



**HAL**  
open science

# Caractérisation de la prévalence et de la diversité des gènes de résistance bactérienne à des antibiotiques dans le sol par des approches métagénomiques : Quel est le potentiel de la dissémination du résistome tellurique?

Joseph Nesme

## ► To cite this version:

Joseph Nesme. Caractérisation de la prévalence et de la diversité des gènes de résistance bactérienne à des antibiotiques dans le sol par des approches métagénomiques : Quel est le potentiel de la dissémination du résistome tellurique?. Autre. Ecole Centrale de Lyon, 2014. Français. NNT : 2014ECDL0015 . tel-01124013

**HAL Id: tel-01124013**

**<https://theses.hal.science/tel-01124013>**

Submitted on 6 Mar 2015

**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.

**THÈSE**

**Présentée devant :**

**ÉCOLE CENTRALE DE LYON**

**Pour obtenir le grade de :**

**DOCTEUR**

de l'école doctorale :

Électronique, électrotechnique, automatique

**UNIVERSITÉ DE LYON**

Spécialité : Ingénierie Du Vivant

Par

**Joseph NESME**

**Caractérisation de la prévalence et de la diversité des gènes de  
résistance bactérienne à des antibiotiques dans le sol par des  
approches métagénomiques:**

***Quel est le potentiel de dissémination du résistome tellurique ?***

soutenue le 16 mai 2014

Devant le jury composé de:

<b>Dr. William GAZE</b> , Senior Lecturer, University of Exeter Medical School, Truro, UK.	<b>Rapporteur</b>
<b>Dr. Edward TOPP</b> , Scientific Researcher, Agriculture and Agri-Food Canada, London ON, Canada.	<b>Rapporteur</b>
<b>Dr. Emmanuelle MAGUIN</b> , DU INRA, UMR 1319 MICALIS, Jouy-En-Josas, France..	<b>Examineur</b>
<b>Dr. Sylvie NAZARET</b> , CR CNRS, Université Claude Bernard Lyon 1, Ecologie Microbienne UMR 5557, Villeurbanne, France.	<b>Examineur</b>
<b>Dr. Fiona WALSH</b> , Lecturer, National University of Ireland, Maynooth, Ireland.	<b>Examineur</b>
<b>Prof. Antoine ANDREMONT</b> , Université Paris-Diderot, Paris, France.	<b>Examineur</b>
<b>Prof. Timothy M. VOGEL</b> , Université Claude Bernard Lyon 1, Ecole Centrale de Lyon, Ecully, France.	<b>Examineur</b>
<b>Dr. Pascal SIMONET</b> , DR CNRS, Ecole Centrale de Lyon, Ecully, France.	<b>Directeur de thèse</b>



**Ph.D. THESIS**  
**Presented before:**  
**ÉCOLE CENTRALE DE LYON**  
**For the title of:**  
**DOCTEUR**  
of doctoral school:  
Électronique, électrotechnique, automatique  
**UNIVERSITÉ DE LYON**  
Speciality: Bioengineering

By  
**Joseph NESME**

**Characterization of antibiotic resistance genes abundance and  
diversity in soil bacteria by metagenomic approaches:  
*What is the dissemination potential of the soil resistome ?***

defended on 16 may 2014

In front of the jury composed of:

<b>Dr. William GAZE</b> , Senior Lecturer, University of Exeter Medical School, Truro, UK.	<b>Reviewer</b>
<b>Dr. Edward TOPP</b> , Scientific Researcher, Agriculture and Agri-Food Canada, London ON, Canada.	<b>Reviewer</b>
<b>Dr. Emmanuelle MAGUIN</b> , DU INRA, UMR 1319 MICALIS, Jouy-En-Josas, France..	<b>Examinator</b>
<b>Dr. Sylvie NAZARET</b> , CR CNRS, Université Claude Bernard Lyon 1, Ecologie Microbienne UMR 5557, Villeurbanne, France.	<b>Examinator</b>
<b>Dr. Fiona WALSH</b> , Lecturer, National University of Ireland, Maynooth, Ireland.	<b>Examinator</b>
<b>Prof. Antoine ANDREMONT</b> , Université Paris-Diderot, Paris, France.	<b>Examinator</b>
<b>Prof. Timothy M. VOGEL</b> , Université Claude Bernard Lyon 1, Ecole Centrale de Lyon, Ecully, France.	<b>Examinator</b>
<b>Dr. Pascal SIMONET</b> , DR CNRS, Ecole Centrale de Lyon, Ecully, France.	<b>Thesis advisor</b>



*"Nothing in biology makes sense, except in the light of evolution"*

**Theodozius Dobzansky** (1900-1975 ; Biologiste)

*"[...] Geneticists described evolution simply as a change in gene frequencies in populations, totally ignoring the fact that evolution consists of the two simultaneous but quite separate phenomena of adaptation and diversification."*

**Ernst Mayr** (1904-2005 ; Biologiste)



*"Orion aveugle cherchant le soleil"*

**Nicolas Poussin** (1594-1650; peintre français)

*"Nani gigantum humeris insidentes"*

*"Nous sommes des nains assis sur les épaules des géants"*

**Bernard de Chartres** (vers 1130-1160 ; philosophe)



## Préambule.

Si ma thèse au laboratoire Ampère a débuté en octobre 2010, cela fait déjà longtemps que les questions auxquelles nous avons tenté d'apporter des éléments de réponse au cours de ces trois dernières années, façonnent mon parcours universitaire. Et si dans le cadre de ce projet de thèse, dont le présent manuscrit constitue l'aboutissement, nous nous sommes penchés sur un cas particulier d'écologie microbienne avec un très fort impact sociétal: la résistance bactérienne à des antibiotiques, mon intérêt pour la science du vivant en général, l'adaptation et les processus évolutifs en particulier, trouvent leur origine dans les interminables discussions au sein du giron familial.

La diversité du vivant, c'est à dire le nombre incroyable d'espèces peuplant, ayant peuplé ou qui peupleront la surface terrestre est tout simplement fantastique. La diversité de formes que peuvent par exemple prendre certains organismes a éveillé l'intérêt de générations de naturalistes, Aristote en tête, qui se sont intéressés de tout temps à tenter de les classer selon des critères les plus objectifs possibles. Les découvertes de la structure de la molécule d'ADN par R. Franklin, J.D. Watson & F. Crick et de son rôle dans l'hérédité génétique ont conduit à l'avènement de la biologie moléculaire. Tous les êtres vivants, si l'on veut bien faire exception des protéines de type prion et éviter d'entrer dans la controverse concernant le classement des virus, utilisent le même principe avec l'acide désoxyribonucléique (ADN) qui assure le transfert des caractères à la descendance tout en étant organisé en structures fonctionnelles (les gènes) qui transcrits en acides ribonucléiques (ARN), eux mêmes traduits en protéines, vont définir le phénotype. Cette unicité des molécules de l'hérédité suggère un ancêtre commun: LUCA (*Last Universal Common Ancestor*). C'est cette hérédité commune soumise au phénomène de sélection naturelle décrit par C. Darwin dans "L'origine des espèces » et les mécanismes moléculaires mis en œuvre dans l'évolution des populations conduisant à l'incroyable diversité des espèces qui m'a toujours fasciné.

Dans ce contexte, la microbiologie a l'avantage expérimental indéniable de permettre de travailler sur de grandes populations, pendant de nombreuses générations, avec des organismes simples et abondamment décrits dans la littérature. Outre cet intérêt, les microorganismes ont un rôle prépondérant dans l'écosystème terrestre et un impact quotidien sur l'Homme, qu'il soit positif et alors plus ou moins délibéré ou négatif et alors subi: industrie alimentaire, maladies infectieuses, fertilisation des sols, etc. J'espère donc

que mon implication enthousiaste à comprendre et tester des concepts fondamentaux d'évolution, d'écologie microbienne apportera néanmoins sa contribution à l'édifice commun de la connaissance et sa mise en œuvre pratique. Puisse la lecture de ce manuscrit vous être profitable.

Joseph.

## **Preamble.**

My thesis at Ampère lab started in October 2010, but the questions we modestly tried to answer to in the past three years were already driving since a long time my scholar education choices. If in this work we focused our attention on a very current subject in microbial ecology, with a strong social impact: bacterial resistance to antibiotics, my interest for life sciences and adaptation processes are the roots of this Ph.D. thesis and the present manuscript.

Biological diversity, that is the number of species living, who lived or who will live on the Earth surface, is incredibly high. This life-form diversity has always fascinated naturalists like Aristotle who have tried since ages to classify them with objective criteria. DNA structure discovery by R. Franklin, J.D. Watson & F. Crick and its role in genetic heredity initiated the molecular biology era. Every biological being are using the same molecules, deoxyribonucleic acid (DNA) and ribonucleic acid (RNA), to code their phenotype into a transferable genotype for their offspring. This inheritance molecules unity suggests a common heredity: LUCA (*Last Universal Common Ancestor*). This heredity question, subjected to natural selection described by C. Darwin in "*On the origin of species*", and the molecular mechanisms involved in population evolution, leading to today's incredible living species diversity, always fascinated me.

In this context, microbiology has the undeniable advantage to allow experimentation on large populations of individuals, over many generations, with simple organisms abundantly described in the literature. Beside this fundamental aspect of research, microorganisms have a preponderant role in the ecosystem and daily impact on Human life: food industry, infectious diseases, soil fertilization, etc. I hope then that my enthusiasm to work on fundamental aspects of microbial ecology is nevertheless contributing to Sciences collective effort and will hopefully lead to practical outcomes. May the reading of this manuscript be worthwhile.

Joseph.

