'ADN extrait peut être utilisée pour estimer la biomasse microbienne (quantité totale de microorganismes vivants) dans les matrices environnementales. La biomasse représente la capacité d'un écosystème à héberger une quantité plus ou moins importante de microorganismes (bactérie, champignons, virus, …). Elle est déterminante pour la qualité biologique des écosystèmes en raison de son rôle dans la régulation et la transformation des nutriments. Elle est également sensible aux perturbations des écosystèmes dues aux activités humaines (agricoles, industrielles, urbaines). L'estimation de la biomasse par quantification de l'ADN microbien représente donc, malgré l'existence de biais d'extraction, un indicateur sensible et robuste de l'état biologique des matrices environnementales : la biomasse moléculaire microbienne.
L'interprétation de l'indicateur se fait par comparaison des valeurs avec un référent local (situation témoin) ou par positionnement dans un référentiel local ou national. Pour la matrice « sol » par exemple, la mesure de la Biomasse Moléculaire Microbienne a été réalisée sur les 2200 échantillons du Réseau de Mesure de la Qualité du Sol (RMQS) ce qui a permis de réaliser une cartographie nationale et de produire un premier référentiel d'interprétation associé à cette mesure (Figure 5) (Dequiedt *et al.*, 2011 ; Horrigue *et al.*, 2016).



- **Figure 5 -** Variations de la biomasse moléculaire microbienne à l'échelle nationale (à gauche) et en fonction des modes d'usage des sols (à droite). La biomasse moléculaire microbienne moyenne des sols français est de 61 µg ADN par gramme de sol. Les sols sous prairie montrent la biomasse microbienne la plus élevée suivi de ceux sous forêts. Les sols sous grande culture possèdent une biomasse microbienne inférieure à la valeur moyenne nationale. Les sols sous vignes et vergers ont la biomasse microbienne la plus faible avec seulement 27 µg ADN par gramme de sol en moyenne (Dequiedt *et al.*, 2011 ; Horrigue *et al.*, 2016).

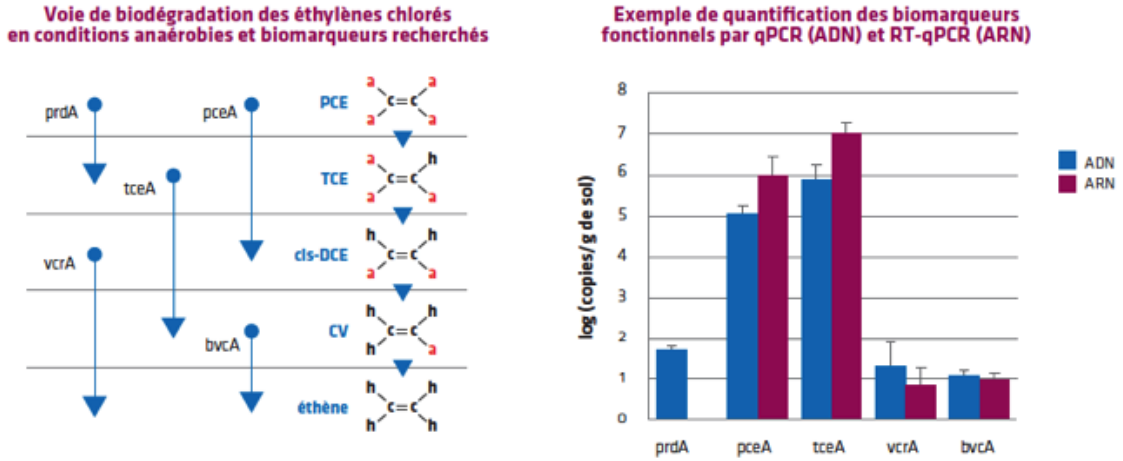
3.2 La détection et/ou le dénombrement *in situ* d'organismes microbiens ou de fonctions microbiennes

La méthode de PCR quantitative en temps réel (qPCR) est une approche moléculaire qui permet de quantifier des marqueurs (*i.e.* gènes) spécifiques dans un échantillon d'ADN extrait de matrices environnementales. La cible de la quantification peut-être un marqueur taxonomique ou fonctionnel représenté par un fragment d'ADN ou d'ADN complémentaire de l'ARNm. Cette méthode de biologie moléculaire représente une alternative aux méthodes microbiologiques basées sur la culture pour la détection et quantification de microorganismes et de leurs activités dans l'environnement. La PCR a pour principe de permettre l'amplification, à partir des acides nucléiques extraits, de fragments d'ADN cibles grâce à une ADN polymérase thermostable. La qPCR a pour particularité de déterminer la quantité initiale de l'ADN cible à partir d'une mesure de la quantité d'ADN produite dans une réaction à l'issue de chaque cycle. Il est donc possible d'estimer le nombre de séquences taxonomiques d'un microbe ou de gènes de fonctions au sein d'une communauté complexe.

La qPCR permet une détermination sensible et robuste de l'abondance de marqueurs génétiques au sein de la communauté microbienne des matrices environnementales. Selon les marqueurs ciblés, cette technique peut permettre de générer différents indicateurs de la qualité de ces matrices:

- les marqueurs taxonomiques (*e.g.* gènes ribosomiques) permettent de déterminer l'abondance des microorganismes. Selon la spécificité des amorces utilisées pour la PCR, cette quantification peut s'effectuer à différents niveaux, de la communauté totale (ex. ensemble des bactéries ou des champignons) à la détermination de l'abondance de familles, espèces ou souches spécifiques (*e.g.* pathogènes, symbiotes, etc). La figure 6 décrit un exemple d'application pour la détection de cyanobactéries dans des environnements lacustres.
- les marqueurs fonctionnels (*i.e.* gènes de fonction qui codent pour des protéines impliquées dans des fonctions d'intérêt : dégradation de polluants, pathogénie, voies de transformation de la matière organique associées aux cycles biogéochimiques, etc). La quantification de ces marqueurs permet une estimation de potentiel d'expression de la fonction ciblée au sein de la communauté (exemple d'application dans la figure 7).

- **Figure 6 -** Détection et quantification par qPCR de l'espèce de cyanobactérie *Planktothrix Rubescens* et de son gène de production de la cyanotoxine au sein d'un lac péri alpin en France en fonction des conditions mésotrophes ou eutrophes (Domaizon *et al.*,2013 ; Savichtcheva *et al.,* 2014) sur la période 1939 - 2008. Cette analyse a permis de mettre en évidence la présence historique d'un génotype unique de *Planktothrix rubescens* présent tout au long du siècle dernier, systématiquement porteur du gène codant pour la production de cyanotoxines, qui est à l'origine d'efflorescences massives (blooms) spécifiquement durant les périodes d'état mésotrophe (lac du Bourget).

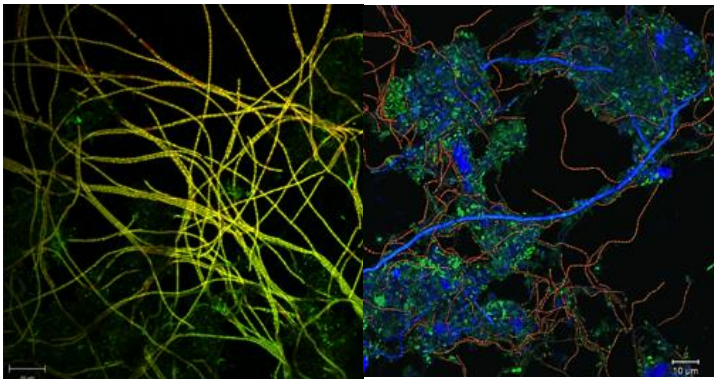


- **Figure 7 -** Voie métabolique et gènes impliqués dans la dégradation des éthylènes chlorés dans un sol contaminé (à gauche) (adapté de Freedman et Gosset, 1989) quantification des gènes impliqués (à droite) (Monier, com. pers.). La quantification des marqueurs par qPCR permet d'établir la présence/absence des microorganismes ou des fonctions de biodégradation d'intérêt. Les données générées servent de support à la prise de décision concernant la mise en place d'un traitement. Les résultats bruts obtenus en conditions de laboratoire doivent être contextualisés afin d'évaluer les meilleures stratégies de dépollution pouvant être mises en place sur site.

- *Détection et Dénombrement par FISH (hybridation in situ par Fluorescence)*.

Apparue à la fin des années 80, le marquage FISH (acronyme anglais pour "Fluorescent *In Situ* Hybridization") permet d'identifier et de dénombrer les cellules de micro-organismes appartenant à un groupe taxonomique ("famille" d'organismes) plus ou moins large. Le principe de la technique FISH est basé sur l'utilisation de petites sondes d'ADN capables de reconnaître et de s'accrocher (hybridation) à ces régions spécifiques (gènes ribosomiques ou fonctionnelles) directement à l'intérieur des cellules de micro-organismes. Ces sondes correspondent à des oligonucléotides de synthèse et portent une molécule fluorescente. Après hybridation, les cellules marquées apparaissent fluorescentes et peuvent être comptées en microscopie à épifluorescence.

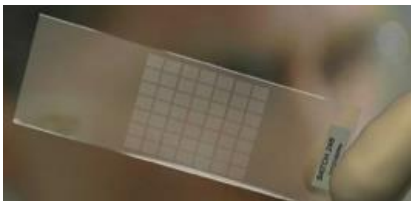
La sensibilité et la spécificité du marquage FISH en font un outil particulièrement efficace pour détecter la présence d'un taxon microbien dans toutes sortes de matrices environnementales (sol, eau, sédiment, air). Il s'agit de la seule méthode permettant de réaliser une quantification directe de l'abondance des cellules. Ces indicateurs (présence et abondance) sont très utiles pour déterminer la qualité microbiologique d'un environnement, détecter la présence d'un taxon ou d'un groupe fonctionnel de microorganismes et estimer les effets des perturbations environnementales sur les communautés microbiennes (exemple d'application sur les eaux usées de station d'épuration, figure 8).



- **Figure 8 -** Détection et quantification par la technique FISH de deux espèces de bactéries filamenteuses responsables de dysfonctionnements biologiques au sein d'une station d'épuration. **A**. *Thiothrix sp.,* **B**. *Microthrix parvicella* (rouge) - *Crédit photo : N. Durban - Irstea* (Durban *et al.*, 2015). Ces analyses ont permis d'observer l'évolution de la morphologie des flocs et de la localisation des bactéries filamenteuses. Plus précisément, elles ont mis en évidence l'internalisation des bactéries filamenteuses, de *M. parvicella* mais également des bactéries du phylum Chloroflexi à l'intérieur des flocs, améliorant ainsi leur compaction et donc leur décantation (évaluée par la mesure de l'indice de boues).

