



HAL
open science

Microbiome du sol

Barbara Pivato, Nicolas Chemidlin Prévost-Bouré, Philippe P. Lemanceau

► **To cite this version:**

Barbara Pivato, Nicolas Chemidlin Prévost-Bouré, Philippe P. Lemanceau. Microbiome du sol. La métagénomique. Développements et futures applications, Editions Quae, 120 p., 2015, Savoir Faire (Quae), 978-2-7592-2293-3 ISSN : 1952-1251. hal-02801832

HAL Id: hal-02801832

<https://hal.inrae.fr/hal-02801832>

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.

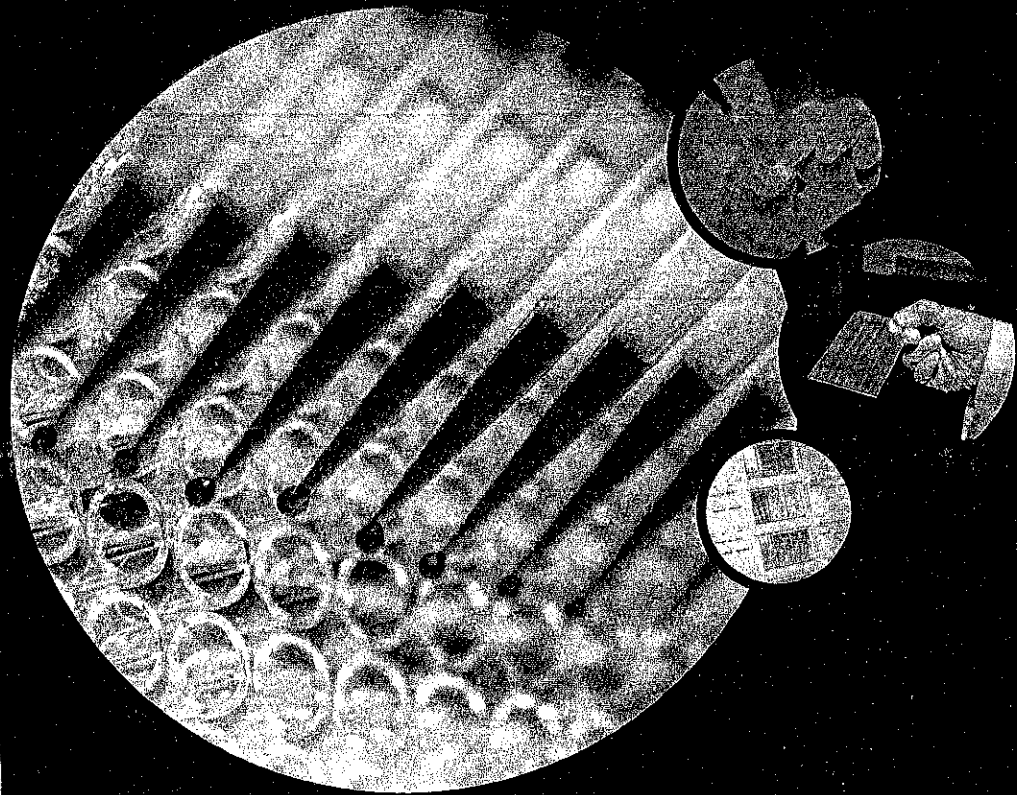
L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Savoir
faire

La métagénomique

Développements et futures applications

M.-C. Champomier-Vergès, M. Zagorec, coord.



éditions
Quæ

Après l'ère de la génomique, la métagénomique a envahi le domaine de la biologie. Permettant l'exploration du contenu génétique d'un échantillon issu d'un environnement complexe (sol, eau, végétal, microbiote animal, aliment, etc.), via le séquençage de l'ADN, la métagénomique repousse les limites imposées par la culture en milieu de laboratoire. Elle se caractérise par une production rapide et massive de données, et des traitements informatiques conséquents afin d'en retirer une information exploitable.

Cet ouvrage fait le point sur les différentes technologies et les méthodes de production et d'analyses de données actuellement disponibles dans ce domaine. Il présente un panel des champs d'applications — agriculture, environnement, agroalimentaire et santé — qui bénéficient de l'apport de ces techniques pour mieux décrypter la diversité du monde vivant microbiologique.

Il permettra aux étudiants, enseignants, chercheurs ou ingénieurs, quelle que soit leur spécialité en sciences de la vie ou de la Terre, de comprendre ces nouvelles approches, leurs mises en œuvre et leurs limites.

Marie Christine Champomier-Vergès, docteur en microbiologie, est directrice de recherche à l'Inra. Au sein de l'unité Microbiologie de l'alimentation au service de la santé (Micalis), elle dirige une équipe dont les travaux, portant sur l'analyse de l'écosystème des produits carnés par la métagénomique et la métatranscriptomique, sont tournés vers l'innovation pour le développement des approches de bioconservation de ces produits.

Monique Zagorec, docteur en génétique moléculaire, est directrice de recherche à l'Inra dans l'unité Sécurité des aliments et microbiologie (Secalim). Après avoir travaillé sur la bactérie lactique emblématique de la viande, *Lactobacillus sakei*, ses recherches portent aujourd'hui sur les communautés bactériennes des aliments carnés, en particulier la viande de volaille afin d'en améliorer la salubrité et la sécurité microbiologique.

En couverture : Préparation d'échantillon de fèces humain pour analyse du microbiote. Plaquette utilisée dans les robots de séquençage et vue d'un écran de traitement informatique de données biologiques dans un laboratoire de la plateforme de métagénomique quantitative MetaQuant (MetaGenoPolis et Micalis). © B. Nicolas/Inra.

éditions
Quæ

Éditions Cirad, Ifremer, Inra, Irstea
www.quae.com

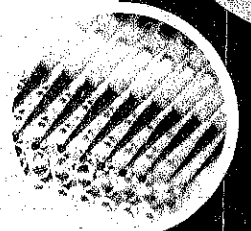
27 €

ISBN : 978-2-7592-2293-3



9 782759 222933

ISSN : 1952-1251
Réf. : 02474



Collection *Savoir-faire*

Faut-il travailler le sol ?
Acquis et innovations pour une agriculture durable
F. Laurent, J. Roger-Estrade, J. Labreuche
2014, 192 p.

Les clémentiniers et autres petits agrumes
C. Jacquemond, F. Curk, M. Heuzet
2014, 368 p.

Torrents et rivières de montagne.
Dynamique et aménagement
A. Recking, D. Richard, G. Degoutte, coord.
2013, 352 p.

Qualité du cacao
L'impact du traitement post-récolte
M. Barel
2013, 104 p.

Analyse de sensibilité et exploration de modèles
Application aux sciences de la nature et de l'environnement
R. Faivre, B. Looss, S. Mahévas, D. Makowski, H. Monod, éd.
2013, 352 p.

De la domestication à la transgénèse
Évolution des outils pour l'amélioration des plantes
A. Gallais
2013, 176 p.

Éditions Quæ
RD 10, 78026 Versailles Cedex, France

© Éditions Quæ, 2015

ISBN 978-2-7592-2293-3

ISSN 1952-1251

Le Code de la propriété intellectuelle interdit la photocopie à usage collectif sans autorisation des ayants droit. Le non-respect de cette disposition met en danger l'édition, notamment scientifique, et est sanctionné pénalement. Toute reproduction, même partielle, du présent ouvrage est interdite sans autorisation du Centre français d'exploitation du droit de copie (CFC), 20 rue des Grands-Augustins, Paris 6^e.

Avant-pr

Chapitre
des donn

Introduct

Design ex

Préparati

Le séquen

Avanta
des tec

Les fut

Les banqu

Conclusio

Références

Chapitre 2
en inform

Introducti

Assemblage

Assignatio

Classificati

Prédiction

Analyse fon

Intégration

Références

Sommaire

Avant-propos	7
Chapitre 1. Techniques et matériels pour la production des données brutes de séquençage métagénomique	9
Introduction.....	9
Design expérimental et échantillonnage.....	10
Préparation des échantillons et extraction des acides nucléiques.....	12
Le séquençage à haut débit	15
Avantages et désavantages pour la métagénomique des technologies NGS les plus populaires	16
Les futures technologies.....	19
Les banques de clones d'ADN métagénomique	19
Conclusion.....	20
Références bibliographiques	20
Chapitre 2. Analyse des données, logiciels, transformation des data en information, utilisation et modélisation des données	25
Introduction.....	25
Assemblage.....	26
Assignation taxonomique	26
Classification non supervisée	29
Prédiction de parties codantes	30
Analyse fonctionnelle	31
Intégration des outils.....	32
Références bibliographiques	32