# Remerciements.

Personne ne travail jamais seul et nous somme tous des nains, assis sur les épaules des géants. On ne choisit pas d'y grimper, nous y sommes, et vous êtes tous des géants. Vous nous faites partager vos théories, vos expériences, vos résultats et vos échecs scientifiques, certes. Mais aussi et surtout: vos joies, vos peines, votre vie et tout ce pourquoi elle vaut d'être vécue. Je veux ainsi remercier ici toutes ces personnes qui, de près ou de loin, m'ont entourées et supportées pendant ces dernières années. Sans vous, jamais ce travail n'aurait vu le jour, et à le réaliser, j'ai pris un énorme plaisir: **Merci à vous tous !**

Un énorme merci d'abord à toi Pascal, pour m'avoir permis de réaliser cette thèse, mais aussi et surtout, de m'avoir fait autant confiance tout au long de la thèse. Jamais je n'aurais pensé acquérir autant sur le métier de chercheur, au niveau scientifique bien sûr, mais aussi social. Merci aussi d'avoir mis des gants et ta blouse blanche pour réaliser devant mes yeux ébahis un gel Eckhardt vertical. Merci aussi pour ne pas avoir insisté... Merci à Tim et à toi de nous avoir permis de participer à autant de congrès scientifiques internationaux et ainsi, de nous permettre de rencontrer autant de personnes. Surtout: merci à vous deux de m'avoir autant appris aussi sur le plaisir qu'il faut prendre en faisant ce métier, car il est bien trop difficile, et inutile, de s'en passer.

Merci Tim de m'avoir toi aussi fait autant confiance et de m'avoir donné la possibilité de réaliser mon stage de master au labo: c'est parce que je suis rentré que je suis resté. Merci aussi Jean-Michel pour m'avoir fait confiance et permis de travailler avec toi sur le projet Générique, qui à donner lieu à un chapitre complet de la thèse.

Un merci tout particulier à toi Jérémy: tu sais que sans toi, beaucoup des pages suivantes n'auraient pas pu être écrites: Merci !

Merci aussi à vous aussi Flore, Lorrie, Laëtita pour m'avoir aidé, aux cours de vos stages au labo, à réaliser ce travail. Merci Nikola à toi aussi, pour toute ton aide, mais surtout ton sourire et ta joie de vivre: travailler avec toi était très agréable. Bonne chances à vous toutes pour la suite!

Merci beaucoup Seb et merci beaucoup Manu', pour m'avoir tellement aidé avec ce bon vieux musclor, écrit autant de scripts... Merci Seb pour tout ton temps à discuter, débogger mes commandes foireuses et ton sudo... Merci aussi et surtout pour tout le temps que tu as pris pour moi, pour filtrer ces résultats de BLAST, pour me donner tes conseils, pour m'aider et aussi pour discuter, discuter, discuter. Merci Cath', toi aussi pour toutes ces discussions,

tout le son que tu m'as fait découvrir, tous ces moments partagés, et aussi où tu m'a rappelé que la réalité est là, devant nous: merci.

Merci Richard, pour tous ces moments de franche rigolade, pour ta bonne humeur, et pour toute ton aide. Merci à toutes: Sandrine, Isabelle, Laure: ce labo existe grâce à vous! Merci pour toutes vos contributions, pour toutes vos discussions, vos commentaires, votre aide, que ce soit de l'organisation de la vie au labo, jusqu'à me remonter le moral: vous êtes des personnes formidables.

Merci Eliana, merci Amal et merci Alban: nous avons commencé notre thèse et passé les trois dernières années ensemble. Combien de discussions enflammées, sur tout et sur rien ? On vous a suivies toutes les deux, pendant que vous deveniez maman: je garderais de très bons souvenirs de tout ce que l'on a pu échanger. Merci pour toutes ces journées passées avec vous! Merci à toi aussi Jérémy, pour tous ces moments au labo, à la maison, pour les chiuhahas et autres...

Merci aussi à vous tous, Laura, Jean-Seb, Christoph, David pour tous ces bons moments, au labo et ailleurs. Mon constat, on finit par arriver au bout... Profitez-en le plus possible !

Merci à vous tous, au labo, pour votre aide tous les jours. Merci Edith, Edwige, Silvia, Christian, pour toutes ces commandes passées, ces ordres de mission envoyés à la dernière minute et pourtant traités: merci. Merci Alice, pour ta gentillesse, ta bonne humeur, ton sourire. Tu vas me manquer. Merci Marie-Christine, pour toute ton aide apportée pour démêler les ficelles de l'administration de l'Ecole Centrale.

Merci à vous tous aussi qui avez maintenant déménagé hors des murs du laboratoire Ampère, vous avez tout de même été mes très proches collègues pendant plus de deux ans! Merci Céline, Sandra et Cédric, et bonne continuation à vous tous.

Merci à toi, maman, merci à toi, papa: vous avez fait ce que je suis.

Merci Elsa et merci Léo, pour être qui vous êtes.

Lorrie, merci pour tout et merci par avance pour tout. Merci de m'avoir supporté à tes côtés pendant ces mois de doutes, de peines, mais aussi de grandes joies: je me tiens à tes côtés.

Merci au Ministère de la Recherche de m'avoir octroyé ce contrat doctoral pendant 36 mois et merci aussi à toi Christian, pour ton geste pour ces deux derniers mois. Merci à Franck, Guy, François et Gérard, qui m'ont accueilli au sein de, respectivement l'Ecole Centrale, le laboratoire Ampère et sa composante ECL et de l'école doctorale EEA, le tout au sein de l'Université de Lyon.

# Résumé.

Les bactéries de l'environnement et du sol en particulier sont des producteurs actifs de molécules antibiotiques et les composés antibiotiques utilisés en médecine ont pour la plupart été isolés de bactéries saprophytes du sol qui ont elle mêmes développés une variété de mécanismes pour contrer les effets des antibiotiques conduisant à un arsenal de gènes de résistance à des antibiotiques dans l'environnement (ARGD). Une évaluation de l'abondance et de la diversité en terme de gènes de résistance à des antibiotiques a donc été conduite. Pour cette analyse, nous avons compilé 71 jeux de données de séquences d'ADN métagénomique environnementale variées: océans, et identifiés des gènes de résistance pour chacun d'eux. Le sol est confirmé par cet étude comme un environnement extrêmement divers en terme de résistance à des antibiotiques. Cet étude *in silico* a été complété d'abord par une approche en microcosmes visant à étudier les effets soit de pollution soit par des molécules antibiotiques pures, soit par des effluents de ferme utilisés pour la fertilisation des sols. Les microcosmes de sols ont été incubés pendant 6 mois au laboratoire en conditions contrôlées. L'abondance de gènes de résistance à des antibiotiques a été évaluée au cours du temps par PCR quantitative. Une seconde étude visant à évaluer l'impact de la consommation de molécules antibiotiques par une population humaine sur son environnement immédiat, dont le sol, a été entreprise. Le village de Trois-Sauts est situé sur les berges du haut-Oyapock en Guyane Française. Les prescriptions antibiotiques sont très récentes dans cette région et les molécules distribuées ont été précisément répertoriées. Un transect de sol de 3km a été échantillonné chaque 600m afin de vérifier l'existence d'un gradient d'anthropisation entre le village (0m) et les échantillons de forêt les plus distants (3000m). Tous nos résultats confirment la présence à une forte abondance de gènes de résistance à des antibiotiques dans l'environnement, et en particulier dans le sol. Les facteurs à l'origine de la sélection et de la dissémination des gènes de résistance restent cependant difficiles à appréhender dans des environnements aussi complexes. C'est cependant avec une meilleure compréhension des phénomènes conduisant à l'émergence et à la dissémination des gènes de résistance à des antibiotiques au sein des flores pathogènes, depuis leur réservoir environnemental, que nous pourrons agir en vu de préserver les antibiotiques encore actifs aujourd'hui et ceux encore à développer.

# Abstract.

Environmental bacteria and especially soil bacteria are active producers of antibiotic molecules and most drugs used nowadays are isolated from saprophytic soil bacteria and these microorganisms have also evolved numerous resistance pathways leading to an arsenal of Antibiotic Resistance Genes Determinants (ARGD) known as the environmental resistome. A survey of ARGD prevalence is required in order to characterize this natural phenomenon with critical implications in our current infectious diseases management. In order to perform such analysis we compiled a set of 71 metagenomic datasets from various environmental origins: soils, oceans, lakes, human feces, indoor air, etc., and compared their sequences with a database of known antibiotic resistance gene determinants (ARGD). ARGD-annotated reads are found in every environment analyzed confirming their ubiquity. Soil is found to be the richest and shares a large part of ARGD with the human gut microbiome, indicating ARGD transfers between these environments. Experiments using qPCR and metagenomic DNA sequencing on soil samples from two sites with known and distinct antibiotic pollution history were conducted to understand how ARGD abundance and diversity in soil are affected when impacted by antibiotic molecules. The first site is a reference soil from a long-term experiment without history of antibiotic pollution (Rothamsted Park Grass, UK). Soil microcosms are setup with addition of either antibiotic-containing animal manure or pure molecules and incubated for 6 months to monitor changes in ARGD concentration following these perturbations. Our second study-site is a very remote settlement in French Guiana where antibiotics are available since recently and may have impacted the local soil microbial community. Soil samples are taken following a line-transect going from the village (antibiotic source) to 3km deep in the forest in a gradient of human-impact. Our results all confirm prevalence of ARGD in soil at significant abundance but also that ARGD distribution is more correlated to environmental factors such as soil type, microbial taxonomy composition or microcosms incubation conditions than antibiotic molecules exposure in both sites. Pathogens ARGD diversity is far lower than ARGD diversity found in the environment and not all the soil resistome is readily accessible for transfer. In order to characterize the soil mobile gene pool, a strategy is proposed to isolate specifically mobile DNA directly from the environment for sequencing purposes. Better knowledge on the microbial ecology factors limiting ARGD transfers to pathogens may greatly help us reduce the current threat on our limited medical antibiotic molecules resource.