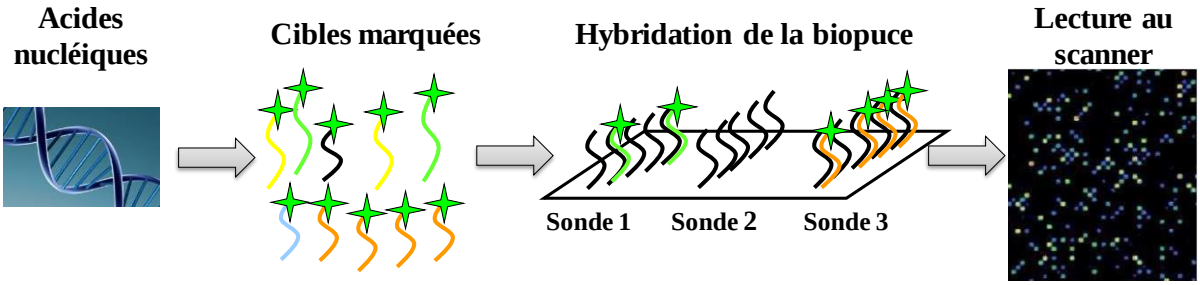
- *Biopuces à ADN*.

Une biopuce ADN est un ensemble de molécules d'ADN, appelées sondes, fixées sur une petite surface composée de verre, de silicium ou de plastique. Grâce aux nouvelles technologies de miniaturisation, des centaines de milliers, voire des millions de sondes différentes peuvent être fixées sur une seule biopuce ADN (figure 9). Il est donc possible d'identifier plusieurs milliers de gènes en une seule expérience.



- **Figure 9 -** Lame en verre comportant dans sa partie centrale les zones de fixation des molécules d'ADN (sondes).

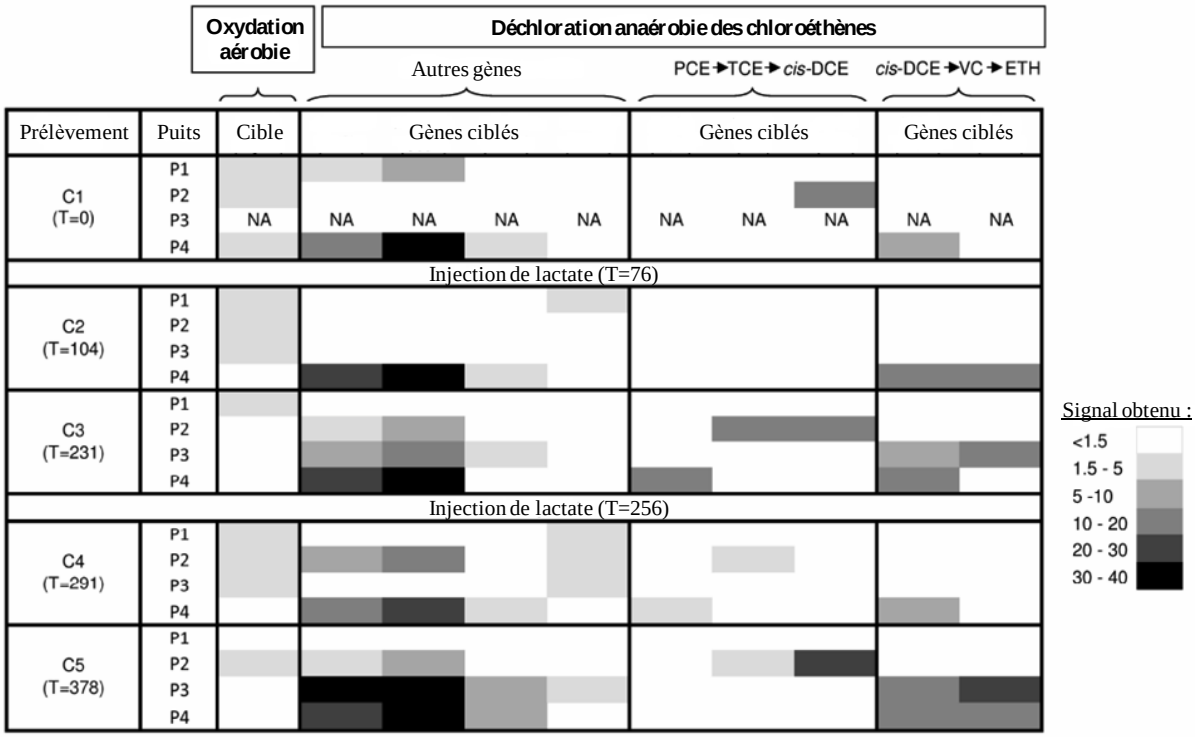
Le principe général des biopuces ADN repose sur une reconnaissance spécifique par appariement des sondes immobilisées, et de leurs cibles (généralement des acides nucléiques comme des produits PCR, de l'ADN ou des ARN), durant l'étape d'hybridation (voir figure 10 ci-dessous). Cette reconnaissance est possible grâce à la propriété qu'ont les bases (**A**dénine, **T**hymine, **C**ytosine et **G**uanine) constituant un brin d'ADN de reformer spontanément une double hélice lorsqu'il se trouve face à un brin d'ADN complémentaire (Joux *et al.*, 2010).



- **Figure 10 -** Principe des biopuces ADN. Les cibles (ADN ou ARN) marquées par un fluorochrome s'hybrident spécifiquement avec les sondes qui leur sont complémentaires, puis l'analyse de l'image obtenue permet de déterminer quels sont les gènes présents ou exprimés dans l'échantillon.

L'utilisation des biopuces ADN nécessite l'obtention d'acides nucléiques de la matrice étudiée (sol, eau, atmosphère, déchets), et donc le passage par une étape préalable d'extraction des acides nucléiques (ADN ou ARN). Concernant les études d'écologie microbienne, deux types de biopuces ADN sont le plus souvent utilisés (Dugat-Bony *et al.*, 2012) : les biopuces phylogénétiques (dont les sondes ciblent des gènes marqueurs taxonomiques) permettant d'identifier les micro-organismes, et les biopuces fonctionnelles (dont les sondes ciblent des gènes codant pour des protéines impliquées dans des fonctions d'intérêt) permettant de révéler les fonctions métaboliques.

La PhyloChip (Brodie, *et al.*, 2006) porte environ 500 000 sondes d'une longueur de 25 bases et couvre quasiment la totalité de la diversité des communautés procaryotes répertoriés dans les bases de données. Elle a été utilisée dans de nombreuses études environnementales. Cependant, en raison d'une connaissance partielle des microorganismes de l'environnement, les informations obtenues par ces approches ne reflètent qu'une partie de la diversité présente dans les environnements complexes. Les biopuces ADN du fait de leurs coûts réduits, de leurs capacités à cibler spécifiquement un grand nombre de marqueurs génétiques tant d'un point de vue qualitatif que quantitatif, de leurs facilités d'utilisation et d'analyses restent encore compétitives face au séquençage massif (Zhou, *et al.*, 2015). Une application de cette approche est présentée dans la figure 11 ci-dessous.



- **Figure 11 -** Réponse des gènes impliqués dans les voies de dégradation des solvants chlorés durant la bioremédiation d'une nappe phréatique polluée. Le signal obtenu pour chaque gène est représenté par une coloration de la case, selon le puit concerné (P1, P2, P3 ou P4), et la date de prélèvement (T : nombre de jours). PCE : perchloroéthylène, TCE : trichloréthylène, DCE : dichloroéthylène, VC : chlorure de vinyle, ETH : éthylène). L'analyse des résultats obtenus avec la biopuce ADN a confirmé l'origine biologique du processus de dégradation et a montré qu'une seule voie métabolique était impliquée : la déchloration réductrice anaérobie. Plusieurs microorganismes ont été identifiés permettant d'assurer la dégradation complète du contaminant chimique en une molécule inoffensive, l'éthylène (Dugat-Bony *et al.*, 2012).

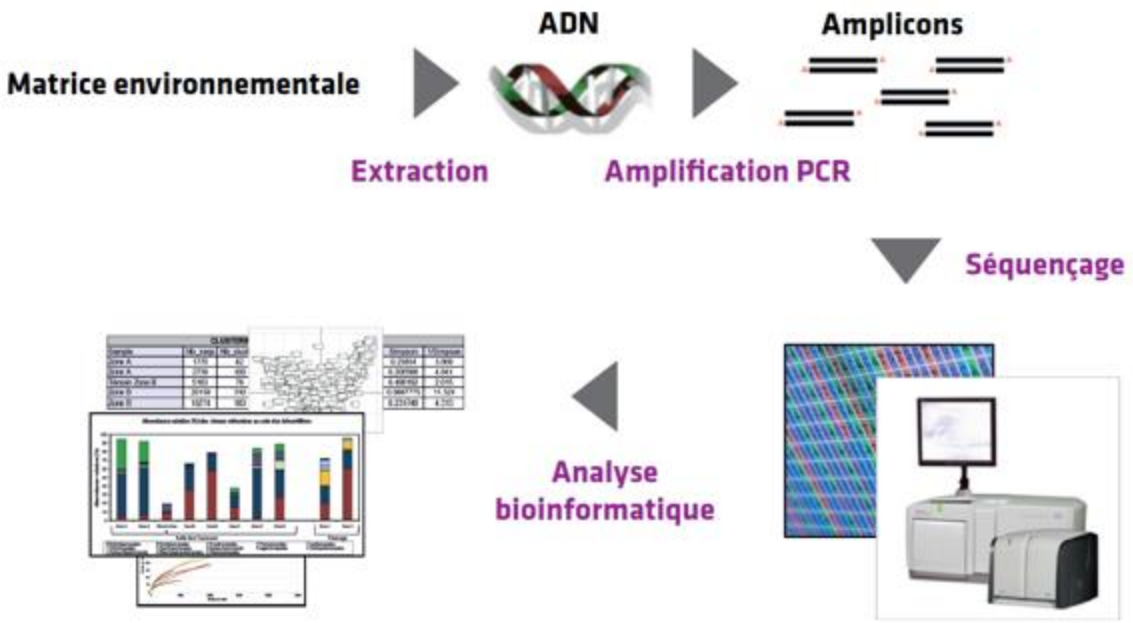
3.3 La diversité et composition taxonomique et fonctionnelle des communautés microbiennes

Si la qualité des écosystèmes est dépendante de l'abondance des microorganismes indigènes, elle est encore plus sous l'influence de la diversité des organismes et des fonctions qu'ils peuvent héberger. En effet, la diversité microbiologique en termes de nombre de taxons ou de fonctions différentes est directement reliée aux fonctions biologiques que peut supporter un écosystème (comme la fertilité, la dégradation des polluants, la production végétale, l'effet barrière aux espèces invasives…) (Maron *et al.*, 2011).