Chapitre 3. Découverte de nouvelles fonctions et familles protéiques : nouveaux défis pour les biotechnologies et l'écologie microbienne.....	35
Introduction.....	35
Stratégies d'échantillonnage.....	35
Le criblage fonctionnel : nouveaux défis pour la découverte de fonctions..	36
La séquence, témoin de l'originalité.....	36
Le criblage d'activité : accélérateur de découverte d'outils biotechnologiques.....	40
Conclusion.....	43
Références bibliographiques	43
Chapitre 4. Microbiote intestinal et typage.....	51
Introduction.....	51
Métagénome du microbiote intestinal humain	52
Dysfonctions du microbiome intestinal et pathologies.....	54
Conclusion.....	54
Références bibliographiques	55
Chapitre 5. Communautés microbiennes des aliments.....	57
Introduction.....	57
Les techniques à haut débit ciblées pour approfondir la connaissance des communautés microbiennes des aliments	58
Les produits fermentés, une grande diversité de denrées et de communautés microbiennes	59
En occident : la chaîne alimentaire, la viande, le pain, le vin, le fromage....	60
Exemples d'études métagénomiques sur des aliments.....	61
Conclusion et perspectives.....	66
Références bibliographiques	67
Chapitre 6. Microbiome du sol.....	71
Introduction.....	71
Exploration et exploitation des ressources génétiques de sols modèles.....	76
Biogéographie et gestion des communautés microbiennes des sols.....	79
Conclusion et perspectives.....	83
Références bibliographiques	84

Chapitre
Des gènes
Commun
L'exemple
Conclusio
Références

Chapitre
Introducti
La phyllos
Premières
Une mine
Limites de
Promesses
Références

Synthèse
L'apport d
La caracté
Les différe
Les limite
Les défis f

Liste des

Chapitre 7. Métagénomique environnementale	89
Des gènes aux génomes	89
Communautés microbiennes complexes	90
L'exemple des écosystèmes arséniés	92
Conclusion et perspectives	93
Références bibliographiques	94
Chapitre 8. Microbiomes de la phyllosphère	97
Introduction	97
La phyllosphère, un habitat extrême pour les micro-organismes	97
Premières études métagénomiques	100
Une mine de données génomiques à explorer	103
Limites de l'approche métagénomique	104
Promesses de l'étude métagénomique des microbiomes de phyllosphère ...	104
Références bibliographiques	106
Synthèse et perspectives	109
L'apport de l'évolution de la technologie de séquençage	109
La caractérisation des écosystèmes de domaines très variés	110
Les différentes facettes de l'exploration du vivant par la métagénomique ...	111
Les limites de la métagénomique	112
Les défis futurs	114
Liste des auteurs	115

Avant-propos

Le terme de métagénomique a été introduit pour la première fois en 1998 par Handelsman et ses collaborateurs¹, même si des études que l'on qualifierait aujourd'hui d'études métagénomiques avaient été publiées dès 1996 par Stein et ses collaborateurs². Ce terme associe le mot « méta » (du grec, signifiant transcendant) au mot génomique, la discipline qui analyse un organisme (un ensemble, une entité vivante) au niveau de son génome. Le génome quant à lui est l'ensemble du matériel génétique, donc des gènes, d'un organisme. La métagénomique est bien effectivement une analyse qui transcende — se situe au-dessus de — l'analyse d'un ensemble de gènes. C'est surtout dans le domaine de la microbiologie que se distinguent de nombreux travaux faisant appel à cette nouvelle discipline. À ce jour, presque tous les éléments présents sur Terre sont étudiés par métagénomique. L'environnement, les sols, les nuages, l'air, l'eau des lacs et des océans, ainsi que les communautés microbiennes associées au monde végétal (la rhizosphère, la phylosphère) et animal (les microbiotes hébergés par les animaux) font l'objet d'études métagénomiques. La communauté scientifique des microbiologistes s'est emparée de cette discipline. Ainsi le sol et l'eau (les océans) ont été parmi les premiers éléments à être investigués par métagénomique. En effet, les limites des méthodes classiques de microbiologie pour décrire les communautés bactériennes extrêmement riches de ces domaines étaient connues. Pour exemple, on estimait que moins de 1 % des bactéries du sol étaient cultivables en milieu de laboratoire. Les microbiotes associés au règne animal et en particulier ceux portés par l'homme ont également bénéficié de la métagénomique car plus de 70 % des bactéries hébergées dans le tractus digestif humain n'étaient pas non plus cultivables avec les milieux et conditions de laboratoire disponibles. L'analyse par métagénomique d'autres écosystèmes, tels que ceux des aliments dont les communautés microbiennes étaient mieux connues, a été plus tardive. Elle a cependant révélé des éléments nouveaux, notamment la présence d'espèces déjà connues mais que l'on considérait comme

1. Handelsman J., Rondon M.R., Brady S.F., Clardy J., Goodman R.M., 1998. Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. *Chem. Biol.*, 5, R245-R249.

2. Stein J.L., Marsh T.L., Wu K.Y., Shizuya H., DeLong E.F., 1996. Characterization of uncultivated prokaryotes: isolation and analysis of a 40-Kilobase-pair genome fragment from a planktonic marine Archaeon. *J. Bacteriol.*, 178 (3), 591-599.

des espèces environnementales et qui n'avaient pas été recherchées dans les aliments. De fait, les aliments véhiculent des communautés microbiennes provenant de différents environnements (sols, eaux, végétaux, microbiotes animaux, etc.). Ces communautés microbiennes des aliments, une fois ingérées, peuvent transiter voire coloniser durablement le tractus digestif humain après consommation. Ainsi, les frontières entre les communautés microbiennes des mondes vivants (animal, végétal) et minéral (sol, air, eau) commencent à tomber, nous faisant prendre conscience de l'omniprésence du monde bactérien dans notre quotidien.

La métagénomique permet une profondeur d'analyse qui n'était pas accessible auparavant. En effet, les techniques de séquençage ont atteint un tel débit que des espèces bactériennes minoritaires peuvent être détectées maintenant parmi des centaines de milliers ou des millions de lectures alors que les approches de microbiologie classique ou même de criblage ne le permettaient pas. Une quantification relative des différentes communautés bactériennes est également possible. Enfin, grâce au déluge de séquences génomiques bactériennes désormais disponibles, les données de métagénomique peuvent aboutir à l'identification des communautés microbiennes au niveau du genre et même de l'espèce.

Avec la même logique et des techniques similaires, la métatranscriptomique a vu le jour. Elle permet de connaître les gènes exprimés ainsi qu'une quantification relative de leur niveau d'expression. Ainsi, non seulement les communautés microbiennes mais également les fonctions qu'elles expriment potentiellement peuvent être connues.

La métagénomique est encore une discipline jeune. On lui reproche souvent de n'être que descriptive. Cependant, elle permet une description bien plus exhaustive qu'auparavant et des analyses approfondies allant au-delà de la description commencent à voir le jour.

bru

Intro

Au
dé
de
de
va
fér
le
tec
d'i
en
de
O
d'i
foi
est
ce
gé
bl
pe
de
téc
sé

- Park E.J., Chun J., Cha C.J., Park W.S., Jeon C.O., Bae J.W., 2012. Bacterial community analysis during fermentation of ten representative kinds of kimchi with barcoded pyrosequencing. *Food Microbiol.* 30, 197-204.
- Penacho V., Valero E., Gonzalez R., 2012. Transcription profiling of sparkling wine second fermentation. *Int. J. Food Microbiol.*, 153, 176-182.
- Roh S.W., Kim K.H., Nam Y.D., Chang H.W., Park E.J., Bae J.W., 2010. Investigation of archaeal and bacterial diversity in fermented seafood using barcoded pyrosequencing. *ISME J.*, 4, 1-16.
- Quigley L., O'Sullivan O., Beresford T.P., Fitzgerald G.F., Corter P.D., 2012. High-throughput sequencing for detection of subpopulations of bacteria not previously associated with artisanal cheeses. *Appl. Environ. Microbiol.*, 78, 5717-5723.
- Sakamoto N., Tanaka S., Sonomoto K., Nakayama J., 2011. 16S rRNA pyrosequencing-based investigation of the bacterial community in nukadoko, a pickling bed of fermented rice bran. *Int. J. Food Microbiol.*, 144, 352-359.
- Shapiro O.H., Kushmaro A., Brenner A., 2010. Bacteriophage predation regulates microbial abundance and diversity in a full-scale bioreactor treating industrial wastewater. *ISME J.*, 4, 327-336.
- Valdés A., Ibáñez C., Simó C., García-Cañas V., 2013. Recent transcriptomics advances and emerging applications in food science. *Trends Anal. Chem.*, 52, 142-154.
- van Hijum S.A., Vaughan E.E., Vogel R.F., 2013. Application of state-of-art sequencing technologies to indigenous food fermentations. *Curr. Opin. Biotechnol.*, 24, 178-186.
- Wecks S., Van der Meulen R., Altemeersch J., Huys G., Vandamme P., Van Hummelten P., De Vuyst L., 2010. Community dynamics of bacteria in sourdough fermentations as revealed by their metatranscriptome. *Appl. Environ. Microbiol.*, 76, 5402-5408.
- Wecks S., Altemeersch J., Van der Meulen R., Vrancken G., Huys G., Vandamme P., Van Hummelten P., De Vuyst L., 2011. Metatranscriptome analysis for insight into whole-ecosystem gene expression during spontaneous wheat and spelt sourdough fermentation. *Appl. Environ. Microbiol.*, 77, 618-626.
- Xie C., Wang L., Gao Q., Yu W., Hong X., Zhao L., Zou H., 2013. Microbial community structure in fermentation process of Shaoling rice wine by Illumina-based metagenomic sequencing. *J. Sci. Food Agric.*, 93, 3121-3125.
- Yeung M., 2012. ADSA Foundation Scholar Award: Trends in culture-independent methods for assessing dairy food quality and safety: Emerging metagenomic tools. *J. Dairy Sci.*, 95, 6831-6842.