# Table of contents.

Preamble.....	07
Abstract.....	11
Table of contents.....	14
List of figures.....	16
List of peer-reviewed publications.....	18

## **Introduction Générale..... 19**

### **Chapter 1 - The soil resistome: a critical review on origins, ecology and dissemination potential of antibiotic resistance in telluric bacteria.**

Résumé.....	34
Abstract.....	34
Introduction.....	36
On the origin of soil biodiversity and the struggle for existence.....	42
The soil resistome diversity: surviving the chemical arsenal.....	47
An ecosystem vision of antibiotic resistance and production in soil.....	52
References.....	59

### **Chapter 2 - Large-scale metagenomic-based study of antibiotic resistance in the environment.**

Résumé.....	70
Abstract.....	71
Introduction.....	72
Results.....	76
Discussion.....	85
References.....	88

### **Chapter 3 - Antibiotic-amended Rothamsted soil microcosms experiment.**

Résumé.....	92
Abstract.....	92
Introduction.....	94
Experimental procedures.....	98
Results.....	105
Abundance of antibiotics molecules in soil microcosms samples measured by LC/MS-MS.....	105
Impact of amendments on soil prokaryotes community profiles obtained by RISA-PCR.....	106
Antibiotic-resistance genes abundance in soil microcosms measured by qPCR.....	110
Discussion.....	118
References.....	123

## **Chapter 4 - Amazonian resistome: evaluating antibiotic resistance abundance and diversity across a French Guiana forest soil gradient through metagenomic approaches.**

Résumé.....	128
Abstract.....	128
Introduction.....	130
Experimental Procedures.....	164
Results.	
Community taxonomy diversity along the soil line-transect.....	145
Antibiotic resistance abundance and diversity in soil shotgun sequencing.....	149
Relative ARGD quantification in soil samples by qPCR.....	152
Discussion.....	154
References.....	166

## **Chapter 5 - Conclusions on experimental approaches.**

Résumé.....	173
Abstract.....	177

## **Annexes.**

<b>Chapter 1 - Annexes.....</b>	<b>182</b>
<b>Chapter 2 - Annexes.....</b>	<b>184</b>
<b>Chapter 3 - Annexes.....</b>	<b>190</b>
<b>Chapter 4 - Annexes.....</b>	<b>192</b>

---

# List of figures.

## **Chapter 1 - The soil resistome: a critical review on origins, ecology and dissemination potential of antibiotic resistance in telluric bacteria.**

Figure 1 – The resistome: potential for antibiotic resistance

Figure 2 - Schematic representation of the soil ecosystem.

Figure 3 - Stacked histogram showing the relative antibiotic and bioactive

## **Chapter 2 - Large-scale metagenomic-based study of antibiotic resistance in the environment.**

Figure 1 – Relative diversity and abundance of ARGD between metagenomes.

Table 1 – Details for the 50 most prevalent resistance types found in all metagenomes.

Figure 2 – Hierarchical clustering based on Mountford diversity.

Figure 3 - Venn Diagram

Table S1 - Metagenome datasets details.

Figure S1 - Figure S1 - Fifty most prevalent resistance type found in all metagenomes.

## **Chapter 3 - Antibiotic-amended Rothamsted soil microcosms experiment**

Figure 1 - Experimental design used in this study to assess the impact of a controlled antibiotic molecules and manure amendment of a reference soil: Rothamsted, Park Grass soil.

Figure 1 - Principal component analysis of RISA-PCR fingerprinting results obtained for each microcosm condition and time.

Figure 2 - Inter-class principal component analysis of RISA-PCR taxonomic fingerprinting profiles for each microcosms at T0, T1, T3 and T6 months.

Figure 4 - Barplots showing quantification for each qPCR assays relative to gene copy quantification at T0 for each effluents amended microcosms.

Figure 5 - Barplots showing quantification for each qPCR assays relative to gene copy quantification relative to water-amended microcosm for each effluents amended microcosms.

Figure 6 - Barplots showing quantification for each qPCR assays relative to each sample gene copy quantification at T0 for antibiotic molecule amended microcosms.

Figure 7 - Barplots showing quantification for each qPCR assays relative to gene copy quantification in water-amended control for each antibiotic molecule amended microcosms.

## **Chapter 4 - Amazonian resistome: evaluating antibiotic resistance abundance and diversity across a French Guiana forest soil gradient through metagenomic approaches.**

Figure 1 – Maps of French Guiana and the Amazonian National Park.

Figure 2 - Chao1 index rarefaction curve for each distance groups of 16S rRNA datasets.

Figure 3 – Relative phylum abundance in 16 rRNA amplicons sequence reads datasets for all samples (A) and per sample (B).

Figure 4 - Plots showing results of unweighted (A) and weighted (B) UniFrac distances principal coordinate analysis.

Figure 5 - Relative abundance (A) and diversity (B) of antibiotic-resistance gene determinants in French Guiana and Rothamsted soils.

Figure 6-7 - Global abundance of antibiotic-resistance genes found in French Guiana soil samples grouped by resisted antibiotic

Figure 8-9 - Barplot qPCR relative quantification

Figure 10 – Correlation between antibiotic exposure and ESBL *Enterobacteriaceae* among volunteers in 2001, 2006, and 2010.

Figure 11-12 - Relative ARGD abundance detected in environmental shotgun

## **Chapter 2 - Annexes.**

Figure S1 - Figure S1 - Fifty most prevalent resistance type found in all metagenomes.

Figure S3 –ARGD-annotated reads SEED subsystems annotation.

## **Chapter 3 - Annexes.**

Figure S1 - PCR amplification of tetracyclines-specific antibiotic-resistance efflux-pump *tetA(G)* in tetracyclines-amended microcosms and PT2 pigs.

Figure S2 - Oxytetracycline decay measured in Rothamsted soil.

## **Chapter 4 - Annexes.**

Figure S1 – Steps involved in de novo positive primer synthesis.

Figure S2 - Relationship between number of sequence reads and number of detected 16S rRNA SSU reads by BLAST.

Figure S3 - Boxplots representing differences in relative abundance as a function of habitat.

Figure S4 - PCoA plots of unweighted UniFrac distance dissimilarity matrix between forest soil samples.

## **Peer-reviewed publications.**

# List of peer-reviewed publications.

**Article - Nesme, J., and Simonet, P.** *The soil resistome: a critical review on origins, ecology and dissemination potential of antibiotic resistance in telluric bacteria.* **Environmental Microbiology**, (under review).

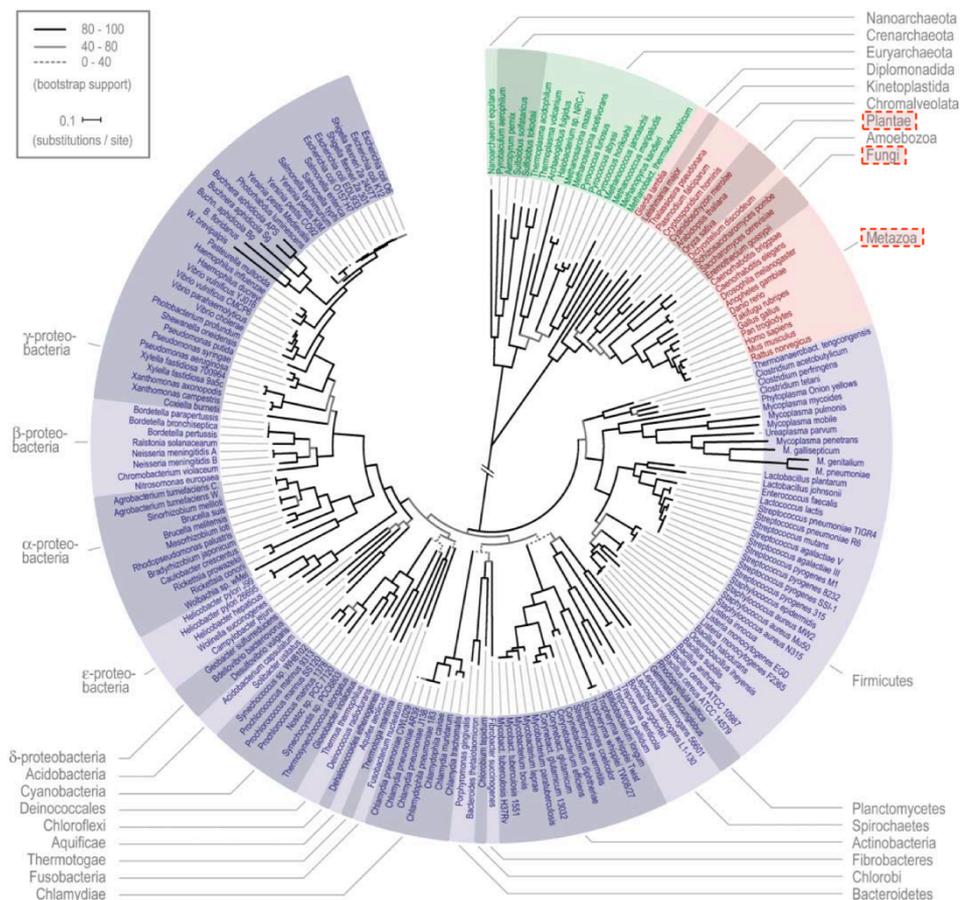
**Article - Nesme, J., Cécillon, S., Delmont, T. O., Monier, J.-M., Vogel, T. M., and Simonet, P.** (2014). Large-Scale Metagenomic-Based Study of Antibiotic Resistance in the Environment. **Current Biology**.

# Introduction Générale.

## De van Leeuwenhoek à Pasteur.

La biosphère terrestre est composée dans sa très vaste majorité d'être vivants unicellulaires et microscopiques, les microorganismes. Ils colonisent tous les environnements et constituent la très vaste majorité des espèces vivantes. Parmi les trois domaines du vivant, Eubactéries, Archées et Eucaryotes, deux sont exclusivement microbiens, les "procaryotes" (i.e. : cellules sans vrai noyau) et les organismes pluricellulaires ne se rencontrent que parmi certains groupes d'espèces d'eucaryotes (Figure 1). La microbiologie est donc la science de la vie invisible à l'œil nu ; observatoire par essence, c'est uniquement lorsqu'il a été techniquement possible de les voir que les microorganismes ont commencé à être considérés. C'est la construction du premier microscope par Antonie van Leeuwenhoek en 1673 qui a permis les premières descriptions d'organismes vivants invisibles à l'œil nu, les microorganismes dévoilant une diversité jusqu'alors inconnue mais ce n'est que deux siècles plus tard et les travaux de Louis Pasteur en 1867 que la microbiologie a pris son véritable essor. Pasteur démontra en effet que des agents biologiques microscopiques, les levures, étaient absolument nécessaires au processus de fermentation et qu'il ne pouvait y avoir de génération spontanée de ces microorganismes comme cela se vérifie si le milieu est stérile où aucune activité biologique n'est alors détectée. Deux siècles ont été nécessaires au développement d'un cadre scientifique rigoureux pour l'étude des microorganismes et l'obtention d'une certaine considération par les autres disciplines scientifiques. Depuis l'origine, la connaissance du monde microbien a été liée aux développements technologiques : microscopie, culture in vitro, physiologie et biochimie, écologie puis biologie moléculaire, bioinformatique, etc.