Au sein d'une matrice environnementale, la diversité et composition microbienne taxonomique peut être abordée par différentes approches comme les puces à ADN (cf paragraphe ci-dessus), le séquençage massif d'amplicons ou le séquençage massif global.

Séquençage massif d'amplicons

Le séquençage d'amplicons a pour objectif de déterminer la séquence de fragments d'ADN amplifiés par PCR. Il nécessite plusieurs étapes techniques puis un travail analytique résumés dans la figure 12 :



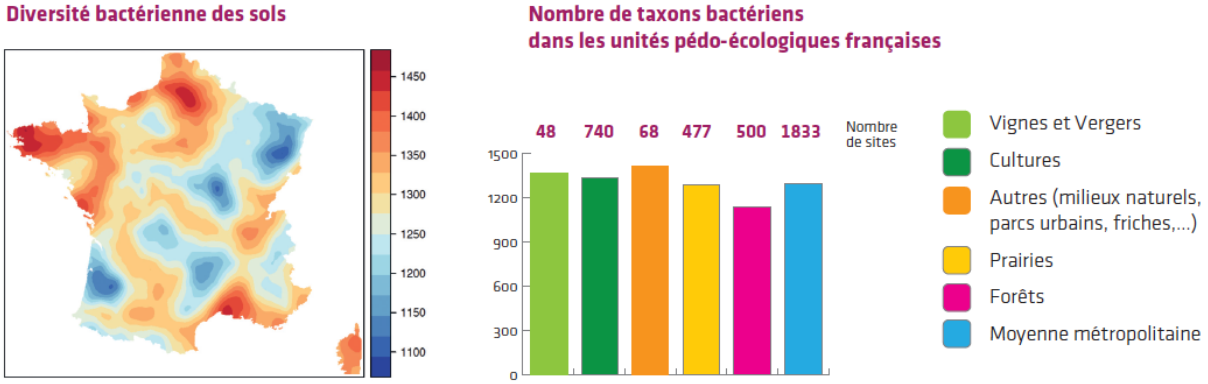
- **Figure 12 -** Principales étapes techniques nécessaires pour le séquençage d'amplicons.

L'ADN doit d'abord être extrait de la matrice étudiée (sol, eau, atmosphère, déchets). A partir de cet ADN, les amplicons sont générés par amplification par PCR de la séquence d'ADN d'intérêt.

A l'issue de l'amplification, le produit PCR est constitué d'un mélange d'amplicons représentatifs de l'abondance et de la diversité du gène ciblé au sein de la communauté microbienne étudiée. Les amplicons sont ensuite séquencés par une technique de séquençage massif sur des séquenceurs automatiques nouvelle génération (pyroséquençeur Roche©, MiSeq Illumina©, ou autre).
Il est ainsi possible d'obtenir plusieurs milliers à plusieurs dizaines de milliers voire plus de séquences du gène ciblé pour un échantillon. Ce grand jeu de données de séquences doit ensuite être analysé par des outils de bioinformatique, organisés dans un pipeline bioinformatique. A ce niveau, il existe plusieurs pipelines disponibles gratuitement sur le net (Mothur, Qiime, Pangea, GnSPipe…) qui permettent de filtrer, trier, classer, regrouper et affilier les séquences produites par références aux séquences contenues dans les bases de données internationales.

Dans les études d'écologie microbienne, le séquençage d'amplicons à partir des ADN issus de matrices environnementales permet de déterminer la variabilité de la composition de gènes d'intérêt entre les populations porteuses de ces gènes au sein de la communauté. Selon les gènes ciblés, cette technique peut permettre de générer différents indicateurs de la qualité des matrices environnementales. Deux types de gènes sont le plus souvent utilisés :

- les gènes ribosomiques qui sont des marqueurs taxonomiques dont le séquençage permet de déterminer la diversité taxinomique et la composition des communautés de microorganismes. La diversité microbienne, est un facteur déterminant dans la qualité biologique des écosystèmes en raison de son rôle dans le recyclage des nutriments, la dégradation des polluants et la stabilité même des écosystèmes. C'est un indicateur sensible des perturbations d'un écosystème (modifications de pratiques agricoles sur un sol, pollutions d'un cours d'eau ou de sédiments, changement de statut trophique d'un lac, étape de maturation d'un compost, etc.) L'exemple d'application décrit dans la figure 13 montre une analyse de la diversité taxonomiques des communautés bactériennes du sol à l'échelle nationale ;
- les gènes de fonction qui codent pour des protéines impliquées dans des fonctions d'intérêt (ex. dégradation de polluants, pathogénie, transformations des cycles biogéochimiques, etc). Le séquençage de ces gènes permet d'évaluer la structure et la diversité fonctionnelle des communautés impliquées dans ces activités. Cette mesure constitue un bon indicateur de l'état de ces communautés dans l'environnement mais aussi de leur évolution en réponse des perturbations anthropiques ou naturelles. Il renseigne sur les potentialités fonctionnelles des communautés ciblées.



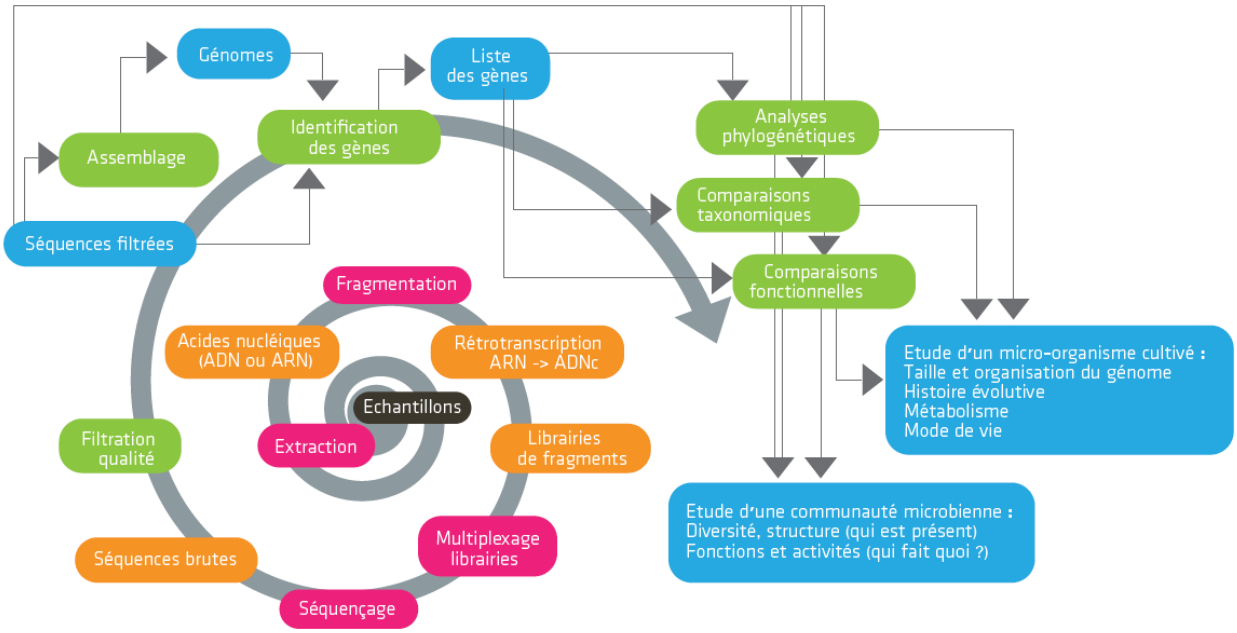
- **Figure 13** - Variation de la diversité bactérienne du sol (en nombre de taxons par gramme de sol) à l'échelle de la France (à gauche) et en fonction des modes d'usage des sols (à droite), (Terrat *et al.*, soumis). Cette carte montre des variations importantes de la diversité avec des zones où le nombre de taxons est important (en rouge sur la carte) et d'autres ou il est faible (en bleu sur la carte). A cette échelle aucune influence du climat ni de la géomorphologie (présence de montagne, de rivière, de bord de mer) n'est démontrée. Des études statistiques poussées montrent que ces variations observées à l'échelle nationale de la diversité bactérienne sont influencées par le type de sol (en termes de pH, de texture et rapport C/N) mais aussi en fonction du mode d'usage (agricole, prairie, forêt). Les sols sous prairie et forêts présentent ainsi les niveaux de diversité les plus bas (1291 et 1136 taxons respectivement) par rapport aux sols agricoles ou viticoles (1363 et 1409, respectivement). Ceci peut s'expliquer par le concept écologique de la "perturbation intermédiaire" qui prédit que la biodiversité d'un écosystème est maximum lorsque ce dernier subit une perturbation d'intensité intermédiaire (ni trop forte et ni trop

faible non plus) et minimum lorsqu'il subit une perturbation faible (phénomène d'exclusion compétitive des espèces) ou forte (sélection des espèces). Ainsi, les sols sous forêts et prairies représentent des écosystèmes qui subissent des perturbations faibles pour les communautés bactériennes par la quasi absence de l'action de l'homme et renferment donc une diversité faible de bactéries. À l'inverse, les sols agricoles et viticoles qui subissent généralement une multitude d'interventions correspondent à des systèmes plus perturbés (mais pas trop) et présentent donc une diversité bactérienne plus élevée.

Séquençage global de l'ADN ou de l'ARN

Le séquençage global de l'ADN extrait d'un échantillon d'une matrice environnementale permet d'accéder aux gènes de l'ensemble des micro-organismes présents sans passer par une étape d'amplification par PCR. Les informations obtenues permettent d'identifier les micro-organismes et leurs fonctions. La même approche de séquençage global peut être utilisée pour séquencer l'ARN d'une communauté microbienne extrait d'un échantillon environnemental et ainsi connaître les micro-organismes actifs et les gènes qui sont exprimés (figure 14).