6

Microbiome du sol

Barbara Pivato, Nicolas Chemidlin Prévost-Boure, Philippe Lemanceau

Introduction

Les sols sont des environnements vivants comprenant une biomasse largement supérieure à celle présente à leur surface. Ainsi, l'ordre de grandeur de cette biomasse est estimé de 1 à 5 tonnes par hectare pour la faune et de 3,5 et 1,5 tonnes par hectare respectivement pour les champignons et les bactéries. Cette biomasse est constituée d'une immense biodiversité, en particulier microbienne. L'analyse de ce microbiome est confrontée à des difficultés majeures. La première est bien évidemment liée à la taille microscopique des organismes considérés, comme leur nom l'indique. Une seconde difficulté majeure est associée au niveau de diversité particulièrement élevé de ce microbiome dans les sols (10^4 à 10^6 génotypes/g sol, Schloss et Handelsman, 2006), dont on réalise juste l'étendue. En effet, jusqu'à récemment, nous n'avions accès qu'aux micro-organismes dits cultivables dont on sait maintenant qu'ils ne représentent qu'une très faible part de la biodiversité totale (Amann *et al.*, 2001). Les études ne permettaient donc qu'une vision très tronquée du microbiome des sols. L'accès à cette biodiversité est de plus rendu difficile par sa localisation cachée, l'hétérogénéité de la matrice tellurique et la variabilité de la distribution des micro-organismes et de leur biodiversité à différentes échelles spatiales (agrégat de sol, parcelle, paysage, territoire, continent) et temporelles (cycles de développement microbien, cycles des cultures végétales, rotations des cultures, changements de paysages). Finalement, les recherches en écologie microbienne du sol ont pendant longtemps souffert du manque de prise en compte de la grande variété des situations environnementales dans lesquelles évolue la biodiversité tellurique (type de sol, de climat, de mode d'usage), de telle sorte que les conclusions pouvaient varier d'une étude à l'autre (Fierer et Jackson, 2006 ; Pasternak *et al.*, 2013).

En dépit des difficultés associées à son étude, l'analyse du microbiome tellurique représente un enjeu majeur du fait de son rôle central dans le fonctionnement

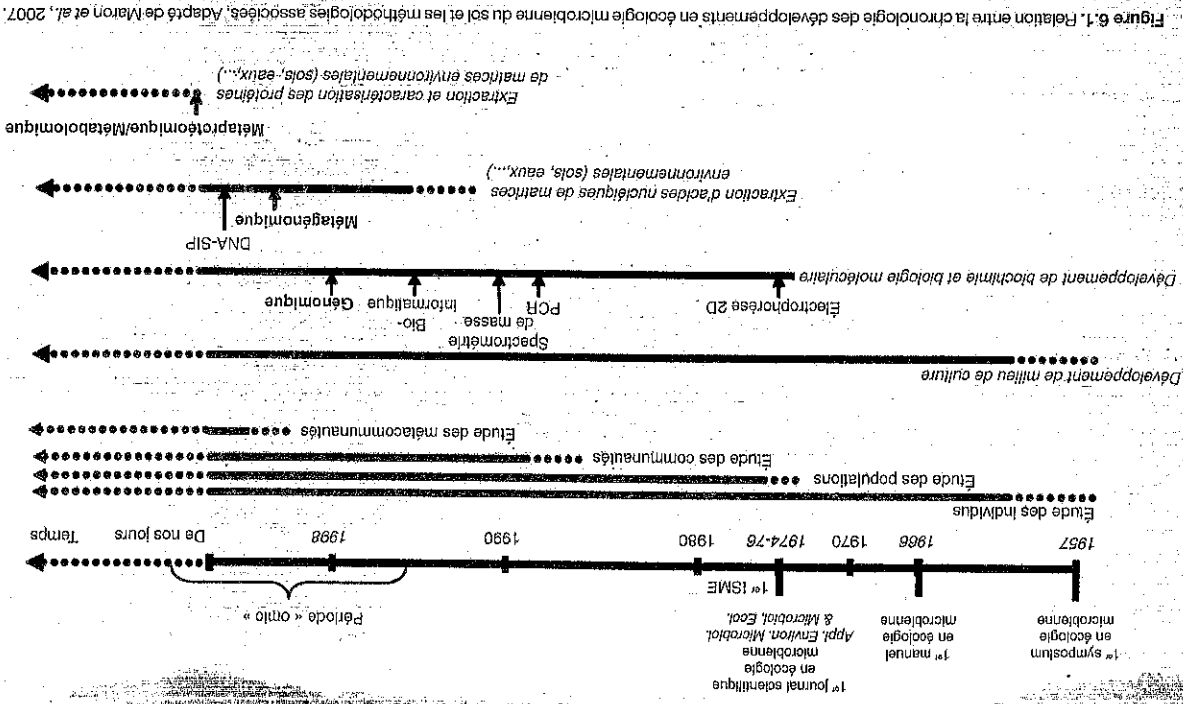
biologique des sols, de sa traduction en services écosystémiques, et du patrimoine que ce microbiome représente comme réservoir de gènes et de métabolites pour des applications potentielles en biotechnologies, pharmacie, phytopharmacie... (Uchiyama et Miyazaki, 2009). Le microbiome contribue en effet à la formation des sols et aux cycles biogéochimiques, en particulier de l'azote et du carbone, avec des conséquences directes sur la nutrition des plantes et la productivité agricole et forestière, mais également sur les services environnementaux associés relatifs à la régulation de l'émission/storage de gaz à effet de serre par les sols (émission de CO₂/stockage de carbone, émission de N₂O/réduction de N₂O en N₂). Les interactions entre plantes et micro-organismes telluriques contribuent à l'adaptation des végétaux aux stress abiotiques comme une faible fertilité et des variations du régime hydrique (Marulanda *et al.*, 2009), et biotiques, *via* la protection contre les agents phytopathogènes (antagonisme microbien et élicitation des réactions de défense de la plante-hôte — van Loon *et al.*, 1998 ; Berendsen *et al.*, 2012). De façon générale, la biodiversité microbienne des sols contribue à la productivité mais également à la stabilité (résistance, résilience) de l'agroécosystème et donc à sa durabilité (Yachi et Loreau, 1999).

Compte tenu des enjeux majeurs représentés par la connaissance et la gestion du microbiome des sols mais également des difficultés associées à son étude, des efforts considérables ont été et sont dédiés pour proposer des méthodologies permettant l'analyse de son abondance, de sa diversité et de son fonctionnement (figure 6.1). Les progrès dans ce domaine ont permis le développement de méthodes pour successivement étudier la physiologie de souches modèles, puis la diversité de populations et ensuite de communautés isolées du sol, respectivement en développant des milieux de culture adaptés (années 1960) et des techniques d'électrophorèse de l'ADN (années 1970-1980). Les progrès se sont ensuite accélérés avec le développement de techniques permettant l'extraction d'ADN des sols évitant la phase de culture microbienne et autorisant ainsi une vue plus exhaustive du microbiome tellurique. Cette étape majeure a ouvert la voie de la métagénomique pour l'analyse de l'ensemble des génomes présents dans un sol (Tringe *et al.*, 2005) (figure 6.2). Cette voie a été grandement favorisée par la diminution spectaculaire du coût du séquençage de l'ADN (<http://www.genome.gov/sequencingcosts/>), grâce aux développements méthodologiques générés par les grands programmes des séquençages du génome humain (Venier *et al.*, 2001) et du microbiome digestif (Cho et Blaser, 2012).

Les approches de génomique et de métagénomique appliquées à l'étude de la biologie des sols permettent et doivent permettre des avancées majeures dans notre connaissance de la biodiversité des sols et de leur fonctionnement à la fois aux niveaux académique et opérationnel.