Pendant de très nombreuses années, l'étude des microorganismes a été strictement limitée à ceux capables d'être cultivés en conditions *in vitro* au laboratoire avant que de récents progrès en biologie moléculaire aient permis de s'affranchir de cette étape favorisant principalement les travaux en écologie microbienne. Les microbiologistes avaient depuis longtemps constaté que de nombreux spécimens étaient récalcitrants à tout développement *in vitro* au laboratoire avec des proportions atteignant 99% dans certains milieux, nécessitant le développement de nouvelles approches pour s'affranchir de ce biais [1].



**Figure 1 – Arbre phylogénétique des génomes séquencés Eukaryotes, Archae et Bactériens.** L'arbre phylogénétique est basé sur la concaténation de l'alignement de 31 familles de protéines universelles et couvre 191 espèces dont le génome est complètement séquencé. En vert, Archées; en rouge, Eucaryotes ; rectangle rouge, groupes contenant des organismes non microscopiques. La branche séparant les Eucaryotes et Archées des Bactéries dans cet arbre non-enraciné a été réduite pour des raisons esthétiques. Adapté de Ciccarelli et al. (2006). Science 311, 1283–1287.

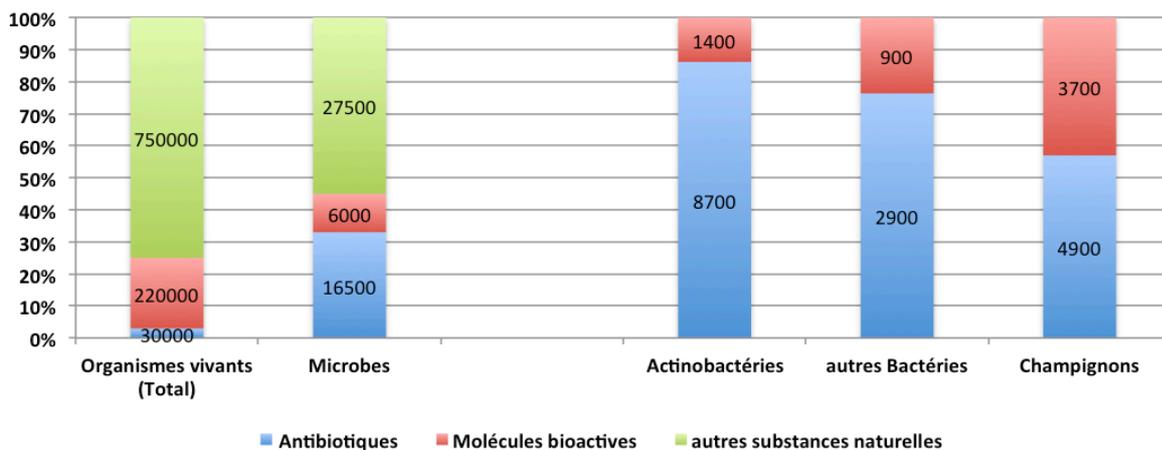
## **Ecologie des microorganismes.**

L'importance des microorganismes dans l'écosystème général terrestre est considérable : les procaryotes représentent une population de plus de  $10^{30}$  cellules pour une masse de plus de  $10^{12}$  tonnes selon certaines estimations [2, 3] et les travaux de physiologie et d'écologie microbienne n'ont depuis cessé de démontrer leur importance fonctionnelle : implication dans les grands cycles géochimiques, photosynthèse, fixation de l'azote atmosphérique, fertilisation des sols par exemple. Les communautés microbiennes sont des systèmes complexes qui imposent de les considérer comme un ensemble lorsque l'on désire comprendre leurs propriétés émergentes [4, 5]. Les travaux sur le microbiote intestinal humain, c'est-à-dire les bactéries colonisant le tube digestif sont, à ce titre, significatifs de la complexité qu'il y a à prendre en compte la problématique dans son ensemble. Le microbiote intestinal a un rôle prépondérant pour assurer la nutrition mais leur influence sur le développement du système immunitaire est maintenant reconnue [6-8]. Comment interpréter alors l'implication de facteurs comme le mode de naissance (voie naturelle ou césarienne) [9, 10], l'alimentation du nourrisson [11], associés aux facteurs génétiques spécifique de l'hôte, sur la composition du microbiote intestinal? Après ce premier exemple d'une influence généralement positive sur leur hôte humain, il faut évoquer ceux responsables des maladies infectieuses qui, quoique beaucoup plus limités en termes de taxons considérés et de diversité n'en ont pas moins un impact considérable sur la santé humaine et ont été responsables de par le passé d'épouvantables épidémies. De ce fait, la microbiologie s'est d'abord développée sur son aspect médical. Les maladies infectieuses sont encore à l'heure actuelle la principale cause de mortalité humaine avec plus de 13,5 millions de morts en 2010, soit un tiers de la mortalité mondiale et jusqu'à 43% dans les pays en voie de développement [12].

## La révolution antibiotique.

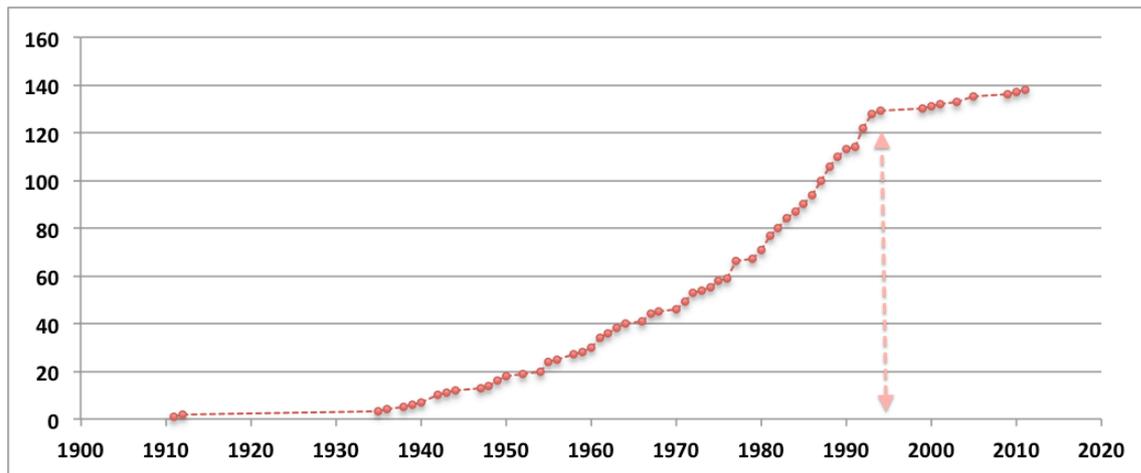
Toutefois, dans les pays développés au moins, les antibiotiques ont permis de considérablement réduire la morbidité et la mortalité des maladies infectieuses. Ils constituent certainement l'un des plus grands succès de la médecine moderne ne serait-ce que par le nombre de vies épargnées. Si les antibiotiques sont aujourd'hui largement disponibles, ce n'est qu'à partir des années 1940 que leur production fut portée à l'échelle industrielle. L'antibiothérapie reste à l'heure actuelle le principal moyen de lutte contre les maladies infectieuses que la résistance croissante des bactéries pathogènes risque toutefois de remettre en question. C'est en 1929 qu'Alexander Fleming publia la démonstration qu'une substance sécrétée par une moisissure, vraisemblablement une souche de *Penicillium rubens* [13], qu'il nomma pénicilline, inhibait spécifiquement la croissance de certaines bactéries [14, 15]. Comme il le dit plus tard à propos de sa découverte: *"J'ai été "accusé" d'avoir inventé la pénicilline. Aucun homme n'aurait pu inventer la pénicilline car elle a été produite, de temps immémorial, par la nature et par une certaine moisissure"*. Fleming avait eu en fait de nombreux prédécesseurs qui entrevoyaient déjà des perspectives thérapeutiques à des relations antagonistes qu'ils avaient observées entre microorganismes [16, 17]. John Tyndall en 1875 dans ses travaux sur les *"vapeurs de putréfaction"* [18] ou Ernest Duchesne en 1897, dans sa thèse de médecine de la Faculté de Lyon [19], avaient par exemple déjà observé les effets antagonistes respectivement de souche de *Penicillium* ou de bactéries sur la croissance d'autres bactéries et entrevoyaient déjà les perspectives thérapeutiques de ces observations. Le terme "antibiotique" est introduit plus tard par Selman Waksman *et al.* en 1941 [20] pour décrire *"toute substance produite par des microorganismes inhibant à faible concentration le développement d'autres microorganismes"* dans la première revue consacrée au réservoir de molécules antibiotiques que constitue le sol [21]. Waksman entrevoyait très clairement les propriétés thérapeutiques qu'on pouvait tirer de ces molécules et, avec ses collaborateurs, a travaillé principalement sur les métabolites produits par des *Actinobactéries* du sol qu'ils avaient

réussi à se faire développer au laboratoire [22]. Ces auteurs ont ainsi largement contribué à développer une pharmacopée riche en molécules antibiotiques: actinomycine (1940), clavacine, streptothricine (1942), streptomycine (1943), griséine (1946) ou néomycine (1948), toutes isolées de bactéries du sol, comme la plupart des antibiotiques découverts à ce jour (Figure 2). Certains de ces antibiotiques, comme la streptomycine et la néomycine sont encore utilisés aujourd'hui en santé humaine. Waksman et les nombreux auteurs qui lui ont succédé ont rapidement constaté la profusion, la diversité et la complexité des mécanismes mis en œuvre par les microorganismes pour synthétiser ces métabolites. Des enzymes, comme les polycétides synthases (ou *PKS* pour *Polyketide synthase* en anglais ; [23-25]) y jouent un rôle considérable et leurs mécanismes d'évolution et de fonctionnement laissent entrevoir des possibilités quasi-illimitées de production de nouvelles molécules [26, 27].