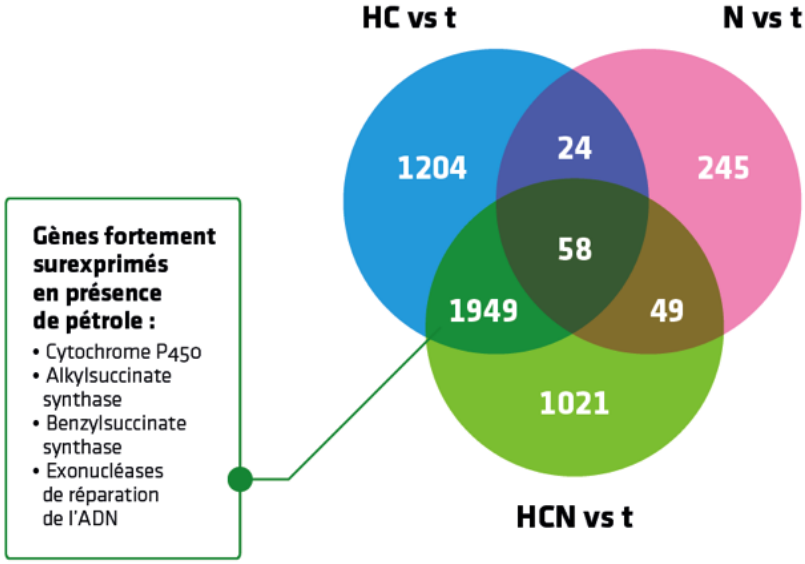
D'un point de vue technique, le séquençage global repose sur deux étapes essentielles : la fragmentation des molécules d'ADN ou d'ARN en une multitude de petits fragments partiellement chevauchants et plus ou moins redondants et la comparaison des séquences de ces fragments pour les assembler et reconstituer la ou les séquences présentes dans l'échantillon analysé.



- **Figure 14 -** Processus de séquençage global d'un (méta)génome ou d'un (méta)transcriptome. La spirale grise représente le processus général allant de l'échantillon jusqu'aux données analysées. Les principales étapes de biologie moléculaire sont en rouge, les résultats de ces étapes sont en orange. Les étapes d'analyses de bioinformatique sont en vert et les résultats correspondant sont en bleu. Les flèches noires représentent la multiplicité des analyses possibles de bioinformatique. Adaptée de Raes et Bork (2008).

Contrairement au séquençage d'amplicons, le séquençage global ne cible aucun gène en particulier. De plus, cette approche ne demande pas d'avoir une connaissance *a priori* des micro-organismes étudiés. Elle peut donc être appliquée à n'importe quelle communauté microbienne et dans n'importe quelle matrice environnementale. Il faut noter que cette approche peut aussi s'appliquer pour séquencer le génome ou le transcriptome de micro-organismes cultivés au laboratoire. Les données obtenues contiennent à la fois des séquences de marqueurs phylogénétiques (gène et transcrits d'ARN ribosomaux) et de marqueurs fonctionnels (gènes codant pour les protéines et ARN messagers).

Un exemple d'application de cette approche pour analyser la réponse des communautés microbiennes à une pollution de sédiments marins par l'apport de pétrole est présenté dans la figure 15 ci-dessous :



- **Figure 15 -** Diagramme de Venn montrant le nombre de gènes sous- ou sur-exprimés dans des sédiments marins côtiers contaminés avec du pétrole comparativement aux sédiments témoins sans aucune addition (t). Trois traitements sont comparés aux sédiments témoins : addition de 5 000 ppm de pétrole brut Rebco (HC), addition de 1 000 ind.m^{-2} de l'annélide marin *Hediste (Nereis) diversicolor* (N), et addition de pétrole et de *H. diversicolor* (HCN). Les nombres situés dans les intersections des cercles correspondent aux gènes dont l'expression est modifiée d'une manière commune dans deux traitements. Par exemple, 1949 gènes sont sous- ou sur-exprimés comparativement aux sédiments témoins (t) dans les traitements HC et HCN (« effet pétrole »). 245 gènes le sont spécifiquement par rapport au témoin lorsque des vers marins sont ajoutés aux sédiments et améliorent la bioturbation des sédiments (aération et remobilisation des contaminants) (Militon *et al.*, 2016).

Dans le cas d'un microorganisme, le séquençage global du génome ou du transcriptome, constitue à l'heure actuelle l'approche la plus performante pour déterminer sa position taxonomique (c'est à dire sa place dans l'arbre du vivant) et surtout, pour connaître, son métabolisme et son mode de vie. Dans le cas d'une communauté microbienne, le séquençage global métagénome ou du métatranscriptome représente l'approche la plus exhaustive pour étudier la diversité et le rôle des micro-organismes dans les écosystèmes.

L'approche globale est assez simple à mettre en œuvre techniquement. Cependant le coût des équipements nécessaires et la complexité des données produites font qu'elle est peu adaptée pour servir d'outil de diagnostic à grande échelle. Néanmoins, le séquençage global peut-être employé en première approche pour obtenir une image détaillée des communautés microbiennes dans un écosystème. Les données générées peuvent ensuite servir à trouver de nouveaux indicateurs qui seront utilisés avec les outils de l'approche ciblée.

4-Tableau de bord opérationnel pour le diagnostic environnemental

Les avancées technologiques majeures de ces 10 dernières années ont permis d'accroitre de façon exponentielle les connaissances scientifiques et techniques en microbiologie moléculaire, ce qui a favorisé le développement de nouvelles techniques et nouveaux marqueurs moléculaires et ouvert de nouvelles perspectives d'application à des fins de diagnostic environnemental (études d'impact, bioremédiation de sites et sols pollués, évaluation des pratiques agricoles, optimisation des bioprocédés de traitement des

déchets…). De fait, il existe une demande de plus en plus forte du monde industriel et agricole pour l'utilisation de ces technologies issues du monde de la recherche.

Il est toutefois important de comprendre que la notion de « technique éprouvée » diffère de la notion de « bioindicateur opérationnel ». En effet, certaines techniques standardisées resteront dans le domaine de la recherche et n'auront jamais d'applications finalisées en tant qu'indicateurs pour des raisons de coûts, de technicité trop importante ou d'interprétation de résultats réservée à des personnels qualifiés.

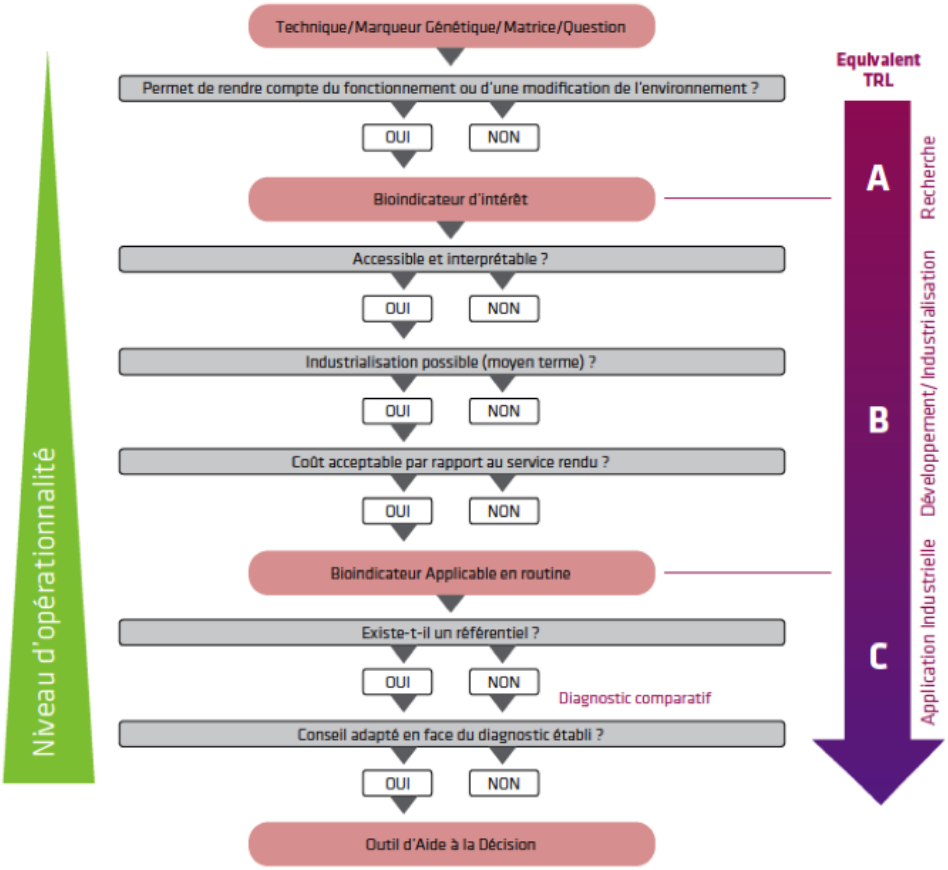
Dans le contexte de la microbiologie moléculaire environnementale, un bioindicateur correspond donc à un binôme technique/marqueur génétique (taxonomique ou fonctionnelle, plus ou moins spécifique). Il permet de détecter, quantifier ou caractériser une communauté, une population, une espèce, une fonction ou un groupe fonctionnel, dans une matrice et un contexte donnés. Choisi avec pertinence, le bioindicateur va permettre de qualifier ou quantifier l'état microbiologique d'un environnement et de renseigner, par exemple, sur l'impact d'une perturbation du système biologique étudié ou évalué.

Pour être opérationnel, un bioindicateur doit être :

- robuste, fiable, précis et spécifique : son interprétation doit être stable et cohérente dans le temps ;
- sensible : refléter même de faibles variations du milieu ou du système d'étude ;
- compréhensible, simple et utilisable par l'ensemble des acteurs ;
- en adéquation avec l'objectif fixé ;
- acceptable en termes de prix par rapport à l'information qu'il donne ;
- pertinent pour favoriser réellement la prise de décision ;
- doit permettre soit de faire suivi dans le temps et dans l'espace d'une évolution biologique de la matrice étudiée soit de caractériser la matrice par comparaison avec un référentiel (base de données).