Il s'agit (figure 6.3) :

- au niveau académique
 - de caractériser la biodiversité des sols,
 - de mettre en relation biodiversité, activités, fonctions et services écosystémiques,
 - d'identifier et hiérarchiser les paramètres abiotiques et biotiques ayant un impact sur la biodiversité du sol et son fonctionnement ;



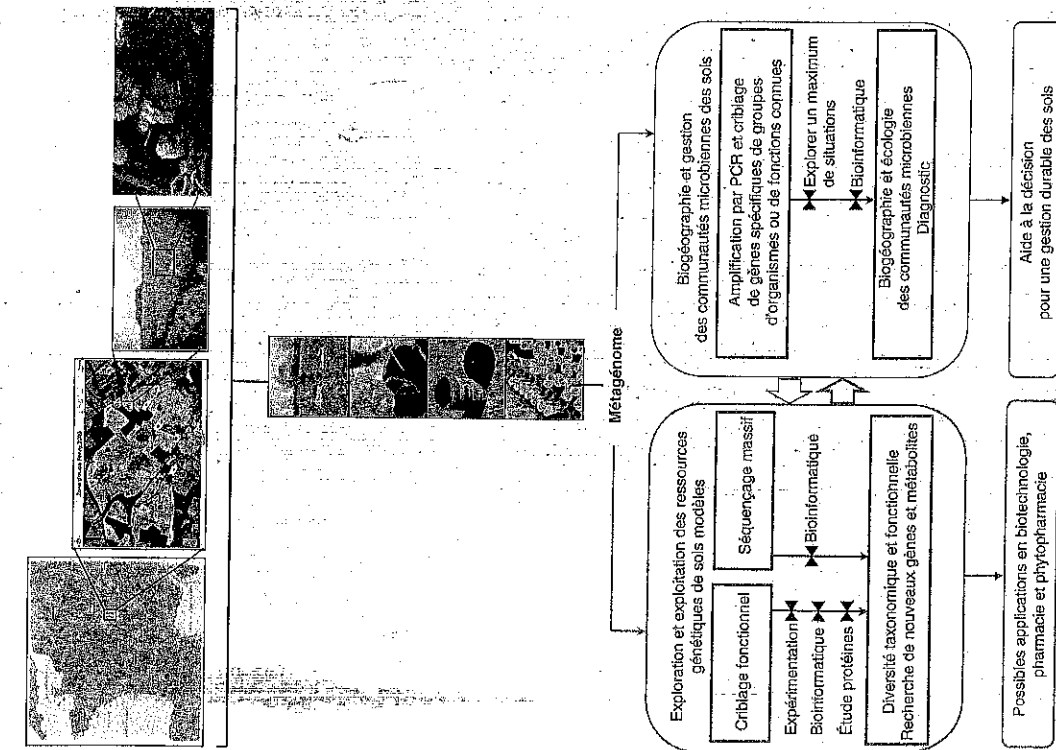


Figure 6.2. Deux types de stratégies pour l'analyse du métagénome du sol reposant sur l'exploration et l'exploitation des ressources génétiques de sols modèles, et la biogéographie et la gestion des communautés microbiennes des sols.

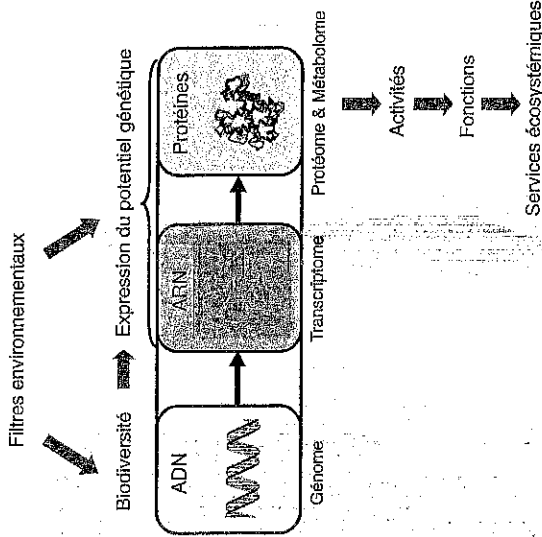


Figure 6.3. Enjeux des recherches en écologie microbienne des sols incluant : l'analyse de la biodiversité ; les relations entre biodiversité, activités, fonctions et finalement services écosystémiques ; et l'influence des filtres environnementaux à la fois sur la biodiversité et sa traduction en activités, fonctions et services écosystémiques. Cette figure correspond au schéma conceptuel du projet européen EcoFinders.

- au niveau opérationnel
 - de proposer des éléments de diagnostic et d'aide à la décision pour la gestion durable des sols,
 - de découvrir de nouveaux gènes et métabolites pour des applications agro-médicales, biotechnologiques, environnementales, pharmaceutiques et phytopharmaceutiques.

Pour répondre à ces enjeux, deux stratégies sont suivies (figure 6.2). La première vise à explorer et exploiter les ressources génétiques de sols modèles. Elle consiste à analyser avec une bonne profondeur la diversité et le potentiel fonctionnel du microbiome tellurique par un séquençage massif de l'ADN directement extrait du sol ou cloné dans des vecteurs. Les clones peuvent être de plus criblés pour la recherche de séquences d'intérêt ; compte tenu de l'importance du séquençage requis, cette stratégie ne peut matériellement être appliquée que sur un nombre limité de sols. La deuxième stratégie vise à décrire l'abondance et la diversité microbienne en utilisant des régions ciblées permettant de classer les micro-organismes selon des groupes à valeur taxonomique ou fonctionnelle ; cette stratégie moins coûteuse peut être appliquée à un plus grand nombre de sols soumis à des conditions environnementales et des modes d'usage variés.

Exploration et exploitation des ressources génétiques de sols modèles

L'exploration de la biodiversité microbienne de sols modèles repose sur l'extraction de l'ADN du sol, son séquençage direct ou après clonage.

Le métagénome a ainsi été étudié dans des environnements complexes tels que la mer des Sargasses (Venter *et al.*, 2004), une mine acide (Lyson *et al.*, 2004), le tube digestif (Chapitre 4), et plus récemment un sol issu de l'essai de longue durée de Rothamsted (Park Grass, RU) (Vogel *et al.*, 2009).

Le défi majeur d'analyser *in extenso* le métagénome de ce sol afin de caractériser sa diversité et son fonctionnement a été lancé par Vogel *et al.* (2009) à travers l'établissement du consortium international TerraGenome. Au niveau national, cette initiative s'est traduite par le projet MetaSoil (<http://www.genomemicrobiologie.org/Projets/METAISOIL.html>), coordonné par Pascal Simonet et Tim Vogel (École Centrale, Lyon), qui a consisté d'une part à séquencer (5 Gbp) directement de l'ADN extrait du sol (pyroséquençage en plateforme 454 GS FLX Titanium), et d'autre part à construire une banque de fosmidés (80 Gbp) dans *Escherichia coli* en clonant des fragments de 40 Kb (deux millions de clones). Sur les 5 Gbp séquencés, 34,5 % des séquences ont été assignées à des fonctions et seulement 1 % des séquences annotées présentaient une forte homologie avec des séquences déjà décrites. Néanmoins, cette étude pionnière a permis d'acquies des informations fonctionnelles sur ce sol de référence (Delmont *et al.*, 2012). À titre d'exemple, il apparaît que les sous-systèmes fonctionnels les plus abondants sont liés aux systèmes de signalisation cAMP et de transport Ton et Tol (Delmont *et al.*, 2012). Les analyses issues du criblage de la banque de fosmidés sont en cours. Depuis lors, d'autres projets d'analyse du métagénome de sols ont été initiés (Myrøld *et al.*, 2014 ; tableau 6.1).

Globalement ces projets sont confrontés à des challenges majeurs liés :

- à l'immensité de la biodiversité des sols ;
- à l'abondance variable des géotypes et des gènes de fonctions, alors que des populations relativement peu abondantes peuvent jouer des rôles importants ;
- au nombre encore faible d'organismes telluriques dont le génome a été séquencé et de bases de données de séquences fonctionnelles.

Ainsi, si l'on considère en première approximation qu'un gramme de sol comporte 10^9 cellules bactériennes avec chacune un génome moyen de l'ordre de 4 Mbp, le métagénome de la communauté bactérienne tellurique serait estimé à 4×10^9 Gbp d'ADN. Ce calcul est bien sûr biaisé car il ne prend pas en compte la représentation en plusieurs exemplaires du même géotype. Il n'en demeure pas moins qu'il permet de mettre en évidence le coût gigantesque de séquençage requis pour obtenir une profondeur d'analyse suffisante permettant d'assurer une description exhaustive du métagénome. Sur la base de ce calcul, les 5 Gbp du projet MetaSoil ne représenteraient donc que 0,000125 % du métagénome bactérien du sol de Park Grass.

La difficulté d'accès aux séquences présentes à de faibles fréquences est illustrée par les résultats du projet européen METACONTROL au cours duquel, en dépit de l'effort majeur de séquençage (entre 6 000 et 60 000 clones pour cinq librairies),

Tableau 6.1. Exemples de synthèses bibliographiques sur l'analyse du métagénome des sols suivant la stratégie décrite p. 76.