**Figure 2 – Répartition de la production de molécules bioactives d'origine naturelle.**

Répartition des quelques 1.000.000 de différentes molécules substances naturelles produites par des organismes vivants en fonction de leurs propriétés bioactives ou antibiotiques connues et de l'origine taxonomique de leur producteur. La majorité des antibiotiques connus est donc produit par les Actinobactéries. Données adaptées de Bérty et al. (2005). *J. Antibiot.* 58, 1–26.



**Figure 3 – Courbe d’accumulation des antimicrobiens disponible en fonction de leur année d’introduction sur le marché.**

La courbe représente l'augmentation du nombre d'antimicrobiens (anti-infectieux et antibiotiques) autorisés en médecine depuis 1911 et l'introduction du Salvarsan (Arsphénamine ; anti-infectieux) et du Neosalvarsan en 1912. Après la découverte des antibiotiques sulfamidés (1935-1940) qui sont des composés de synthèse, les premiers antibiotiques d'origine naturelle sont commercialisés dès 1942. Ce sont la pénicilline G (benzylpénicilline), isolée de *Penicillium chrysogenum* et la gramicidine S, isolée de *Bacillus brevis* ; deux bactéries du sol. On peut noter un très net ralentissement de l'apparition sur le marché de nouvelles molécules entre la fin des années 80 et le début des années 90. La dernière molécule antibiotique approuvée pour usage médical est la fidaxomicin, isolée de *Dactylosporagium aurantiacum*, une actinobactérie. Données adaptées de Lewis, K. (2013) *Nat Rev Drug Discov* 12, 371–387 et 1. Powers, J. H. (2004). *Clin Microbiol Infect* 10 Suppl 4, 23–31.

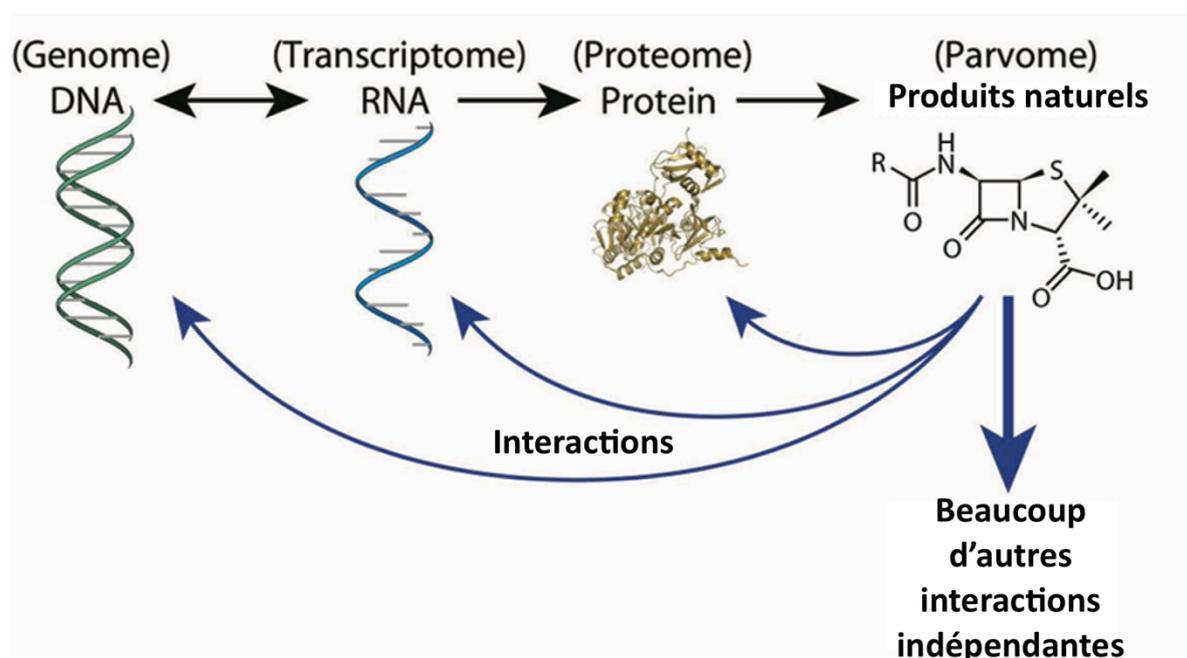
### **Abondance et pénurie des molécules antibiotiques.**

L'essor rapide de l'antibiothérapie est directement lié à son succès avec une recherche de nouvelles molécules antibiotiques très importante et couronnée de succès jusqu'aux années 1980 et une grande efficacité de ces médicaments. Lorsque la pénicilline fut introduite sur le marché, la plupart des bactéries pathogènes y étaient sensibles même si des souches résistantes de Staphylocoques furent très vite découvertes [28]. Il en fut de même pour les autres classes de molécules antibiotiques qui furent ensuite utilisées (Table 1), à nouveau pour la plupart isolées à partir de microorganismes du sol. Depuis la découverte de nouveaux antibiotiques s'est considérablement ralentie (Figure 3) du fait d'une plus grande difficulté technique pour trouver des molécules originales par les technologies classiques et le désintérêt des industriels du domaine pour investir de plus en

plus de moyens pour détecter et développer des médicaments de plus en plus rapidement rendus inopérants par les phénomènes de résistance [29]. De ce fait, les réglementations relatives à l'introduction sur le marché de nouvelles molécules en santé humaine ou même vétérinaire sont devenues plus contraignantes à partir des années 1960 et la création respectivement aux Etats-Unis et en France du Office of New Drugs (U.S Food and Drug Administration) et de la commission d'Autorisation de Mise sur le Marché, en 1959. Tests cliniques garantissant sécurité et efficacité de nouvelles molécules ont été demandées quand, dans le même temps, la dissémination des résistances aux antibiotiques parmi les pathogènes aurait nécessité le recours à de nouvelles molécules tout en essayant d'en limiter l'apparition et la dispersion des mécanismes de résistance. Le développement de nouvelles molécules antibiotiques par l'industrie pharmaceutique apparaît donc aujourd'hui capital même si elle constitue un véritable défi scientifique, technologique, industriel et financier. Si la majeure partie des molécules antibiotiques connues et encore utilisées à ce jour ont leur origine chez la très faible fraction des Actinobactéries du sol cultivées *in vitro* (Figure 2) [30, 31] l'énorme proportion de souches non cultivées laisse augurer que le réservoir de molécules bioactives non encore découvertes est considérable [32, 33]. Les avancées récentes en métagénomique fonctionnelle notamment ont déjà permis d'isoler de nouvelles molécules prometteuses comme la Tetarimycin A, efficace contre les staphylocoques résistants à la méticilline [34]. Ces avancés techniques laissent entrevoir une nouvelle voie pour cribler la diversité de ces molécules et leurs propriétés biologiques [35, 36]. Étant donnée la diversité de structures et de modes d'action déjà observés parmi la relativement faible proportion connue de ces molécules c'est réellement un réservoir quasi-inextinguible de molécules pouvant avoir un intérêt thérapeutique qui est à "portée de main" dans le sol [31, 32].

## Les antibiotiques dans l'environnement.

Les êtres vivants secrètent une quantité impressionnante de petites molécules ou métabolites secondaires (Figure 2) dont certaines ont des propriétés bioactives, comme les antibiotiques. Leur rôle *in-situ* pour les organismes producteurs reste cependant encore mal défini et plusieurs hypothèses ont pu être avancées [37, 38]. Le caractère spécifique de la toxicité peut suggérer une forme de compétition par inhibition sélective dans le cas d'une intense compétition pour la ressource, comme dans la rhizosphère [39]. D'autres auteurs ont mis en évidence une fonction de signal impliqué dans la communication bactérienne [40, 41] ou comme facteur de transcription dose-dépendant [42].



**Figure 4 – Schématisation des interactions entre les molécules de dogme moléculaire et le parvome.**

L'ensemble des petites molécules secrétées par tous les êtres vivants, le parvome, a une influence considérable sur les différentes étapes du dogme moléculaire en agissant comme facteur de réplication, transcription, traduction par exemple. Il a aussi été montré que ces molécules ont de nombreux autres effets comme par exemple leur utilisation comme molécule signal dans le dialogue moléculaire entre cellules. Adaptée de Davies, J. , and Ryan, K. S. (2012). ACS Chem. Biol. 7, 252–259.

Ces substances naturelles sont extrêmement diverses en termes de structures moléculaires ou de propriétés chimiques. Elles partagent cependant des caractéristiques

communes comme leur petite taille et leur activité à très faible concentration qui a incité à définir cet ensemble de molécules par un terme généraliste, le parvome. Le parvome représenterait donc l'ensemble des petites molécules bioactives pouvant avoir toute une variété d'effets sur le métabolisme (Figure 4) [43]. Ces molécules ciblent des processus essentiels de la physiologie des bactéries : division cellulaire, traduction des protéines, transcription de l'ARN, réplication de l'ADN (Table 1). Ces processus sont réalisés par les mêmes familles de protéines chez la plupart des microorganismes, voir des êtres vivants, pour les protéines très conservées liées, par exemple, à la synthèse des acides-nucléiques. On comprend alors que certaines molécules antibiotiques aient un spectre très large. La spécificité de certaines molécules provient alors plus de leur capacité à atteindre une concentration suffisante au niveau de leur cible (ribosome, membrane cellulaire, cytoplasme). Les antibiotiques de la classe des beta-lactames par exemple inhibent un type de paroi spécifique des bactéries à Gram-positif, composée de feuilletts de peptidoglycane liés selon une certaine conformation. Celles qui ne possèdent pas cette cible, les protéobactéries (ex bactéries à Gram-négatif) sont donc naturellement résistantes. En revanche, des molécules antibiotiques ayant une cible universelle, comme les tétracyclines qui bloquent le site A de la sous-unité 30S du ribosome, présentent un spectre d'action uniquement limité par l'affinité de la molécule pour une cible existant dans toutes les cellules. Les tétracyclines sont donc parfaitement capables d'inhiber la traduction des protéines des cellules eucaryotes en se fixant aux sous-unités 70S ou 80S du ribosome mais c'est leur plus forte affinité pour la sous-unité ribosomique 30S des bactéries qui permet de les utiliser comme antibiotiques à des concentrations suffisamment faibles pour limiter les effets secondaires sur l'hôte eucaryote. Les concentrations auxquelles l'Homme utilise les propriétés bactéricides ou bactériostatiques dose-dépendantes des antibiotiques sont de

l'ordre du microgramme et interpellent sur les concentrations habituellement retrouvées dans le sol, infimes, et leurs effets sur la physiologie des bactéries. Différentes hypothèses ont été proposées sur l'origine, l'évolution et le rôle dans l'environnement des bactéries productrices et résistantes à des antibiotiques. Une synthèse en sera proposée dans un premier chapitre avec une emphase particulière sur l'environnement sol qui est le principal réservoir de la diversité bactérienne incluant les producteurs d'antibiotiques et les bactéries qui leur résistent.