L'échelle TRL (pour "Technology Readiness Level") permet l'évaluation le degré de maturité des technologies en matière d'innovation (figure 16). Cette échelle commence avec l'observation et la description de principes de base (niveau 1 - concept) jusqu'à l'application opérationnelle réelle (hors contexte expérimental) d'une technologie (niveau 9). Cette échelle peut être simplifiée en ne considérant que trois niveaux de maturité technologique :

- **A** : Découverte d'un marqueur génétique d'intérêt et identification de la technique adaptée à son étude - TRL 1 à 3 - recherche fondamentale ;
- **B** : Mise en œuvre du bioindicateur dans un contexte expérimental - TRL 4 à 6 : développement – standardisation - recherche appliquée ;
- **C** : Mise en œuvre du bioindicateur dans un contexte opérationnel répondant à un marché et pour lequel des prestataires sont en mesure de le proposer en routine – TRL 7 à 9 : industrialisation - prestation, mise sur le marché.



- **Figure 16 -** Logigramme d'opérationnalité d'un bioindicateur.

Sur la base de cette échelle TRL, il est possible de positionner les différents indicateurs présentés dans la section III selon leur niveau d'opérationnalité (tableau 1). Il existe une très importante variabilité des propriétés des matrices environnementales entre le sol, les sédiments, l'eau, les déchets et l'atmosphère (variabilité spatiale, hétérogénéité, densité cellulaire,…). Il en résulte que les contraintes techniques pour le développement des outils moléculaires dépendent de la matrice étudiée. De plus, l'opérationnalité n'est pas définie seulement sur la base de la maturité technique de la méthode moléculaire utilisée, mais aussi sur d'autres critères essentiels comme le coût, la facilité de mise en œuvre ou encore l'existence de référentiels d'interprétation… qui dépendent également de la matrice étudiée (figure 16). En conséquence, le niveau d'opérationnalité d'un indicateur ne peut pas être défini dans l'absolu mais doit être relié à la matrice environnementale (tableau 1).

Matrice environnementale	Sol	Sédiments	Eau	Déchets / Bioprocédés	Atmosphère
Biomasse moléculaire	TRL C	TRL B	TRL B	TRL B	NO
Détection/abondance d'organismes ou de fonctions *- PCR quantitative* *- FISH* *- Biopuce ADN*	 TRL C TRL B TRL C	 TRL C TRL B TRL B	 TRL C TRL B TRL C	 TRL C TRL B TRL C	 TRL C NO TRL A
Diversité taxonomique et fonctionnelle *- Séquençage massif d'amplicons* *- Séquençage global*	TRL C TRL A	TRLB TRL A	TRLB TRL A	TRLB TRL A	TRL B TRL A

Tableau 1 - Niveau d'opérationnalité des indicateurs moléculaires microbiens en fonction de chaque matrice environnementale et selon l'échelle TRL décrite dans la figure 16. NO : Non opérationnel sur cette matrice.

Le référencement de l'opérationnalité des indicateurs présentés dans le tableau ci-dessus n'est évidemment pas figé mais représente seulement une image instantanée définie à la date de publication de cet article. L'évolution rapide des outils moléculaires, des questionnements et des connaissances dans les différentes matrices environnementales entrainera indéniablement une évolution de ces indicateurs vers une opérationnalité plus aboutie qui permettra à terme leur déploiement pour un diagnostic de la qualité de notre environnement et des différentes matrices le constituant.

5- Conclusion

L'analyse des communautés microbiennes a d'abord reposé sur des approches culturales, dites pasteuriennes, qui sont relativement longues à mettre en œuvre. Elles ne permettent donc d'étudier et de caractériser qu'un nombre limité d'échantillons. De plus, on ne sait cultiver qu'une minorité (quelques pourcents) de micro-organismes présents dans les environnements. Malgré leur intérêt, notamment pour tout ce qui concerne la recherche de pathogènes dans les aliments ou dans les eaux de baignade et l'enrichissement des bases de données avec les génomes d'organismes séquencés après isolement, la mise en place d'indicateurs opérationnels utilisant ces approches reste limitée. Grâce à l'avènement de la biologie moléculaire dans les années 1990 et, plus récemment, à la révolution des approches "omiques", plusieurs techniques moléculaires permettent de détecter, quantifier ou caractériser une communauté, une population, une espèce, une fonction ou un groupe fonctionnel microbien de manière précise et robuste. Bien qu'il existe encore plusieurs limitations à leur utilisation (*e.g.*, absence de référentiel pour certaines matrices encore peu ou pas assez étudiées comme les écosystèmes océaniques ; pour certaines techniques, volume et complexité des données générées ; biais d'extraction, mesure des populations dominantes), elles offrent la possibilité de suivre et de caractériser finement le patrimoine microbiologique et même l'état fonctionnel de différentes matrices environnementales.

Plusieurs bioindicateurs tels que la biomasse moléculaire microbienne, la détection et le dénombrement *in situ* de certains organismes ou gènes de fonction microbiens, la diversité taxonomique microbienne, la diversité fonctionnelle microbienne ou encore la composition taxonomique des communautés microbiennes peuvent ainsi être définis. Appliqués au diagnostic environnemental et du fait du rôle majeur joué par les microorganismes, ils permettent de contribuer à déterminer l'état d'un système biologique mais aussi à suivre son évolution temporelle et/ou spatiale suite à une perturbation et/ou une réhabilitation, ou encore à évaluer l'efficacité des bioprocédés.

Il est toutefois important de ne pas oublier que la qualité des informations obtenues dépendra, (i) de la pertinence de la stratégie d'échantillonnage et du soin apporté aux échantillons à analyser (conditionnement, transport, manipulation, stockage) et, (ii) des protocoles de bioinformatiques utilisées (choix des banques de données, des approches de traitement et d'analyse statistiques des jeux de données obtenus). La bioinformatique est une discipline en plein essor dont beaucoup de protocoles analytiques n'ont pas encore étaient standardisés. Le traitement des informations obtenues, particulièrement en ce qui concerne le séquençage NGS, devra donc être confié à un expert en la matière (prestataire).

Malgré son fort potentiel, la microbiologie moléculaire reste encore peu exploitée dans le domaine de la bioindication et de l'écotoxicologie. Nul doute toutefois que le diagnostic environnemental, basé sur l'utilisation de bioindicateurs microbiens de plus en plus opérationnels, est appelé à fortement se développer dans les années à venir. Au-delà du diagnostic, ces bioindicateurs permettront à court terme la mise en place d'un véritable conseil indispensable pour une meilleur gestion et durabilité des écosystèmes.

6- Références

Bailly J., Fraissinet-Tachet L., Verner M.C., Debaud J.C., Lemaire M., Wesolowski-Louvel M. et Marmeisse R., 2007 - Soil eukaryotic functional diversity, a metatranscriptomic approach. ISME J, 1, 7, pp. 632–42.

Baker B.J., Sheik C.S., Taylor C.A., Jain S., Bhasi A., Cavalcoli J.D. et Dick G.J., 2013 - Community transcriptomic assembly reveals microbes that contribute to deep-sea carbon and nitrogen cycling. ISME J, 7, 10, pp. 1962-1973.

Barrios E., 2007 - Soil biota, ecosystem services and land productivity. Ecol Econ, 64, 2, pp. 269-285.

Bertin P.N., Heinrich-Salmeron A., Pelletier E., Goulhen-Chollet F., Arsène-Ploetze F., Gallien S., *et al.* , 2011 - Metabolic diversity among main microorganisms inside an arsenic-rich ecosystem revealed by meta- and proteo-genomics. ISME J, 5, 11, pp. 1735-1747.

Bouchez T., Blieux AL, Dequiedt S., Domaizon I., Dufresne A., Ferreira S., Godon JJ., Hellal J., Joulian C., Quaiser A., Martin-Laurent F., Mauffret A., Monier JM., Peyret P., Schmitt-Koplin P., Sibourg O., D'oiron E., Bispo A., Deportes I., Grand C., Cuny P., Maron PA. et Ranjard L., 2016 - Molecular microbiology methods for environmental diagnosis. Environmental Chemistry Letters, 14, 4, pp. 423–441.

Brodie E.L., DeSantis T.Z., Joyner D.C., Baek S.M., Larsen J.T., Andersen G.L., Hazen T.C., Richardson P.M., Herman D.J., Tokunaga T.K., Wan J.M. et Firestone M.K., 2006 - Application of a High-Density Oligonucleotide Microarray Approach To Study Bacterial Population Dynamics during Uranium Reduction and Reoxidation. Appl Environ Microbiol, 72, 9, pp. 6288-6298.

Buitenhuis E.T., Hashioka T. et Le Quéré C., 2013 - Combined constraints on global ocean primary production using observations and models. Global Biogeochem Cycles, 27, pp.

Chaparro J.M., Badri D.V. et Vivanco J.M., 2014 - Rhizosphere microbiome assemblage is affected by plant development. ISME J, 8, 4, pp. 790–803.

Delmont T.O., Robe P., Clark I., Simonet P. et Vogel T., 2011 - Metagenomic comparison of direct and indirect soil DNA extraction approaches. J Microbiol Methods, 86, 3, pp. 397-400.

Dequiedt S., Saby N.P.A., Lelievre M., Jolivet C., Thioulouse J., Toutain B., Arrouays D., Bispo A., Lemanceau P. et Ranjard L., 2011 - Biogeographical Patterns of Soil Molecular Microbial Biomass as Influenced by Soil Characteristics and Management. Global Ecol Biogeogr, 20, 4, pp. 641-652.

Dugat-Bony E., Peyretaillade E., Parisot N., Biderre-Petit C., Jaziri F., Hill D., Rimour S. et Peyret P., 2012 - Detecting unknown sequences with DNA microarrays: explorative probe design strategies. Environ Microbiol, 14, 2, pp. 356-371.

Durban N., Juzan L., Krier J. et Gillot S. 2016 - Control of Microthrix parvicella by aluminium salts addition. Water Sci Technol, 73, 2, pp. 414-22.

Dumont M.G., Pommerenke B. et Casper P., 2013 - Using stable isotope probing to obtain a targeted metatranscriptome of aerobic methanotrophs in lake sediment. Environ Microbiol Rep, 5, pp. 757–64.

Hayatsu M., Tago K. et Saito M., 2008 - Various players in the nitrogen cycle: Diversity and functions of the microorganisms involved in nitrification and denitrification. Soil Sci Plant Nutr Soil, 54, 1, pp. 33-45.