Thème de la synthèse	Points particuliers	Référence
Application d'approches de métagénomique pour l'identification de nouvelles molécules antimicrobiennes		Banik et Brady, 2010
Métagénomique et antibiotiques		Garmendia <i>et al.</i> , 2012
Avancées et perspectives en métagénomique. Proposition des pistes pour optimiser les études en métagénomique		Kowalchuk <i>et al.</i> , 2007
Exploration de la résistance aux antibiotiques dans les sols à travers une approche de métagénomique		Monier <i>et al.</i> , 2011
Description des différentes techniques de séquençage appliquées en métagénomique	Figures illustrant la biochimie à la base des différentes techniques de séquençage	Morey <i>et al.</i> , 2013
Utilisation des approches de métagénomique pour comprendre les processus microbiens du sol	Tableaux résumant les études effectuées sur le sol avec des approches de métagénomique, métranscriptomique et métabolomique	Myrøld <i>et al.</i> , 2014
État des connaissances dans le domaine de l'analyse des données métagénomiques		Scholz <i>et al.</i> , 2012
Introduction à l'analyse des données de métagénomique		Sharpton, 2014
Approches en métagénomique		Simon et Daniel, 2011
Métagénomique fonctionnelle et découverte de nouvelles enzymes	Tableau résumant les approches utilisées pour la découverte de nouvelles enzymes	Uchiyama et Miyazaki, 2009

seule une très faible proportion des clones (< 0,05 %) contenait des séquences codant les polykétide synthases (PKS), enzymes impliquées dans la production de métabolites secondaires, tels que le sidérophore ferricrocine et les gibbérellines, dont on connaît pourtant le rôle déterminant dans les interactions biotiques (van Elsas *et al.*, 2008). En effet, un calcul théorique proposé par Leveau (2007) estime qu'il faudrait tester une librairie de 57 000 clones avec des inserts de 40 Kbp pour atteindre une probabilité 99 % de retrouver le gène d'intérêt qui représenterait 1 % du métagénome. Cette difficulté est représentée par l'image faite par Kowalchuk *et al.* (2007) de la recherche d'une aiguille dans une botte de foin.

Finalement, l'assemblage des données issues du séquençage haut débit (*reads*) en séquences plus longues continues et ordonnées (*contigs*) demeure un verrou bio-informatique majeur. Les progrès dans ce domaine requièrent clairement la mise à disposition d'un plus grand nombre de séquences de génomes microbiens d'origine environnementale ainsi que leur organisation en bases de données centralisées et accessibles.

Les résultats attendus du séquençage massif de sols modèles, au-delà de la connaissance de leur biodiversité, doivent nous permettre de progresser dans la connaissance du fonctionnement biologique du sol en classant les protéines putatives à partir des séquences du métagnôme. Cependant, à nouveau les bases de données demeurent très insuffisantes et seule une faible fraction des familles (en moyenne, moins de 50 %) peut être associée à des gènes de fonctions. De plus, la présence de ces gènes dans l'environnement ne signifie bien sûr pas qu'ils soient exprimés au moment de l'échantillonnage (Sharpton, 2014). Une autre approche de métagnomique fonctionnelle repose sur le clonage préalable de l'ADN extrait du sol dans différents vecteurs (plasmides, cosmides, fosmidés, *bacterial artificial chromosomes* — BAC) selon la longueur du fragment d'ADN à intégrer. Ces vecteurs sont introduits dans des cellules bactériennes hôtes, notamment *E. coli*, pour constituer des banques de clones. Ces banques sont criblées pour la recherche d'activités d'intérêt et des gènes qui les codent. L'identification de ces activités et gènes présente le double intérêt de progresser dans la connaissance du fonctionnement biologique du sol et de permettre la valorisation de nouveaux métabolites ou de nouvelles fonctions pour des applications potentielles en biotechnologie, pharmacie, phytopharmacie... Cette approche a d'ores et déjà permis l'identification de nouveaux gènes fonctionnels issus du sol. Parmi ceux-ci, on peut citer

- des biocatalyseurs (estérases, lipases, oxygénases et xylanases),
- des lacramases, impliquées dans la résistance aux antibiotiques,
- des molécules servant de signal dans le processus de *Quorum Sensing* (Banik et Brady, 2010 ; Garmendia *et al.*, 2012 ; Uchiyama et Miyazaki, 2009).

Étant donné le coût considérable que représente le séquençage complet d'un sol, il ne peut s'appliquer qu'à un nombre limité de sols modèles qui pourraient ensuite être utilisés pour mettre en œuvre une approche de métagnomique comparative (à la fois taxonomique et fonctionnelle) à l'image de celle suivie pour le microbiome du tube digestif (Arumugam *et al.*, 2011). Ces sols représenteraient en effet des sols de référence à l'image des entérotypes définis pour le métagnôme du tube digestif. Il serait alors possible de comparer chaque nouveau microbiome tellurique à ceux de référence afin d'identifier leurs proximités respectives et les différences comparées au sol de référence le plus proche. Il ne s'agit bien sûr à ce stade que d'une perspective dont il est difficile d'apprécier le réalisme compte tenu de l'immense varié des types de sols et des situations environnementales. Fierer *et al.* (2012) ont effectué une première tentative de métagnomique comparative sur deux sols soumis à différents régimes de fertilisation azotée et ont ainsi mis en évidence que les phyla coprotrophes sont favorisés par cette fertilisation. Afin de minimiser la variabilité liée aux types de sols, on pourrait, à l'image des recherches

conduites sur le microbiome du tube digestif, envisager d'appliquer cette approche de métagnomique comparative plutôt aux communautés associées à différents hôtes végétaux. Dans ce contexte, il est intéressant de noter la proposition récente d'un microbiome cœur associé à la plante-hôte indépendamment du sol dans lequel elle est cultivée (Bulgarelli *et al.*, 2012 ; Lundberg *et al.*, 2012).

Biogéographie et gestion des communautés microbiennes des sols

Cette stratégie vise non plus à caractériser l'ensemble du métagnôme mais plutôt des séquences particulières choisies pour leur valeur taxonomique (par exemple le gène de l'ADNr 16S) ou pour les activités qu'elles codent (gènes de fonction). Cette caractérisation de la diversité taxonomique et du potentiel fonctionnel, ciblée sur des régions particulières de l'ADN microbien, permet de redéployer les ressources analytiques décrites dans la section précédente sur un grand nombre d'échantillons de sols simultanément (haut débit analytique) et permet ainsi le développement d'approches à de larges échelles spatiales et temporelles (figure 6.4). Ce type de recherches s'inscrit dans le domaine de la biogéographie qui vise à caractériser la distribution de la biodiversité dans l'espace et le temps (Martiny *et al.*, 2006). Cette discipline scientifique a été développée à partir du XVIII^e siècle pour l'étude des macro-organismes (plantes, animaux) et a permis d'identifier et de hiérarchiser les mécanismes générant et maintenant leur diversité (spéciation, sélection, dispersion, interactions biotiques, etc.) trouvant ainsi des ouvertures en écologie de la conservation (Whittaker *et al.*, 2005 ; Harte *et al.*, 2009). Même si le premier postulat a été émis en 1934 par Bas Becking : « Tout est partout, mais, l'environnement sélectionne », l'application de la biogéographie aux micro-organismes telluriques s'est développée à partir des années 2000 avec l'avènement des méthodes moléculaires de caractérisation de la diversité (méthodes d'empreintes moléculaires), particulièrement aujourd'hui avec les méthodes de séquençage à haut débit. Ces méthodes permettent en effet l'analyse d'un grand nombre d'échantillons de sols issus d'environnements variés : types de sols, de climats, de modes d'usages (Green *et al.*, 2004 ; Martiny *et al.*, 2006 ; Fierer et Jackson, 2006). Les recherches en biogéographie microbienne ont pour objectif de comprendre les règles d'assemblage des micro-organismes dans les sols et en particulier de déterminer si les lois de distribution identifiées pour les macro-organismes s'appliquent aux micro-organismes ou si des règles spécifiques régissent cette distribution en raison de leurs particularités (taille microscopique, temps de génération très court, extrême diversité, grandes capacités adaptatives). Comme pour les macro-organismes, ces recherches visent également à identifier et hiérarchiser les paramètres environnementaux (édaphiques, climatiques, usage des sols, pratiques agricoles) impactant l'assemblage et la distribution des communautés microbiennes du sol à grande échelle. Les résultats obtenus indiquent clairement que les micro-organismes telluriques sont distribués de façon hétérogène et structurés selon des patrons biogéographiques tant au niveau du groupe microbien (Cho et Tiedje, 2000 ; Papke et Ward, 2004 ; Whitaker *et al.*, 2005 ; Psey *et al.*, 2007) qu'à celui de la communauté microbienne (Fierer et Jackson 2006 ; Martiny *et al.*, 2006 ; Dequiedt *et al.*, 2009 ; Rausard *et al.*, 2013 ; Chemidlin Prévoost-Bouré

et al., 2014). Ces conclusions incitent logiquement à identifier les processus à l'origine de la structuration décrite.