Classe d'antibiotique ; exemple	Année de découverte	Année d'introduction	Année d'observation de la résistance	Mécanisme d'action (conséquence physiologique)	Spectre d'activité ou cible
Sulfonamides; prontosil	1932	1936	1942	Inhibe l'activité dihydropteroate synthétase (inhibe la production d'acide folique)	Bactéries Gram+
β-lactams; pénicilline	1928	1938	1945	Inactivation des transpeptidases (inhibe la synthèse de la paroi cellulaire)	Large spectre
Aminoglycosides; streptomycine	1943	1946	1946	Perturbe la fixation des aminoacyl-ARNt au site A du ribosome (traduction de protéines non-fonctionnelles)	Large spectre
Chloramphenicols; chloramphenicol	1946	1948	1950	Inhibe l'activité péptidyl-transférase de la sous-unité 50S du ribosome (inhibe la traduction des protéines)	Large spectre
Macrolides; erythromycin	1948	1951	1955	Inhibe l'activité péptidyl-transférase de la sous-unité 50S du ribosome (inhibe la traduction des protéines)	Large spectre
Tétracyclines; chlortétracycline	1944	1952	1950	Inhibe la fixation des aminoacyl-ARNt au site A du ribosome (inhibe la traduction des protéines)	Large spectre
Rifamycins; rifampicin	1957	1958	1962	Inactive la sous-unité β de l'ARN polymérase (inhibe la transcription)	Bactéries Gram+
Glycopeptides; vancomycin	1953	1958	1960	Immobilise les précurseurs lipopéptidiques peptidyl-D-Ala-D-Ala du PGL* (inhibe la synthèse de la paroi cellulaire)	Bactéries Gram+
Quinolones; ciprofloxacine	1961	1968	1968	Inhibe l'activité des ADN gyrases (inhibe la réplication)	Large spectre
Streptogramines; streptogramine B	1963	1998	1964	Inhibe l'activité péptidyl-transférase de la sous-unité 50S du ribosome (inhibe la traduction des protéines)	Bactéries Gram+
Oxazolidinones; linezolid	1955	2000	2001	Inhibe l'activité péptidyl-transférase de la sous-unité 50S du ribosome (inhibe la traduction des protéines)	Bactéries Gram+
Lipopéptides; daptomycine	1986	2003	1987	Insertion et aggrégation dans la membrane cellulaire induisant une perte du potentiel de membrane (inhibe réplication, transcription et traduction)	Bactéries Gram+
Fidaxomicin (ciblant <i>Clostridium difficile</i> )	1948	2011	1977	Inhibe l'activité ARN polymérase (inhibe la transcription)	Bactéries Gram+
Diarylquinolines; bedaquiline	1997	2012	2006	Inhibe l'activité F1FO-ATPase (inhibe la synthèse d'ATP)	Spectre réduit ( <i>Mycobacterium tuberculosis</i> )

\*PGL : péptidoglycane, composant majoritaire de la paroi cellulaire des bactéries Gram+

**Table 1 - Chronologie de la découverte, de l'introduction et des premières résistances observées pour quinze classes de molécules antibiotiques différentes.**

Adapté de Lewis, K. (2013). Nat Rev Drug Discov 12, 371–387.

## **Origines et mécanismes de la résistance antibiotique chez les pathogènes humains.**

Des infections jadis combattues par des antibiotiques courants sont aujourd'hui incurables car causées par des souches multi-résistantes et font que le problème de la résistance à des antibiotiques est devenu aussi aigu. Staphylocoques résistants à la méticilline (SARM) [44], entérocoques résistants à la vancomycine (VRE), *Escherichia coli* pathogènes multi-résistants, proches de ceux de la flore commensale de l'Homme et possédant des facteurs de virulence accrue, sont devenus courants [45]. Le même problème existe avec les bactéries pathogènes opportunistes, infectant les patients immunodéprimés. Dans ces cas, les bactéries responsables sont plus proches de taxons également retrouvés dans l'environnement extérieur comme: *Acinetobacter baumannii* [46]; *Burkholderia cepacia* [47] ou *Stenotrophomonas maltophilia* [48].

Le problème de la résistance à des antibiotiques en milieu hospitalier serait donc lié à l'acquisition de résistance par des bactéries pathogènes, l'accumulation des phénotypes de résistance au sein d'une population et l'expansion clonale de ces souches multi-résistantes. Dans ce contexte, quel rôle l'environnement extérieur et notamment le sol, peut-il avoir eu dans la transmission vers les pathogènes humains des mécanismes de résistance qui se sont développés pendant les milliards d'années de leur évolution chez les organismes saprophytes telluriques ? Quels mécanismes de transfert, de sélection ont-ils été mis en œuvre pour permettre l'émergence et la sélection de bactéries pathogènes qui étaient demeurées jusqu'à très récemment sensibles à ces molécules ? Ce sont ces questions qui ont motivé ce travail de thèse. Si l'hypothèse d'une origine environnementale des mécanismes de résistance aujourd'hui détectés chez les pathogènes devait être confirmée ce sera d'abord en démontrant l'occurrence et leur diversité dans les microbiotes environnementaux et leurs potentialités à être transférés latéralement. Les progrès récents

en biologie moléculaire, le débit des technologies de séquençage de nouvelles générations et la somme des connaissances recueillies dans le cadre de la surveillance épidémiologique de la résistance antibiotique clinique, sont autant d'éléments conceptuels et de moyens techniques et expérimentaux qui permettent à ces questions très généralistes d'être abordées. Rechercher dans les génomes des bactéries de l'environnement la présence, la diversité des gènes de résistance identifiés chez les souches pathogènes à des antibiotiques est une des premières questions qui peut être posée. Le niveau quantitatif, c'est-à-dire leur représentation au sein des communautés est aussi un élément capital à prendre en compte car il renseigne sur le potentiel adaptatif que ces gènes peuvent conférer à leurs hôtes bactériens mais aussi de leur prévalence dans des microbiotes d'environnements très diversifiés.

### Références.

1. Schloss, P. D., and Handelsman, J. (2005). Metagenomics for studying unculturable microorganisms: cutting the Gordian knot. *Genome Biol* 6, 229.
2. Whitman, W. B., Coleman, D. C., and Wiebe, W. J. (1998). Prokaryotes: the unseen majority. *Proc Natl Acad Sci USA* 95, 6578–6583.
3. Torsvik, V. (2011). Prokaryotic Diversity--Magnitude, Dynamics, and Controlling Factors. *Science* 296, 1064–1066.
4. Baquero, F., and Moya, A. (2012). Intelligibility in microbial complex systems: Wittgenstein and the score of life. *Front Cell Infect Microbiol* 2, 88.
5. Kaneko, K. (2011). Approach of Complex-Systems Biology to Reproduction and Evolution. In *Principles of Evolution, The Frontiers Collection* (Springer Berlin Heidelberg), 241–259.
6. Turnbaugh, P. J., Ley, R. E., Mahowald, M. A., Magrini, V., Mardis, E. R., and Gordon, J. I. (2006). An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* 444, 1027–1031.
7. Lepage, P., Leclerc, M. C., Joossens, M., Mondot, S., Blottiere, H. M., Raes, J., Ehrlich, D., and Dore, J. (2012). A metagenomic insight into our gut's microbiome. *Gut* 62, 146–158.

8. Arumugam, M., Raes, J., Pelletier, E., Le Paslier, D., Yamada, T., Mende, D. R., Fernandes, G. R., Tap, J., Bruls, T., Batto, J.-M., et al. (2011). Enterotypes of the human gut microbiome. *Nature* 473, 174–180.
9. Dominguez-Bello, M. G., Costello, E. K., Contreras, M., Magris, M., Hidalgo, G., Fierer, N., and Knight, R. (2010). Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns.
10. Grönlund, M.-M., Lehtonen, O.-P., Eerola, E., and Kero, P. (1999). Fecal Microflora in Healthy Infants Born by Different Methods of Delivery: Permanent Changes in Intestinal Flora After Cesarean Delivery. *Journal of Pediatric Gastroenterology and Nutrition* 28, 19.
11. Harmsen, H. J. M., Wildeboer Veloo, A. C. M., Raangs, G. C., Wagendorp, A. A., Klijn, N., Bindels, J. G., and Welling, G. W. (2000). Analysis of Intestinal Flora Development in Breast-Fed and Formula-Fed Infants by Using Molecular Identification and Detection Methods. *Journal of Pediatric Gastroenterology and Nutrition* 30, 61.
12. World Health Organization (2010). *World Health Statistics 2010*.
13. Houbraken, J., Frisvad, J. C., and Samson, R. A. (2011). Fleming's penicillin producing strain is not *Penicillium chrysogenum* but *P. rubens*. *mycology* 2, 87–95.
14. Fleming, A. (1929). On the Antibacterial Action of Cultures of a *Penicillium*, with Special Reference to Their Use in the Isolation of *B. Influenzae*. *Br J Exp Pathol.* 10, 226-236.
15. Fleming, A. (1932). On the specific antibacterial properties of penicillin and potassium tellurite. Incorporating a method of demonstrating some bacterial antagonisms. *J. Pathol.* 35, 831–842.
16. Florey, H. W. (1946). The Use of Micro-organisms for Therapeutic Purposes. *The Yale Journal of Biology and Medicine.* 1, 101-117
17. Waksman, S. A. (1925). The Soil Population. *Proceedings of the National Academy of Sciences.* 11, 476-481.
18. Tyndall, J. (1876). Observations on the Optical Department of the Atmosphere in Reference to the Phenomena of Putrefaction and Infection. *Br Med J* 1, 121–124.
19. Duchesne, E. (1897). MD thesis Faculté de Lyon.
20. Waksman, S. A., and Woodruff, H. B. (1941). *Actinomyces antibioticus*, a New Soil Organism Antagonistic to Pathogenic and Non-pathogenic Bacteria. *J Bacteriol.*
21. Waksman, S. A., and Woodruff, H. B. (1940). The Soil as a Source of Microorganisms Antagonistic to Disease-Producing Bacteria. *J Bacteriol* 40, 581–600.