Horrigue W., Dequiedt S., Chemidlin Prévost-Bouré N., Jolivet C., Saby N., Arrouays D., Bispo A., Maron P.A. et Ranjard L., 2016 - Predictive Model of Soil Molecular Microbial Biomass. Ecol Indic, 64, pp. 203-211.

Fiehn O., Kopka J., Dörmann P., Altmann T., Trethewey R.N. et Willmitzer L. 2000 - Metabolite profiling for plant functional genomics. Nature biotechnol, 18, 11, pp. 1157–1161.

Figeys D. 2000 - The Achilles' heel of proteomics. Trends Biotechnol 18, 12, pp. 483.

Freedman D.L. et Gossett J.M., 1989 - Biological reductive dechlorination of tetrachloroethylene and trichloroethylene to ethylene under methanogenic conditions. Appl Environ Microbiol, 55, 9, pp. 2144-2151.

Frias-Lopez J., Shi Y., Tyson G.W., Coleman M.L., Schuster S.C., Chisholm S.W. et Delong E.F. 2008 - Microbial community gene expression in ocean surface waters. Proc Natl Acad Sci USA 105, 10, pp. 3805–3810.

Geisen S., Tveit A., Clark I., Richter A., Svenning M.M., Bonkowski M. et Urich T., 2015 - Metatranscriptomic census of active protists in soils. ISME J, 9, 10, pp. 2178-2190.

Ghai R., Mizunon C.M., Picazo A., Camacho A. et Rodriguez-Valera F., 2014 - Key roles for freshwater Actinobacteria revealed by deep metagenomic sequencing. Mol Ecol, 23, 24, pp. 6073–6090.

Gifford S.M., Sharma S. et Moran M.A., 2014 - Linking activity and function to ecosystem dynamics in a coastal bacterioplankton community. Front Microbiol, 5, pp. 185.

Gilbert J., Field D., Huang Y., Edwards R., Li W., Gilna P. et Joint I., 2008 - Detection of large numbers of novel sequences in the metatranscriptomes of complex marine microbial communities. PLoS One, e3042. doi:10.1371/journal.pone.0003042

Joux F., Bertrand J.C., De Wit R., Grossi V., Intertaglia L., Lebaron P., Michotey V., Normand P., Peyret P., Raimbault .P, Tamburini C. et Urios L. 2011 - Les biopuces en écologie microbienne. Méthodes d'études des micro-organismes dans l'environnement. Chapitre 17. In: Bertrand JC, Caumette P, Lebaron P, Matheron R, Normand P (eds) Ecologie Microbienne. Microbiologie des milieux naturels et anthropisés, Presses Universitaires de Pau et des Pays de l'Adour, Pau, pp 799-876.

Keiblinger K.M., Wilhartitz I.C., Schneider T., Roschitzki B., Schmid E., Eberl L., Riedel K. et Zechmeister-Boltenstern S., 2012 - Soil metaproteomics - comparative evaluation of protein extraction protocols. Soil Biol Biochem, 54, pp. 14-24.

Lankadurai B.P., Nagato E.G. et Simpson M.J., 2013 - Environmental metabolomics: an emerging approach to study organism responses to environmental stressors. Environ Rev, 21, 13, pp. 180-205.

Leary D.H., Hervey IV W.J., Deschamps J.R., Kusterbeck A.W. et Vora G.J., 2013 - Which metaproteome? The impact of protein extraction bias on metaproteomic analyses. Mol Cell Probes, 27, 5-6, pp. 193-9.

Lesniewski R.A., Jain S., Anantharaman K., Schloss P.D. et Dick G.J., 2012 - The metatranscriptome of a deep-sea hydrothermal plume is dominated by water column methanotrophs and lithotrophs, ISME J, 6, 12, pp. 2257–2268.

London R.E. et Houck D.R., 2004 - Introduction to metabolomics and metabolic profiling. In: Hamadeh HK, Afshari CA (eds) Toxicogenomics: Principles and Applications, John Wiley & Sons, New York, pp 299–340.

Lucio M., 2009 - Datamining Metabolomics: The Convergence Point of Non-target Approach and Statistical Investigation, 189 pp.

Mao Y., Xia Y., Wang Z. et Zhang T., 2014 - Reconstructing a Thauera genome from a hydrogenotrophic-denitrifying consortium using metagenomic sequence data. Appl Microbiol Biotechnol, 98, 15, pp. 6885-95.

Maron P.A., Ranjard L., Mougel C. et Lemanceau P., 2007 - Metaproteomics: a new approach for studying functional microbial ecology. Microbial Ecol, 53, 3, pp. 486-493.

Maron P.A., Mougel C. et Ranjard L., 2011 - Soil microbial diversity: spatial overview, driving factors and functional interest. CRAS Biology II, 334, 5-6, pp. 403-411.

McCarren J., Becker J.W., Repeta D.J., Shi Y., Young C.R., Malmstrom R.R., Chisholm S.W. et DeLong E.F., 2010 - Microbial community transcriptomes reveal microbes and metabolic pathways associated with dissolved organic matter turnover in the sea. Proc Natl Acad Sci, 107, 38, pp.16420-16427.

Militon C., Atkinson A., Michotey V., Jeziorski C., Cravo-Laureau C., Durand R., Bonin P. et Cuny P., 2016 - Metatranscriptomes of marine coastal sediment affected by oil addition and/or by the bioturbating activity of the marine polychaete *Hediste diversicolor*: who are the microbial players? Mar Genomics, 29, pp- 55-59.

Mondav R., Woodcroft B.J., Kim E.H., McCalley C.K., Hodgkins S.B., Crill P.M., Chanton J., Hurst G.B., VerBerkmoes N.C., Saleska S.R., Hugenholtz P., Rich V.I. et Tyson G.W., 2014 - Discovery of a novel methanogen prevalent in thawing permafrost. Nat commun, 5, pp. 3212.

Mushtag M.Y., Choi Y.H., Verpoorte R. et Wilson E.G., 2013 - Extraction for Metabolomics: Access to The Metabolome. Phytochem Anal, 25,4, pp. 291-306.

Nicholson J. K., Lindon J.C. et Holmes E., 1999 - 'Metabonomics': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. Xenobiotica, 29,11, pp. 1181–1189.

Nielsen K.M., Johnsen P.J., Bensasson D. et Daffonchio D., 2007 - Release and persistence of extracellular DNA in the environment. Environ Biosafety Res, 6, 1-2, pp. 37–53.

Ottesen E., Young C.R., Eppley J.M., Ryan J.P., Chavez F.P., Scholin C.A. et DeLonga E.F., 2013 - Pattern and synchrony of gene expression among sympatric marine microbial populations. Proc Natl Acad Sci U S A, 110, 6, E488–E497.

Pelletier E., Kreimeyer A., Bocs S., Rouy Z., Gyapay G., Chouari R., Rivière D., Ganesan A., Daegelen P., Sghir A., Cohen G.N., Médigue C., Weissenbach J. et Le Paslier D., 2008 - "Candidatus Cloacamonas acidaminovorans":genome sequence reconstruction provides a first glimpse of a new bacterial division. J Bacteriol, 190,7, pp. 2572-2579.

Poretsky R.S., Bano N., Buchan A., LeCleir G., Kleikemper J., Pickering M., Pate W.M., Moran M.A. et Hollibaugh J.T., 2005 - Analysis of microbial gene transcripts in environmental samples. Appl Environ Microbiol, 71, 7, pp. 4121-6.

Poretsky R.S., Hewson I., Sun S., Allen A.E., Zehr J.P. et Moran M.A., 2009 - Comparative day/night metatranscriptomic analysis of microbial communities in the North Pacific subtropical gyre. Environ Microbiol, 11, 6, pp. 1358–1375.

Pulleman M., Creamer R., Hamer U., Helder J., Pelosi C., Pérès G. et Rutgers M., 2012 - Soil biodiversity, biological indicators and soil ecosystem services-an overviewof European approaches. Curr. Opin Environ Sustain, 4, 5, pp. 529-538.

Quaiser A., Bodi X., Dufresne A., Naquin D., Francez A.J., Dheilly A., Coudouel S., Pédrot M. et Vandenkoornhuyse P., 2014 - Unraveling the stratification of an iron-oxidizing microbial mat by metatranscriptomics. PLoS One, 9, 7, e102561.

Raes J. et Bork P., 2008 - Molecular eco-systems biology: towards an understanding of community function, Nat Rev Microbiol, 6, 9, pp. 693-699.

Ramachandran N., Hainsworth E., Bhullar B., Eisenstein S., Rosen B., Lau A.Y., Walter J.C. et LaBaer J., 2004 - Self-assembling protein microarrays. Science, 305, 5680, pp. 86-90.

Rames E.K., Smith M.K., Hamill S.D. et De Faveri J., 2013 - Microbial indictors related to yield and disease and changes in soil microbial community structure withginger farm management practices. Aust Plant Pathol 42, 6, pp. 685–692.

Maier R.M. et Gentry T.J., 2015 - Microorganisms and Organic Pollutants. *In* Environmental Microbiology (Third Edition) [0-12-394626-3], pp. 377-413

Ritz K., Black H.I.J., Campbell C.D., Harris J.A. et Wood C., 2009 - Selecting indicators for monitoring soils:a framework for balancing scientific and technical opiniont o assist policy development. Ecol Ind, 9, 6, 1212-1221.

Roane T.M., Pepper I.L. et Gentry Terry J., 2015 - Microorganisms and Metal Pollutants. *In* Environmental Microbiology (Third Edition) [0-12-394626-3], pp.415-439.

Savichtcheva O., Debroas D., Kurmayer R., Villar C., Jenny J.P., Arnaud F., Perga M.E. et Domaizon I., 2011 - Quantitative PCR enumeration of total/toxic Plankthotrix rubescens and total Cyanobacteria in preserved DNA isolated from lake sediments. Applied Env. Microbiol, 77, 24, pp. 8744-8752.

Sharma S.K., Ramesh A., Sharma M.P., Om Joshi O.P., Govaerts B., Steenwerth K.L. et Karlen DL., 2011 - Microbial community structure and diversity as indicators forevaluating soil quality. In: Lichtfouse E (ed) Biodiversity, Biofuels, Agroforestryand Conservation Agriculture, Springer, pp. 317-358.