Une manière de discriminer les processus de diversification et d'évaluer leur influence relative est de considérer les variations spatiales des communautés à travers la relation aire-espèces (Gleason, 1922 ; Harte et al., 1999, 2009 ; Morlon et al., 2008 ; Ranjard et al., 2013 ; Chemidlin, Prévost-Bouré et al., 2014). Cette relation prend en compte les variations de diversité biologique lorsque l'on augmente l'aire échantillonnée ou la distance entre les sites échantillonnés (figure 6.4). Elle est principalement employée pour étudier les processus de sélection et de dispersion car les marqueurs moléculaires servant à la caractérisation taxonomique des communautés microbiennes (ADNr 16S pour les prokaryotes, ADNr 18S pour les eukaryotes) sont peu affectés par les processus de dérive écologique ou de mutation. De nombreuses études ont démontré que l'assemblage des communautés est en premier lieu sous le contrôle de processus de sélection (Hanson et al., 2012). Le second processus le plus important semble être la dispersion même s'il reste fortement discuté. En effet, son importance a été démontrée pour des groupes microbiens particuliers (genre *Pseudomonas*, champignons ectomycorhiziens ; Cho et Tiedje, 2000 ; Peay et al., 2007, 2012) mais sa contribution à l'échelle de la communauté microbienne n'est pas claire. Les paramètres environnementaux ont un fort impact sur la structuration des communautés microbiennes telluriques. Ainsi, elles sont fortement structurées par les propriétés physico-chimiques des sols (pH, texture, teneur et qualité de la matière organique du sol ; Fierer et Jackson, 2006 ; Ramette et Tiedje, 2007 ; Lauber et al., 2008 ; Dequiedt et al., 2009, 2011) et le mode d'usage des sols et les pratiques agricoles (Liénhard et al., 2013, 2014 ; Lauber et al., 2008). En revanche, les paramètres climatiques ont un impact moindre dans la gamme de conditions restées (Pasternak et al., 2013), même si cette influence fait l'objet de discussions (Mulder et al., 2005 ; Drenovsky et al., 2010).

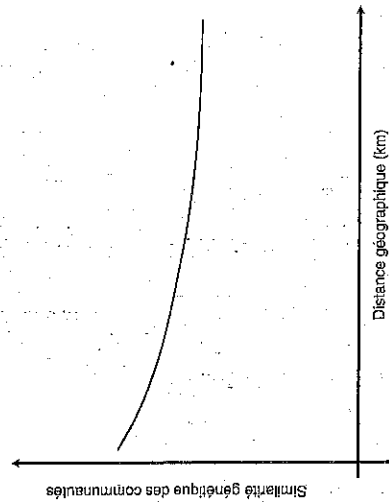


Figure 6.4. Relation aire-espèces.

Ces recherches ont fait et font encore l'objet de grands programmes nationaux (ANR) et internationaux (projets européens). À l'échelle française, le projet ECOMIC-RMQS se place comme un pionnier de la biogéographie microbienne. Débuté en 2006, ce projet avait pour objectifs de mieux comprendre les processus de diversification des communautés microbiennes et leurs importances relatives, et de mieux hiérarchiser les filtres environnementaux impliqués dans la structuration spatiale de l'abondance et de la diversité microbienne des sols. Pour atteindre ces objectifs, le projet ECOMIC-RMQS s'est appuyé sur le Réseau de mesure de la qualité des sols (Arrouays et al., 2002) géré par le GIS Sol (www.gissol.fr). Ce réseau constitue un échantillonnage systématique des sols français suivant une grille de 16 km de côté, soit un total de 2 150 échantillons de sol analysés. Les résultats de ces analyses ont permis de mettre en évidence la distribution hétérogène des communautés microbiennes du sol à l'échelle du territoire national français tant en termes d'abondance que de structure génétique (Dequiedt et al., 2009, 2011 ; Chemidlin, Prévost-Bouré et al., 2014 ; figure 6.5). Ils ont aussi révélé que les processus de dispersion et de sélection sont tous deux impliqués dans la diversification des communautés microbiennes (Ranjard et al., 2013 ; Chemidlin, Prévost-Bouré et al., 2014). Il a aussi permis de mettre en évidence que le processus de sélection est principalement basé sur des paramètres locaux (types de sol et occupation) plutôt que des paramètres distaux (climat et géomorphologie), et que certains modes d'usage des sols peuvent avoir des impacts forts sur les communautés bactériennes du sol (particulièrement des pratiques agricoles) conduisant à une réduction importante de la biomasse microbienne et modifiant la structure génétique des communautés bactériennes (Dequiedt et al., 2009, 2011). L'influence de ces différents facteurs environnementaux se traduit à l'échelle du paysage puisqu'il a été observé que plus un paysage était diversifié en termes d'habitats, plus les communautés bactériennes tendent à se diversifier (Ranjard et al., 2013).

En plus de ces apports fondamentaux, l'acquisition exhaustive des données sur les communautés bactériennes des sols a abouti à des sorties appliquées importantes telles que :

- la définition des niveaux d'abondance et de diversité des communautés bactériennes dans les sols français,
- l'évaluation de l'impact de l'occupation des sols et des activités anthropiques sur l'abondance et la diversité bactérienne des sols,
- l'identification de bio-indicateurs spécifiques d'une occupation du sol ou d'un impact anthropique.

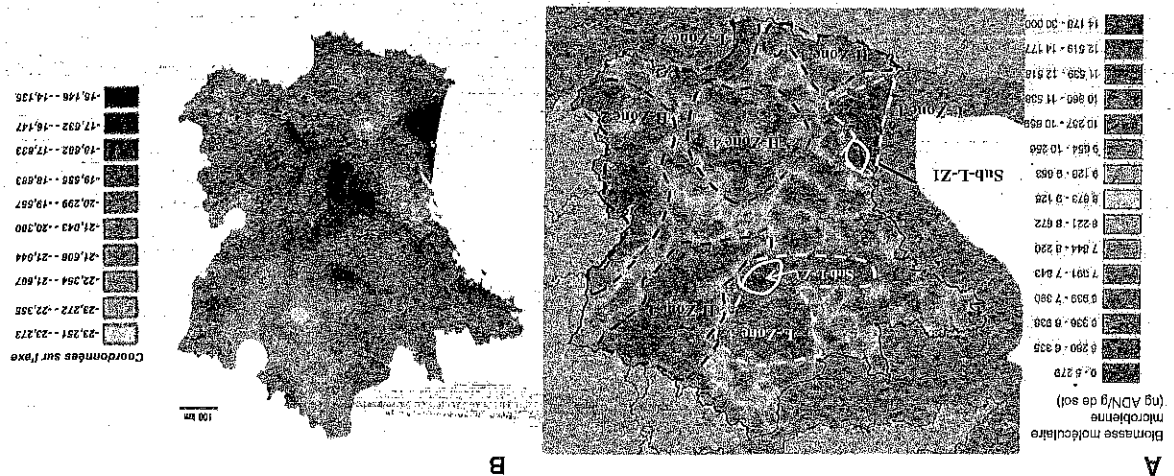
Ces sorties sont aujourd'hui mises en œuvre en association avec l'ensemble des acteurs du monde agricole. D'autres ont vu le jour dans ce domaine au cours des années récentes à des échelles nationales en particulier aux Pays-Bas (Rungers et al., 2009), au Royaume-Uni (Countrywide Survey, www.countrywidesurvey.org.uk, Griffiths et al., 2011), mais également à des niveaux plus larges en Europe (EcoFINDERS, <http://www.ecofinders.eu>, Lemanceau, 2011) et aux États-Unis (Fierer et Jackson, 2006), confirmant et élargissant les conclusions obtenues.

L'ensemble de ces études montre que cartographier les communautés microbiennes des sols sur de vastes territoires est possible et apporte des réponses à des questions à la fois fondamentales et opérationnelles. La compréhension des processus, l'identification et la hiérarchisation des facteurs faisant varier la diversité des communautés microbiennes des sols permettent l'identification de leviers de gestion mais aussi de développer des approches de modélisation pour prédire l'impact de changement d'usage des sols. Outre les progrès dans la compréhension de l'écologie des communautés telluriques, les études de biogéographie conduisent en effet à l'établissement de référentiels indiquant les gammes de variations normales de diversité selon les types de sols et leur mode d'usage. Ces référentiels permettent la formulation de diagnostic sur les propriétés biologiques des sols comme cela est fait depuis longtemps pour leurs propriétés physico-chimiques. Sur la base de ces diagnostics, il sera ensuite possible d'envisager des éléments d'aide à la décision grâce aux recherches actives menées dans le domaine de l'agroécologie (Lemanceau *et al.*, 2015).