22. Waksman, S. A., Horning, E. S., Welsh, M., and Woodruff, H. B. (1942). Distribution of Antagonistic Actinomycetes in Nature. *Soil Science* 54, 281.
23. Shen, B. (2003). Polyketide biosynthesis beyond the type I, II and III polyketide synthase paradigms. *Curr Opin Chem Biol*.
24. Dutta, S., Whicher, J. R., Hansen, D. A., Hale, W. A., Chemler, J. A., Congdon, G. R., Narayan, A. R. H., Håkansson, K., Sherman, D. H., Smith, J. L., et al. (2014). Structure of a modular polyketide synthase. *Nature* 510, 512–517.
25. Jenke-Kodama, H., and Dittmann, E. (2009). Evolution of metabolic diversity: Insights from microbial polyketide synthases. *Phytochemistry* 70, 1858–1866.
26. Yoon, V., and Nodwell, J. R. (2013). Activating secondary metabolism with stress and chemicals. *J Ind Microbiol Biotechnol* 41, 415–424.
27. Hutchinson, C. R., and Fujii, I. (1995). Polyketide synthase gene manipulation: a structure-function approach in engineering novel antibiotics. *Annu Rev Microbiol* 49, 201–238.
28. Kirby, W. M. M. (1945). Properties of a penicillin inactivator extracted from penicillin-resistant Staphylococci. *Journal of Clinical Investigation* 24, 170–174.
29. Silver, L. L., and Bostian, K. A. (1993). Discovery and development of new antibiotics: the problem of antibiotic resistance. *Antimicrob Agents Chemother*.
30. Waksman, S. A., and Woodruff, H. B. (1942). Selective Antibiotic Action of Various Substances of Microbial Origin. *J Bacteriol*.
31. Bérdy, J. (2005). Bioactive microbial metabolites. *J. Antibiot.* 58, 1–26.
32. Pidot, S. J., Coyne, S., Kloss, F., and Hertweck, C. (2014). Antibiotics from neglected bacterial sources. *International Journal of Medical Microbiology* 304, 14–22.
33. Smith, T. (2000). Antibiotics from soil bacteria. *Nat Struct Biol* 7, 189–190.
34. Kallifidas, D., Kang, H.-S., and Brady, S. F. (2012). Tetarimycin A, an MRSA-active antibiotic identified through induced expression of environmental DNA gene clusters. *J. Am. Chem. Soc.* 134, 19552–19555.
35. Cheng, G., Hu, Y., Yin, Y., Yang, X., Xiang, C., Wang, B., Chen, Y., Yang, F., Lei, F., Wu, N., et al. (2012). Functional screening of antibiotic resistance genes from human gut microbiota reveals a novel gene fusion. *FEMS Microbiol Lett*.
36. Davies, J. (2011). How to discover new antibiotics: harvesting the parvome. *Curr Opin Chem Biol* 15, 5–10.
37. Aminov, R. I. (2009). The role of antibiotics and antibiotic resistance in nature. *Environ Microbiol* 11, 2970–2988.

38. Martinez, J. L. (2008). Antibiotics and Antibiotic Resistance Genes in Natural Environments. *Science* *321*, 365–367.
39. Haas, D., and Défago, G. (2005). Biological control of soil-borne pathogens by fluorescent pseudomonads. *Nat Rev Microbiol* *3*, 307–319.
40. Liu, Z., Wang, W., Zhu, Y., Gong, Q., Yu, W., and Lu, X. (2013). Antibiotics at subinhibitory concentrations improve the quorum sensing behavior of *Chromobacterium violaceum*. *FEMS Microbiol Lett* *341*, 37–44.
41. Silva, A. J., Benitez, J. A., and Wu, J.-H. (2010). Attenuation of bacterial virulence by quorum sensing-regulated lysis. *J Biotechnol* *150*, 22–30.
42. Fajardo, A., and Martínez, J. L. (2008). Antibiotics as signals that trigger specific bacterial responses. *Curr Opin Microbiol* *11*, 161–167.
43. Davies, J., and Ryan, K. S. (2012). Introducing the parvome: bioactive compounds in the microbial world. *ACS Chem. Biol.* *7*, 252–259.
44. Thompson, J. M., Gündoğdu, A., Stratton, H. M., and Katouli, M. (2013). Antibiotic resistant *Staphylococcus aureus* in hospital wastewaters and sewage treatment plants with special reference to methicillin-resistant *Staphylococcus aureus* (MRSA). *J Appl Microbiol* *114*, 44–54.
45. Boucher, H. W., Talbot, G. H., Bradley, J. S., Edwards, J. E., Gilbert, D., Rice, L. B., Scheld, M., Spellberg, B., and Bartlett, J. (2009). Bad Bugs, No Drugs: No ESKAPE! An Update from the Infectious Diseases Society of America. *Clin Inf Dis*.
46. Lortholary, O., Fagon, J. Y., Hoi, AB, Mahieu, G., and Gutmann, L. (1998). Colonization by *Acinetobacter baumannii* in intensive-care-unit patients. *Infection Control and Hospital Epidemiology* *19*, 188–190.
47. Compant, S., Nowak, J., and Coenye, T. (2008). Diversity and occurrence of *Burkholderia* spp. in the natural environment. *FEMS microbiology ....*
48. Looney, W. J., Narita, M., and Mühlemann, K. (2009). *Stenotrophomonas maltophilia*: an emerging opportunist human pathogen. *Lancet Infect Dis*.

# Chapitre 1.

## ***The soil resistome: a critical review on origins, ecology and dissemination potential of antibiotic resistance in telluric bacteria.***

### **Résumé.**

Le sol est un très grand réservoir de diversité microbienne et la majorité des composés antimicrobiens encore utilisés aujourd'hui en médecine humaine ou vétérinaire ont été isolés à partir de microorganismes du sol. L'hypothèse darwinienne d'une "course à l'armement" entre producteurs d'antibiotiques et bactéries résistantes est souvent citée pour expliquer l'origine et la diversité des déterminants génétiques de la diversité bactérienne. L'exposition environnementale aux molécules antibiotiques et l'abondance de gènes de résistance retrouvés dans l'environnement ne sont cependant pas systématiquement corrélées. De nombreux autres facteurs contribuent à l'émergence, la sélection et la dissémination des gènes de résistance à des antibiotiques dans l'environnement. Le sol est un environnement complexe et hétérogène composé d'un vaste spectre de niches écologiques différentes, à une échelle spatiale restreinte, et abritant de nombreuses espèces différentes. Une telle diversité biologique dans une si petite dimension spatiale favorise considérablement les échanges de matériel génétique par le biais des transferts horizontaux de gènes. Ce chapitre a pour objectif de passer en revue les différentes hypothèses écologiques permettant d'expliquer la prévalence des ARGD dans le sol.

# Chapter 1.

## ***The soil resistome: a critical review on origins, ecology and dissemination potential of antibiotic resistance in telluric bacteria.***

### **Abstract.**

Soil is a large reservoir of microbial diversity and the majority of antimicrobial compounds used today in human and veterinary healthcare have been isolated from soil microorganisms. The Darwinian hypothesis of an “arms-shields race” between antibiotic producer and resistant strains is often cited to explain antibiotic resistance gene determinants (ARGD) origins and diversity. ARGD and antibiotic molecule exposure are, however, not systematically linked and many other factors contribute to resistance gene emergence, selection and dissemination in the environment. Soil is complex and heterogeneous with a broad spectrum of different ecological niches varying in terms of abiotic and biotic factors concentrated in a relatively small scale and harboring different bacterial species. Such a large genetic diversity at such a small spatial scale favors the exchange of genetic material by means of horizontal gene transfer (HGT) that may lead to introduction of ARGD into pathogen genomes, threatening antibiotic therapies. The objective of this chapter is to review our current knowledge on the extent of the soil resistome abundance and diversity, the different ecological hypotheses explaining their high prevalence in soil and how they are transferred to pathogens.

## Introduction

Since their introduction in medicine to treat millions of people, antibiotics have considerably reduced infectious diseases casualties and may account for the biggest medical revolution in the 20th century. Penicillin, one of the first widely used antibiotics, was isolated in 1929 by Alexander Fleming [1] and introduced onto the market in 1938 [2], in time to significantly reduce the number of deaths caused by infected wounds during World War II [3]. The antibiotics golden age lasted from the 1940s to the 1990s with an exponential growth of antibiotic molecules discovered and introduced into the medical market. Most of antibiotics are from environmental origin, either bacterial (mostly Actinobacteria such as *Streptomyces* spp.) or fungal with, for most of them, a soil origin [4, 5]. Their immediate success was directly linked to pathogen sensitivity in the early age of antibiotic therapy that dramatically reduced infectious diseases mortality and morbidity. The increasing resistance of pathogenic strains seen in hospitals however very quickly hindered this success: first observation of penicillin-resistant *Staphylococcus* spp. occurred as early as 1945 [6] and seems unavoidable when introducing new antibiotics molecules [2, 7]. According to the World Health Organization (WHO), antibiotic resistance is defined by "an increase in the minimum inhibitory concentration of a compound for a previously sensitive strain" (WHO Fact sheet N°194). This definition of antibiotic resistance is ad-hoc for clinical microbiology but neglects concepts of intrinsic [8] or resistance genes naturally occurring in the environment [9, 10]. In a broader view of microbial ecology, we still have to elucidate the origins and selection processes of the diverse mechanisms of antibiotic resistance that we are currently observing in clinical and environmental settings. Findings on antibiotic resistance genes ecology *in situ* could have considerable impact on clinical

microbiology as well as being a fantastic context for studying traits-adaptation in bacterial population. Are resistance genes a result of rapid adaptation through gene shuffling or mutations [11, 12]? Are they acquired through horizontal gene transfer (HGT) from non-pathogenic microbial lineages [13, 14]? Is this transfer-to-host chain of events impacted by other anthropogenic impacts than sole antibiotic molecule selective pressure [15]? Urbanization factors such population density and building density are for example shown to be positively correlated with integron-related genes frequently implied in transfer of antibiotic resistance genes in soil samples [16]. What is the fitness-benefit for producing antibiotic molecules, or resisting it, in natural environments with intense competition such as soil [17, 18]? What is the extent of ARGD diversity in nature and what is the share of currently known ones [19]?