Terrat S., Plassart P., Bourgeois E., Ferreira S., Dequiedt S., Adele-Dit-De-Renseville A., Lemanceau P., Bispo A., Chabbi A., Maron P.A. et Ranjard L., 2015 - Meta-barcoded evaluation of the ISO standard 11063 DNA extraction procedure to characterize soil bacterial and fungal community diversity and composition. Microbial Biotechnolog, 8, 1, pp. 131-142.

Tseng C.H. et Tang S.L., 2014 - Marine Microbial Metagenomics: From Individual to the Environment. Int J Mol Sci, 15, 5, pp. 8878-8892.

Turner T.R., Ramakrishnan K., Walshaw J., Heavens D., Alston M., Swarbreck D., Osbourn A., Grant A. et Poole P.S., 2013 - Comparative metatranscriptomics reveals kingdom level changes in the rhizosphere microbiome of plants. ISME J, 7,12, pp. 2248–58.

Tveit A., Schwacke R., Svenning M.M. et Urich T., 2013 - Organic carbon transformations in high-Arctic peat soils: key functions and microorganisms. ISME J, 7, 2, pp. 299-311.

Urich T., Lanzén A., Qi J., Huson D.H., Schleper C. et Schuster S.C., 2008 - Simultaneous Assessment of Soil Microbial Community Structure and Function through Analysis of the Meta-Transcriptome. PLoS One, 3, 6, e2527.

Vila-Costa M., Sharma S., Moran M.A. et Casamayor E.O., 2013 - Diel gene expression profiles of a phosphorus limited mountain lake using metatranscriptomics. Environ Microbiol, 15, 4, pp. 1190-1203.

Wilmes P. et Bond P.L., 2006 - Metaproteomics:studying functional gene expression in microbial ecosystems. Trends Microbiol, 14, 2, pp. 92-97.

Yates J.R. 3rd, Speicher S., Griffin P.R. et Hunkapiller T., 1993 - Peptide mass maps: a highly informative approach to protein identification. Anal. Biochem, 214, 2, pp. 397-408.

Zhou J., He Z., Yang Y., Deng Y., Tringe S.G. et Alvarez-Cohen L., 2015 - High-throughput metagenomic technologies for complex microbial community analysis: open and closed formats. MBio, 27, 6, 1, e02288-14. doi:10.1128/mBio.02288-14

GLOSSAIRE

- Acides nucléiques : macromolécules d'origine biologique dont l'unité de base est le nucléotide.
- ADEME : Agence de l'Environnement et la Maitrise de l'Energie.
- ADN : l'acide désoxyribonucléique, ou ADN, est une macromolécule biologique présente dans toutes les cellules ainsi que chez de nombreux virus. L'ADN contient toute l'information génétique, appelée génome, permettant le développement et le fonctionnement des êtres vivants.
- ADN complémentaire : L'ADN complémentaire (ou ADNc) est un simple brin artificiellement synthétisé à partir d'un ARNm, représentant ainsi la partie codante de la région du génome ayant été transcrit en cet ARN. L'acide ribonucléique messager, ARN messager ou ARNm est une copie transitoire d'une portion de l'ADN correspondant à un ou plusieurs gènes. L'ARNm est utilisé comme intermédiaire par les cellules pour la synthèse des protéines.
- ADN polymerase : complexe enzymatique intervenant dans la réplication de l'ADN au cours du cycle cellulaire, mais aussi dans des processus de réparation et de recombinaison de l'ADN.
- Amplicons : produits obtenus après une réaction d'amplification génique (PCR) ; Fragment d'ADN à amplifier.
- Analyse en composante principale : méthode de la famille de l'analyse des données et plus généralement de la statistique multivariée, qui consiste à transformer des variables liées entre elles (dites "corrélées" en statistique) en nouvelles variables décorrélées les unes des autres. Ces nouvelles variables sont nommées "composantes principales", ou axes principaux. Elle permet au praticien de réduire le nombre de variables et de rendre l'information moins redondante.
- Analyse multi-tableaux : ou analyse en tableaux ou multi-blocs, approche statistique permettant d'analyser des données provenant de différentes sources de mesures ou techniques analytiques sur les mêmes échantillons.
- Aquifères : nappe phréatique de faible profondeur.
- ARN : l'acide ribonucléique est présent chez pratiquement tous les êtres vivants, et est très proche chimiquement de l'ADN. Il est synthétisé dans les cellules à partir d'une matrice d'ADN. Les cellules vivantes utilisent en particulier l'ARN comme un support intermédiaire des gènes pour synthétiser les protéines dont elles ont besoin.
- ARNm : l'ARNm est une copie simple brin linéaire de l'ADN composée d'ARN, qui comprend la région codant une protéine, encadrée de régions non codantes. Il est synthétisé sous forme de précurseur dans le noyau de la cellule lors d'un processus appelé transcription.
- Bioindicateur : indicateur constitué par une espèce végétale, microbienne ou animale ou par un groupe d'espèces ou communautés dont la présence ou l'état renseigne sur certaines caractéristiques écologiques (c'est-à-dire physico-chimiques, microclimatique, biologiques et fonctionnelle) de l'environnement ou sur l'incidence de certaines pratiques.
- Biomasse microbienne : quantité de microorganismes dans un environnement ou une matrice naturelle évaluée par la mesure du carbone vivant ou plus récemment par la mesure de la quantité d'ADN (biomasse moléculaire microbienne).
- Biopuces : ou puce à ADN, est un ensemble de molécules d'ADN fixées en rangées ordonnées sur une petite surface qui peut être du verre, du silicium ou du plastique. Cette biotechnologie récente permet d'analyser par hybridation la présence ou le niveau d'expression des gènes dans un organisme ou une communauté d'organismes issue d'une matrice naturelle, à un moment donné et dans un état donné par rapport à un échantillon de référence.
- Biosphère : ensemble d'organismes dans les matrices naturelles.
- Chimiomètrie : application des outils mathématiques, en particulier statistiques, pour obtenir le maximum d'informations à partir des données chimiques.
- Chromatographie : méthode physico-chimique de séparation des molécules présentes dans un échantillon en phase homogène liquide ou gazeuse.
- CNRS : Centre National de la Recherche Scientifique
- Communauté : ensemble de populations d'organismes vivants au sein d'un même écosystème.
- Composition taxonomique : Composition d'une communauté d'organismes vivants basée sur le regroupement de ces individus en taxons.
- Contaminant : Toute substance organique ou inorganique décelée dans un environnement (sol, air, eau, sédiments, ...) ou elle ne se trouve pas normalement ou qui va s'y trouver en quantité anormale. Il peut aussi s'agir d'organismes (bactéries, virus, moisissures, ...) ou des productions d'organismes (toxines) indésirables dans un produit (par exemple des aliments) ou dans l'environnement du produit (air, eau, surface).

- Couche euphotique : couche éclairée superficielle des écosystèmes aquatiques (aussi appelée photique) où se produit la photosynthèse et dont la profondeur est définie comme étant celle où ne subsiste plus que 1% de l'énergie lumineuse qui arrive en surface. L'épaisseur de cette couche est variable et dépend de la concentration en composés dissous et sous forme particulaires (dont l'abondance en microalgues).

- Cox : cytochrome C oxidase, gène mitochondrial.

- Hybridation (ADN) : appariement entre des brins d'ADN ou entre ADN et ARN

- Dénitrification : processus microbien de respiration alternatif qui est opéré par des bactéries spécifiques, satisfaisant leur besoin en oxygène en environnement hypoxique ou anoxique, par une désoxygénation des ions nitrates.

- Diagnostic : raisonnement menant à l'identification de la cause (l'origine) d'une défaillance, d'un problème ou d'une maladie.

- Diversité fonctionnelle : chez les microorganismes elle est définie comme la diversité des gènes portant des fonctions ou la diversité des fonctions elles-mêmes à l'échelle d'un groupe ou d'une communauté d'individus.

- Diversité génétique : désigne le degré de variétés des gènes au sein d'une même espèce ou d'une communauté d'espèces.

- Écologie microbienne : domaine scientifique qui aborde la place et le rôle des microorganismes dans un habitat (environnement, écosystème) ainsi que les interactions des microorganismes entre eux et avec leur milieu.

- Écosystème : ensemble formé par une association ou communauté d'êtres vivants (ou biocénose) et son environnement biologique, géologique, édaphique, hydrologique, climatique, etc. (le biotope). L'écosystème représente aussi une unité fonctionnelle, qui intègre les interactions entre les espèces et leur milieu de vie mais aussi les rapports existant entre ces espèces.

- Écotoxicologie : discipline scientifique qui étudie le comportement et les effets d'agents "polluants" sur les écosystèmes, qu'il s'agisse d'agents d'origine artificielle (incluant médicaments, perturbateurs endocriniens, etc.) ou d'agents naturels dont l'homme modifie la répartition et/ou les cycles dans les différents compartiments de la biosphère.

- Électrophorèse : techniques utilisées en biochimie ou biologie moléculaire pour la séparation des protéines ou des acides nucléiques basées sur des différences de charge électrique et pour des charges identiques, en fonction de leur taille.

- Empreintes génétiques : caractérisation d'un microorganisme ou d'une communauté microbienne sur la base de ses différences génétiques.

- Eucaryote : domaine du vivant regroupant tous les organismes, unicellulaires ou pluricellulaires, qui se caractérisent par la présence d'un noyau et de mitochondries dans leurs cellules. Il s'oppose aux domaines des Eubacteria et des Archaea.

- Fluorescence : émission lumineuse provoquée par l'excitation d'une molécule (généralement par absorption d'un photon) immédiatement suivie d'une émission spontanée. La fluorescence peut entre autres servir à caractériser un matériau ou une molécule.

- Fluorochrome : substance chimique capable d'émettre de la lumière de fluorescence après excitation.

- Fonctionnalité : ensemble des activités potentielles portées par un microorganisme ou une communauté de microorganismes.

- Gènes fonctionnels : gènes codant pour les enzymes impliquées dans le métabolisme primaire et secondaire d'un microorganisme.

- Génome : ensemble du matériel génétique d'un individu ou d'une espèce codée dans son acide désoxyribonucléique (ADN).

- Génomique : discipline de la biologie moderne qui étudie le fonctionnement d'un organisme microbien à l'échelle du génome, au lieu de se limiter à l'échelle d'un seul gène.

- INRA : Institut National de la Recherche Agronomique

- Irstea : Institut national de recherche en sciences et technologies pour l'environnement et l'agriculture

- ISO (International Organization for Standardization) : organisme de normalisation international composé de représentants d'organisations nationales de normalisation de 165 pays.