Conclusion et perspectives

La métagnomique des sols ouvre des perspectives particulièrement stimulantes pour l'analyse de leur biodiversité et de leur fonctionnement. Ce type d'analyses doit permettre à terme de découvrir et répertorier les organismes peuplant ce qui a longtemps représenté une boîte noire et qui constitue pourtant une des plus grandes sources de biodiversité de la planète. De telles recherches bénéficient des développements méthodologiques et des réductions de coût spectaculaires du séquençage à haut débit acquis lors de programmes emblématiques tels que le séquençage du génome humain et celui du tube digestif. En France, elles bénéficient également du soutien de programmes spécifiquement dédiés tels que le métaprogramme Inra Méta-omiques des écosystèmes microbiens, MEM (<http://www.mem.inra.fr/>). Ces études demeurent cependant confrontées à la profondeur de séquençage requise pour une analyse exhaustive ainsi qu'aux difficultés d'analyse et d'assemblage des séquences. Elles représentent donc des défis qui nécessitent une mobilisation de moyens financiers mais également analytique au niveau mondial. On peut citer dans ce domaine l'initiative de Pascal Simonet et Tim Vogel, mais également celle du projet plus large « Earth Microbiome » (<http://www.earthmicrobiome.org/>). De même, l'analyse de la diversité des sols à des échelles spatiales et temporelles importantes ne peut évidemment pas s'arrêter aux frontières entre pays établies par l'homme. Un consensus international doit être trouvé pour disposer d'un corpus de méthodologies standardisées autorisant la comparaison des données issues de différents projets et ainsi la construction de bases de données et de référentiels. Dans ce contexte, on peut saluer la mise en place de la *Global Soil Biodiversity Initiative* (<http://globalsoilbiodiversity.org/>) qui œuvre dans cette direction. Les progrès qui devraient en résulter sont non seulement une augmentation spectaculaire de nos connaissances du fonctionnement des sols mais également la découverte de nouveaux gènes et métabolites ainsi que l'établissement de diagnostics de l'état des sols afin de proposer des modes de gestion durable de ce patrimoine essentiel mais fragile. Il s'agit donc d'un défi majeur pour l'humanité.

Figure 65. Cartographie de l'abondance des communautés microbiennes et de la composition bactérienne des sols de France métropolitaine. A. Carte de biomasse microbienne moléculaire des sols (Daguidet *et al.*, 2011), i.e. de la quantité d'ADN extraits du sol, correspondant à l'abondance des micro-organismes (bactéries et champignons). La biomasse moléculaire microbienne est distribuée de manière hétérogène suivant des patterns biogéographiques. Zones H et L : zones présentant des biomasses moléculaires microbiennes fortes ou faibles. Les flèches indiquent des sous-zones présentant des biomasses particulièrement faibles, valeurs souvent associées à l'utilisation du sol (e.g. sous-zone L-Z1 : culture viticole) ou à un type de sol particulier (e.g. sous-zone L-Z3 : sols sableux). B. Cartographie des variations de structure génétique des communautés bactériennes du sol à l'échelle de la France. La structure génétique des communautés bactériennes du sol est caractérisée au travers d'analyses multivariées sous contrainte spatiale (i.e. prenant en compte les relations de voisinage entre sites) qui permettent de résumer l'information contenue dans les données d'inventaire sur un nombre d'axes restreint. Les coordonnées des sites sur l'axe 1 de l'analyse sont représentées ici. Par cette approche, des couleurs divergentes sur la carte indiquent une divergence de communautés bactériennes du sol sont distribuées de manière hétérogène suivant des patterns biogéographiques (280 km de diamètre), ces variations étant en lien avec l'utilisation du sol et le type de sol à chaque site.



Références bibliographiques

- Amann R., Fuchs B.M., Behrens S., 2001. The identification of microorganisms by fluorescence *in situ* hybridisation. *Curr. Opin. Biotechnol.*, 12, 231-236.
- Arrouays D., Jolivet C., Bouillon L., Bodineau G., Saby N., Grolleau E., 2002. A new projection in France: a multi-institutional soil quality monitoring network. Une initiative nouvelle en France : la mise en place d'un réseau multi-institutionnel de mesure de la qualité des sols (RMQS). *C. R. Académie d'Agriculture de France*, 88, 93-105.
- Arunugam M., Raes J., Peletier E., Le Paslier D., Yamada T., Mendel D.R., *et al.*, 2011. Enterotypes of the human gut microbiome. *Nature*, 473, 174-180.
- Banik J.J., Brady S.F., 2010. Recent application of metagenomic approaches toward the discovery of antimicrobials and other bioactive small molecules. *Curr. Opin. Microbiol.*, 13, 603-609.
- Beas Becking L.G.M., 1934. *Geobiologie of inleiding tot de milieukunde*. W.P. Van Stockum & Zoon, The Hague, The Netherlands (in Dutch).
- Berendse R.L., Pieterse C.M.J., Bakker P.A.H.M., 2012. The rhizosphere microbiome and plant health. *Trends Plant Sci.*, 17, 478-486.
- Bulgarelli D., Rott M., Schlaeppli K., Van Themaat E.V.I., Ahmadinejad N., Assenza F., Rauf P., Huettel B., Reinhardt K., Schmelzer E., Peplins J., Glockner F.O., Amanur R., Eichhöfer T., Schulze-Lefert P., 2012. Revealing structure and assembly cues for Arabidopsis root-inhabiting bacterial microbiota. *Nature*, 488, 91-95.
- Chemidlin Prévost-Bouré N., Dequiedt S., Thioulouse J., Lelièvre M., Saby N.P.A., Jolivet C., Arrouays D., Plassart P., Lemanceau P., Ranjard L., 2014. Similar processes but different environmental filters for soil bacterial and fungal community composition turnover on a broad spatial scale. *PLoS ONE*, 9 (11), e111667.
- Cho J.-C., Tiedje J.M., 2000. Biogeography and degree of endemism of fluorescent *Pseudomonas* strains in soil. *Appl. Environ. Microbiol.*, 66, 5448-5456.
- Cho I., Blaser M.J., 2012. The human microbiome: at the interface of health and disease. *Nat. Rev. Genet.*, 13, 260-270.
- Delmont T.O., Prestat E., Keegan K.P., Faubladiet M., Robe P., Clark L.M., Pelletier E., Fritsch P.R., Meyer E., Gilbert J.A., Le Paslier D., Simonet P., Vogel T.M., 2012. Structure, fluctuation and magnitude of a natural grassland soil metagenome. *ISME J.*, 6, 1677-1687.
- Dequiedt S., Thioulouse J., Jolivet C., Saby N.P.A., Lelièvre M., Arrouays D., Martin M.P., Chemidlin Prévost-Bouré N., Toutain B., Arrouays D., Lemanceau P., Ranjard L., 2009. Biogeographical patterns of soil bacterial communities. *Environ. Microbiol. Rep.*, 1, 251-255.
- Dequiedt S., Saby N.P.A., Lelièvre M., Jolivet C., Thioulouse J., Toutain B., Arrouays D., Bispo A., Lemanceau P., Ranjard L., 2011. Biogeographical patterns of soil molecular microbial biomass as influenced by soil characteristics and management. *Global Ecol. Biogeography*, 20, 641-652.
- Drenovsky R.E., Steenwerth K.L., Jackson L.E., Scow K.M., 2010. Land use and climatic factors structure regional patterns in soil microbial communities. *Global Ecol. Biogeography*, 19, 27-39.
- Fierer N., Jackson R.B., 2006. The diversity and biogeography of soil bacterial communities. *Proc. Natl. Acad. Sci. USA*, 103, 626-631.
- Fierer N., Lauber C.L., Ramirez K.S., Zaneveld J., Bradford M.A., Knight R., 2012. Comparative metagenomic, phylogenetic and physiological analyses of soil microbial communities across nitrogen gradients. *ISME J.*, 6, 1007-1017.
- Garmendia L., Hernandez A., Sanchez M.B., Martinez J.L., 2012. Metagenomics and antibiotics. *Clin. Microbiol. Infect.*, 18, 27-31.
- Gleason A.H., 1922. On the relationship between species and area. *Ecology*, 3, 158-162.
- Green J.L., Holmes A.J., Westoby M., Oliver I., Briscoe D., Dungefield M., Gillings M., Beattie A.J., 2004. Spatial scaling of microbial eukaryote diversity. *Nature*, 432, 747-750.
- Griffiths R.I., Thomson B., James P., Bell T., Bailey M., Whiteley A.S., 2011. The bacterial biogeography of British soils. *Environ. Microbiol.*, 13, 1642-1654.
- Hanson C.A., Fuhrman J.A., Horner-Devine M.C., Martiny J.B.H., 2012. Beyond biogeographic patterns: processes shaping the microbial landscape. *Nat. Rev. Microbiol.*, 10, 497-506.
- Harte J., Kinzig A., Green J., 1999. Self-similarity in the distribution and abundance of species. *Science*, 284, 334-336.
- Harte J., Smith A.B., Storch D., 2009. Biodiversity scales from plots to biomes with a universal species-area curve. *Ecol. Lett.*, 12 (8), 789-797.
- Kowalchuk G.A., Spetskyjder A.G.C.L., Zhang K., Goodman R.M., Van Veen J.A., 2007. Finding the needles in the metagenome haystack. *Microb. Ecol.*, 53, 475-485.
- Lauber C.L., Strickland M.S., Bradford M.A., Fierer N., 2008. The influence of soil properties on the structure of bacterial and fungal communities across land-use types. *Soil Biol. Biochem.*, 40, 2407-2415.
- Lemanceau P., 2011. EcoFINDERS, caractériser la biodiversité et le fonctionnement des sols en Europe. *Biofuture*, 326, 56-58.
- Lemanceau P., Maron P.-A., Mazurier S., Mougel C., Pivato B., Plassart P., Ranjard L., Revellin C., Tardy V., Wipf D., 2015. Understanding and managing soil biodiversity: a major challenge in agroecology. *Agronomy for Sustainable Development*, 35 (1), 67-81.
- Leveau J.-H.J., 2007. The magic and menace of metagenomics: prospects for the study of plant growth-promoting rhizobacteria. *Eur. J. Plant Pathol.*, 119, 279-300.
- Lienhard P., Terrat S., Mathieu O., Leveau J., Chemidlin Prévost-Bouré N., Nowak V., Régner T., Favre C., Sayphoummie S., Panyasiri K., Tivet F., Ranjard L., Maron P.-A., 2013. Soil microbial diversity and C turnover modified by tillage and cropping in Laos tropical grassland. *Environ. Chem. Lett.*, 11, 391-398.
- Lienhard P., Terrat S., Prévost-Bouré N., Nowak V., Régner T., Sayphoummie S., Panyasiri K., Tivet F., Mathieu O., Leveau J., Maron P.-A., Ranjard L., 2014. Pyrosequencing evidences the impact of cropping on soil bacterial and fungal diversity in Laos tropical grassland. *Agronomy for Sustainable Development*, 34, 525-533.
- Lundberg D.S., Lebeis S.L., Paredes S.H., Youngster S., Gehring J., Malfatti S., Tremblay J., Engelbrektson A., Kunin V., Rio T.G.D., Edgar R.C., Eickholt T., Ley R.E., Hugenholtz P., Tringe S.G., Dang J.L., 2012. Defining the core Arabidopsis thaliana root microbiome. *Nature*, 488, 86-90.
- Maron P.-A., Ranjard L., Mougel C., Lemanceau P., 2007. Metaproteomics : a new approach for studying functional microbial ecology. *Microb. Ecol.*, 53, 486-493.
- Martiny J.B.H., Bohannan B.J.M., Brown J.H., Colwell R.K., Fuhrman J.A., Green J.L., Horner-Devine M.C., Kane M., Krumins J.A., Kuske C.R., Morin P.J., Naeem S., Ovreas L., Reyersbach A.L., Smith V.H., Staley J.T., 2006. Microbial biogeography: putting microorganisms on the map. *Nat. Rev. Microbiol.*, 4, 102-112.