- Tapis microbiens : biofilm d'épaisseur variable (de quelques micromètres à quelques centimètres) constitué essentiellement de microorganismes, qui peut se former dans différentes matrices environnementales (eau, sol, aliment, tissus …). Il est le lieu d'une intense activité microbienne.

- Marqueur génétique: séquence d'ADN ou d'ARN cible, permettant d'identifier des organismes (ou groupe d'organismes) ou des fonctions spécifiques.

- Mésotrophe : caractérise un milieu moyennement riche en nutriments, se situe entre les états oligotrophe (moins riche) et eutrophe (plus riches).

- Métabolomique : science très récente qui étudie l'ensemble des métabolites (sucres, acides aminés, acides gras, etc.) présents dans une cellule, un organe, un organisme. C'est l'équivalent de la génomique pour l'ADN. Elle utilise la spectrométrie de masse et la résonance magnétique nucléaire.

- Métagénomique : méthode d'étude du contenu génétique d'échantillons issus d'environnements complexes (ex : intestin, océan, sols, air, etc.). Cette approche, via l'extraction et le séquençage direct de l'ADN présent dans l'échantillon, permet une description génomique du contenu de l'échantillon, mais aussi un aperçu du potentiel fonctionnel d'un environnement. En microbiologie, elle permet d'étudier les microorganismes directement dans leur environnement sans passer par une étape de culture en laboratoire.

- Métaprotéome : ensemble des protéines contenues dans un échantillon complexe (ex : intestin, océan, sols, air, etc.)

- Métaprotéomique : méthode d'étude du contenu protéique d'échantillons issus d'environnements complexes (ex : intestin, océan, sols, air, etc.). Cette approche, via l'extraction et la caractérisation directe des protéines présentes dans l'échantillon, permet une description du potentiel fonctionnel d'un environnement. En microbiologie, elle permet d'étudier les fonctions microbiennes directement dans leur environnement sans passer par une étape de culture en laboratoire.

- Métatranscriptome : ensemble des molécules issues de la transcription de l'ADN (ARN) dans un échantillon complexe (ex : intestin, océan, sols, air, etc.).

- Microorganismes : organismes de taille généralement inférieure à 1 mm constitués de procaryotes (bactéries et archées) mais aussi d'une grande diversité d'organismes eucaryotes (microeucaryotes) appartenant à différents groupes taxonomiques.

- Minéralisation de la matière organique : processus biologique de transformation de la matière organique de l'environnement (sol, eau, déchets) en matière minérale (CO2, NH3, SO2…).

- Next Generation Sequencing (NGS) : technologie de séquençage basée sur des processus chimiques et/ou enzymatiques et permettant d'augmenter significativement les capacités de séquençage de l'ADN en comparaison avec la technique historique dite de Sanger.

- Nitrification : processus microbiologique qui transforme l'ammonium ($NH4^+$) en nitrates en deux étapes appelées respectivement nitritation et nitratation.

- Nucléotide : molécule organique qui est l'élément de base d'un acide nucléique tel que l'ADN ou l'ARN. Il est composé d'une base nucléique (ou base azotée), d'un ose à cinq atomes de carbone, dit pentose, dont l'association forme un nucléoside, et enfin de un à trois groupes phosphate.

- Oligonucléotides : courts segments d'acides nucléiques (ARN ou ADN) de quelques dizaines de nucléotides. Ils sont en général obtenus par synthèse chimique, sous forme de simple brin.

- Omique : possibilité de caractériser la diversité génétique et fonctionnelle des microorganismes dans leur ensemble, sans *a priori* et par l'analyse en haut débit des ADN (génomique), ARN (transcriptomique), protéine (protéomique) ou des métabolites (métabolomique).

- Opérationnalité : capacité d'un indicateur à être facilement utilisé en routine et interprété par des utilisateurs novices.

- Pathogène : organismes vivants induisant une maladie chez un autre organisme vivant.

- PCR (Polymerase Chain Reaction) : ou réaction en chaîne par polymérase, méthode de biologie moléculaire basée sur l'acitvité enzymatique de la Taq Polymerase qui perment l'amplification d'ADN *in vitro*.

- Pétrole brut : pétrole non transformé, issu directement de l'exploitation d'un puits de pétrole.

- pH : Le potentiel hydrogène (ou pH) est une mesure de l'activité chimique des ions hydrogènes H^+ (appelés aussi couramment protons) en solution. Plus couramment le pH mesure l'acidité ou la basicité d'une solution.

- Phylum : deuxième niveau de classification classique des espèces vivantes.

- PhyloChip : biopuce à ADN permettant l'identification rapide et simultanée de microorganismes ciblés, mais aussi une estimation du nombre de taxons présents. Plus de 8000 microorganismes -bactéries et archées- peuvent être recherchés en même temps dans des échantillons environnementaux ou des échantillons médicaux.

- Poids moléculaire : ou masse moléculaire relative, rapport de la masse réelle d'une molécule et l'unité de masse atomique (qui est par convention, le douzième de la masse de l'isotope le plus abondant du carbone, le carbone 12) exprimées en daltons (Da).

- Polluant : élément chimique, biologique ou physique, qui provoque directement ou indirectement, des effets délétères sur les organismes ou, plus généralement, un impact négatif sur tout ou partie d'un écosystème ou de l'environnement.

- Procaryotes : microorganismes généralement unicellulaires dépourvus de noyau appartenant soit au domaine (empire) des Bactéries ou des Archées.

- Protéomique : science qui étudie l'ensemble des protéines présentes dans une cellule, un organe, un organisme.

- Référentiel : accumulation de grands jeux de données obtenus à partir d'environnements variés qui permet de connaître la gamme de variation d'un indicateur et la valeur de référence pour un environnement donné.

- Rétrotranscription : réaction inverse de la transcription. C'est la synthèse d'un brin d'ADN à partir d'une matrice ARN grâce à une ADN polymérase ARN dépendante ou encore *transcriptase inverse* ou *rétrotranscriptase.*

- Ribosome : complexes ribonucléoprotéiques (c'est-à-dire composés de protéines et d'ARN) présents dans les cellules eucaryotes et procaryotes. Leur fonction est de synthétiser les protéines en décodant l'information contenue dans l'ARN messager.

- RMN (Résonance Magnétique Nucléaire) : propriété de certains noyaux atomiques possédant un spin nucléaire (par exemple ^{1}H, ^{13}C, ^{17}O, ^{19}F, ^{31}P, ^{129}Xe…), placés dans un champ magnétique. Lorsqu'ils sont soumis à un rayonnement électromagnétique (radiofréquence), le plus souvent appliqué sous forme d'impulsions, les noyaux atomiques peuvent absorber l'énergie du rayonnement puis la relâcher lors de la relaxation. L'énergie mise en jeu lors de ce phénomène de résonance correspond à une fréquence très précise, dépendant du champ magnétique et d'autres facteurs moléculaires. Ce phénomène permet donc l'observation des propriétés quantiques magnétiques des noyaux dans les phases gaz, liquide ou solide. Seuls les atomes dont les noyaux possèdent un moment magnétique donnent lieu au phénomène de résonance.

- Sédiments : dépôt meuble formé par des particules non consolidées laissé par les eaux, le vent et les autres agents d'érosion. Il peut être d'origine fluviatile, glaciaire, lacustre ou marine.

- Sensibilité : capacité d'un indicateur/outil à détecter une perturbation même minime.

- Séquençage haut débit : Nouvelle technologie de séquençage (pyroséquençage, Illumina…) qui permet de séquencer des millions de molécules d'ADN de taille moyenne (200-500 pb) en une seule réaction.

- Service écosystémique : bénéfices que les humains retirent des écosystèmes sans avoir à agir pour les obtenir. Ce services sont majoritairement supportés par la biodiversité des écosystèmes.

- Sondes : sondes ADN de petites tailles (5- 30 nucléotides) utilisées pour des réactions de PCR ou d'hybridation de fragments plus grands.

- Spectrométrie : étude expérimentale du spectre d'un phénomène physique, c'est-à-dire de sa décomposition sur une échelle d'énergie, ou toute autre grandeur se ramenant à une énergie (fréquence, longueur d'onde, etc.).

- Spectromètrie de masse : technique physique d'analyse permettant de détecter et d'identifier des molécules d'intérêts par mesure de leur masse, et de caractériser leur structure chimique.

- Taxon : élément d'une classification systématique quel que soit son rang. Cela peut être, par exemple, une espèce, un genre ou encore une famille.

- Transcription : mécanisme qui permet de « recopier » les données des gènes ADN en ARN, ce qui permet leur utilisation pour créer de la matière biologique en assemblant des acides aminés en protéines selon ce code.

- Transcriptome : ensemble des ARN issus de la transcription du génome.

- TRL : technology Readiness Level, système de mesure développé par la NASA (National Aeronotics and Space Administration) permettant de mesurer la maturité d'une technologie.

- Ubiquiste : organisme présent dans des milieux écologiques très différents, dans différentes zones géographiques.

L'ADEME EN BREF

L'Agence de l'Environnement et de la Maîtrise de l'Energie (ADEME) participe à la mise en œuvre des politiques publiques dans les domaines de l'environnement, de l'énergie et du développement durable. Elle met ses capacités d'expertise et de conseil à disposition des entreprises, des collectivités locales, des pouvoirs publics et du grand public, afin de leur permettre de progresser dans leur démarche environnementale. L'Agence aide en outre au financement de projets, de la recherche à la mise en œuvre et ce, dans les domaines suivants : la gestion des déchets, la préservation des sols, l'efficacité énergétique et les énergies renouvelables, la qualité de l'air et la lutte contre le bruit.

L'ADEME est un établissement public sous la tutelle conjointe du ministère de l'Ecologie, du Développement durable et de l'Energie, et du ministère de l'Éducation nationale, de l'Enseignement supérieur et de la Recherche.

ADEME
Agence de l'Environnement et de la Maîtrise de l'Energie

Liberté • Égalité • Fraternité
RÉPUBLIQUE FRANÇAISE

MINISTÈRE DE L'ENVIRONNEMENT, DE L'ÉNERGIE ET DE LA MER

MINISTÈRE DE L'ÉDUCATION NATIONALE, DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE

ADEME
20, avenue du Grésillé
BP 90406 | 49004 Angers Cedex 01

www.ademe.fr