- Marulanda A., Barca J.-M., Azcon R., 2009. Stimulation of plant growth and drought tolerance by native microorganisms (AM fungi and bacteria) from dry environments: mechanisms related to bacterial effectiveness. *J. Plant Growth Regul.*, 28, 115-124.
- Monier J.-M., Demanèche S., Delmont T.O., Mathieu A., Vogel T.M., Simonet P., 2011. Metagenomic exploration of antibiotic resistance in soil. *Curr. Opin. Microbiol.*, 14, 229-235.
- Morey M., Fernández-Marmiesse A., Castiella D., Fraga J.M., Couce M.L., Cocho J.A., 2013. A glimpse into past, present, and future DNA sequencing. *Mol. Genet. Metabol.*, 110, 3-24.
- Morlon H., Chuyong G., Condit R., Hubbell S., Kenfack D., Thomas D., Valencia R., Green J.L., 2008. A general framework for the distance-decay of similarity in ecological communities. *Ecol. Lett.*, 11, 904-917.
- Mulder C., Van Wijnen H.J., Van Wezel A.P., 2005. Numerical abundance and biodiversity of below-ground taxocenes along a pH gradient across the Netherlands. *J. Biogeogr. phy.*, 32, 1775-1790.
- Myrold D.D., Zeglin L.H., Jansson J.K., 2014. The Potential of metagenomic approaches for understanding soil microbial processes. *Soil Sci. Soc. Am. J.*, 78, 3-10.
- Papke R.T., Ward D.M., 2004. The importance of physical isolation to microbial diversification. *FEMS Microbiol. Ecol.*, 48, 293-303.
- Pasternak Z., Al-Ashhab A., Gatica J., Gafny R., Avraham S., Minz D., Gilor O., Jurkevitch I., 2013. Spatial and temporal biogeography of soil microbial communities in arid and semiarid regions. *PLoS ONE*, 8, e69705.
- Peay K.G., Bruns T.D., Kennedy P.C., Bergemann S.E., Garbelotto M., 2007. A strong species-area relationship for eukaryotic soil microbes: island size matters for ectomycorrhizal fungi. *Ecol. Lett.*, 10, 470-480.
- Peay K.G., Schubert M.G., Nguyen N.H., Bruns T.D., 2012. Measuring ectomycorrhizal fungal dispersal: macroecological patterns driven by microscopic propagules. *Mol. Ecol.*, 21, 4122-4136.
- Ramete A., Tiedje J.M., 2007. Multiscale responses of microbial life to spatial distance and environmental heterogeneity in a patchy ecosystem. *Proc. Natl. Acad. Sci. USA*, 104, 2761-2766.
- Ranjard L., Dequiedt S., Prevost-Boure N.C., Thioulouse J., Saby N.P.A., Lelievre M., Maron P.A., Morin F.E.R., Bispo A., Jolivet C., Arrouays D., Lemanceau P., 2013. Turnover of soil bacterial diversity driven by wide-scale environmental heterogeneity. *Nature Com.*, 4, 1434.
- Rungtjes M., Schouren A.J., Bloem J., van Eekeren N., de Goede R.G.M., Akkerhuis G.A.J.M., van der Wal A., Mulder C., Brussaard L., Breure A.M., 2009. Biological measurements in a nationwide soil monitoring network. *Env. J. Soil Sci.*, 60, 820-832.
- Schlösser P.D., Handelsman J., 2006. Toward a census of bacteria in soil. *PLoS Comput. Biol.*, 2, e92.
- Scholz M.B., Lo C.-C., Chain P.S.G., 2012. Next generation sequencing and bioinformatic bottlenecks: the current state of metagenomic data analysis. *Curr. Opin. Biotechnol.*, 23, 9-15.
- Sharpton T.J., 2014. An introduction to the analysis of shotgun metagenomic data. *Front. Plant Sci.*, 5, 209.
- Simon C., Daniel R., 2011. Metagenomic Analyses: past and future trends. *Appl. Environ. Microbiol.*, 77, 1153-1161.
- Tringe S.G., Von Mering C., Kobayashi A., Salamov A.A., Chen K., Chang H.W., Podar M., Short J.M., Mathur E.J., Deter J.C., Bork P., Hugenholtz P., Rubin E.M., 2005. Comparative metagenomics of microbial communities. *Science*, 308, 554-557.
- Tyson G.W., Chapman J., Hugenholtz P., Allen E.E., Ram R.J., Richardson P.M., Solovyev V.V., Rubin E.M., Rokhsar D.S., Banfield J.F., 2004. Community structure and metabolism through reconstruction of microbial genomes from the environment. *Nature*, 428, 37-43.
- Uchiyama T., Miyazaki K., 2009. Functional metagenomics for enzyme discovery: challenges to efficient screening. *Curr. Opin. Biotechnol.*, 20, 616-622.
- Van Elsas J.D., Costa R., Jansson J., Sjöling S., Bailey M., Nalin R., Vogel T.M., Van Overbeek L., 2008. The metagenomics of disease-suppressive soils - experiences from the METACONTROL project. *Trends Biotechnol.*, 26, 591-601.
- Van Loon L.C., Bakker P.A., Pieterse C.M.J., 1998. Systemic resistance induced by rhizosphere-bacteria. *Ann. Rev. Phytopathol.*, 36, 453-483.
- Venter J.C., Adams M.D., Myers E.W., Li P.W., Mural R.J., Sutton G.G., et al., 2001. The sequence of the human genome. *Science*, 291, 1304-1351.
- Venter J.C., Remington K., Heidelberg J.F., Halpern A.L., Rusch D., Eisen J.A., et al., 2004. Environmental genome shotgun sequencing of the Sargasso Sea. *Science*, 304, 66-74.
- Vogel T.M., Simonet P., Jansson J.J., Hirsch P.R., Fiedje J.M., van Elsas J.D., Bailey M.J., Nalin R., Philippot L., 2009. Editorial TerraGenome: a consortium for the sequencing of a soil metagenome. *Nat. Rev. Microbiol.*, 7, 252.
- Whittaker R.J., Aratjo M.B., Paul J., Ladle R.J., Watson J.E.M., Willis K.J., 2005. Conservation biogeography: assessment and prospect. *Diversity and Distributions*, 11, 3-23.
- Yachi S., Loreau M., 1999. Biodiversity and ecosystem productivity in a fluctuating environment: The insurance hypothesis. *Proc. Natl. Acad. Sci. USA*, 96, 1463-1468.