Data already gathered in environmental microbiology studies shows that ARGD have been found in diverse environmental samples, such as soil [20-22], ocean [23, 24] and pristine environments [25, 26]. These results support an abundant environmental reservoir hypothesis and especially in soil where most antibiotic producers are extracted. Studying resistant environmental bacteria with high-throughput methods, Forsberg *et al.* showed that a significant amount of resistance genes are shared between the soil and the gut microbiota [27]. The transfer of antibiotic resistance genes from the environment to pathogen genomes is one of the most probable hypotheses of resistance genes emergence in the clinical environment [10, 28]. While not indicating the direction of transfer this work shows that commensal and pathogenic bacteria share antibiotic resistance genes with soil bacteria. Tracking directions of gene transfer is certainly one of the most intense research topics in evolutionary microbial ecology. In this respect, studying the environment reservoir of antibiotic-resistance genes and its links to the clinic is of crucial importance and

emphasizes on the need for thorough environmental surveys of this situation in natural environments but also in environments impacted by human-activities. Experimental studies tracking confidently antibiotic-resistance gene reservoirs are still scarce and most often limited to cultivable microorganisms and well-defined classes of resistance genes. Such studies, while tracking down the most probable route from the environmental reservoir to pathogen hosts and subsequent dissemination are limited by 1) prior knowledge of resistance in the clinic and will therefore miss any genes not similar to known antibiotic resistance genes and 2) exhaustive cultivation of both the clinical and environmental hosts of the same gene, a risky bet considering the wide majority of not yet cultured microorganisms. These studies have so far enabled to trace the evolutionary and hosts route of several types of antibiotic-resistance genes such as: plasmid-mediated *qnrA* (quinolones resistance) [29]; *bla<sub>OXA</sub>* [30]; *bla<sub>CTX-M</sub>* [31, 32] and *AmpC* types of beta-lactamases or tetracyclines ribosomal-protection resistance genes [33, 34]. In these studies, authors are searching for a discrepancy between gene localization (i.e. plasmid/other mobile element versus chromosome) a habitat (i.e. pathogens or commensal versus soil or water-borne) for a given gene sequence (i.e. an allele of the *bla<sub>CTX-M</sub>* genes family) in several hosts, considering these examples as signature of, respectively, clinical or environmental antibiotic-resistance.

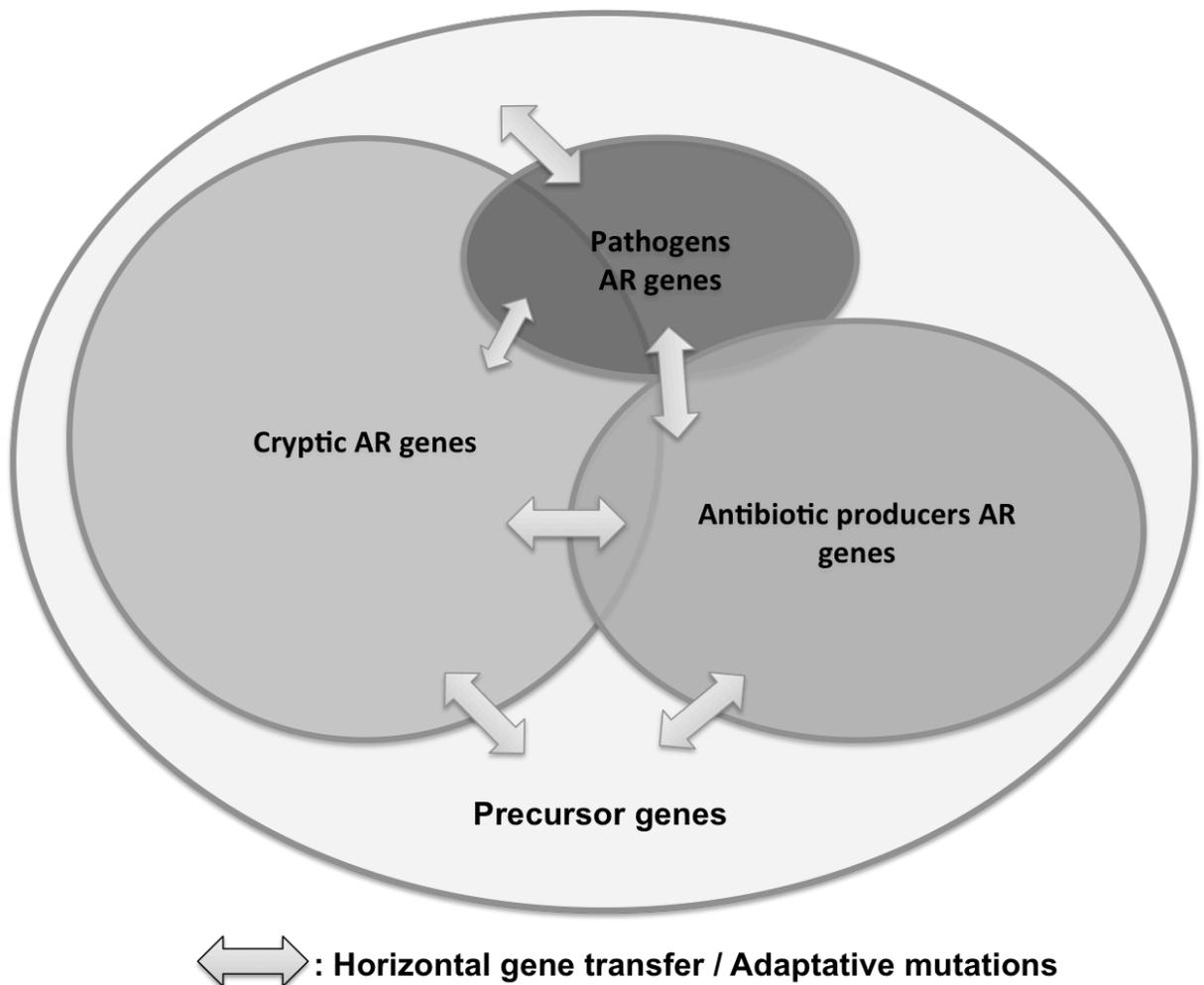
The fast pace of antibiotic resistance increase in clinical environment suggests a pre-existing pool of antibiotic resistance genes in “natural” environmental reservoirs and a transferability of these genes: every known bacterial pathogen has developed resistance to one or more antimicrobial molecule used in the clinic. In this context, a major question of clinical microbiologists is to track reservoirs of resistance genes and limits their dissemination to hosts where their expression is problematic, such as in pathogens and

understanding factors driving selection and dissemination of environmental antibiotic resistance genes will considerably enhance our ability to prevent dissemination of antibiotic-resistance genes in the clinic [35], especially when considering that a better management of currently still-efficient antibiotic molecules used in veterinary and human medicine is critically needed [36, 37]. It is admitted that horizontal gene-transfer is a major mechanism of bacteria adaptation, notably in the case of antibiotic resistance [38-40].

Deciphering the different chains of gene transfer events leading to antibiotic-resistance phenotype acquisition in previously sensitive pathogenic bacteria is a major challenge for microbial ecology. Untangling this network requires to carefully take into account its different aspect from gene source to drastic host's microbiome population selection following antibiotic therapy. Obviously, natural selection is not acting on these traits with the same fitness cost depending on gene host or environmental factors. These theoretical statements can now be more thoroughly tested experimentally using recent advances in molecular biology, bioinformatics and increased sequence data-throughput. While these technologies reaffirmed clearly that complex environmental matrices such as soil harbor a diversity far more important than previously expected [41, 42], combining data provided with these new tools and previously developed statistical algorithms for community ecology enable a vast field of experimentations for microbial ecologists. In the context of antibiotic-resistance, a trait selected for resistance in the clinic may be costly for the same organism in the environment in the absence of high antibiotic-molecule selective pressure, for example when such traits are encoded on high-copy number plasmids, but strongly selected in the clinic in presence of antibiotics [43, 44]. Gene fluxes between different hosts are strongly shaping microbial populations and have considerable effects on general microbial ecology and physiology, allowing for example access to new niches upon

acquisition of a new functional trait [45]. Antagonistic relationships driven by antibiotic-molecules secretion and resistance are also a strong cohesion factor between related *Vibrio* spp. population, as previously shown [24] that may explain for a part existence of sympatric speciation in related bacteria sharing a common habitat but different ecological niches (i.e.: *ecotypes*, see Frederick M. Cohan: [46]). Both from their practical and fundamental aspects studies on antibiotic resistance are of the utmost interest and promise to bring considerable new experimental data considering the recent technological advances cited above (i.e.: molecular biology, bioinformatics and increased sequence data-throughput). Soil is the original habitat of most currently known bacteria producing antibiotic molecules and certainly one of the richest habitats in terms of microbial diversity. Methods to study environmental bacteria in complex ecosystems such as soil are still limited and only recent advances in sequencing technology and molecular biology have started to fill the gap. A majority of environmental microbes resist cultivation with standard laboratory techniques and the few cultured ones do not accurately represent overall biodiversity. These uncultured microorganisms remained mostly a black box until the metagenomic revolution [47]. The principle of metagenomics (from “*meta-*” = all and genomics) *sensu lato* is to use standard molecular techniques on DNA extracted directly from environmental samples, such as a gram of soil, alleviating cultivation biases. Development of metagenomics jointly with next-generation sequencing technologies alleviated a number of previous critical technical limitations. Details on metagenomics methods developed to study antibiotic resistance in soil are available elsewhere [48]. In the second chapter of thesis results on ARGD abundance and diversity in environmental shotgun metagenomic sequencing are presented and is an example of how new technical concepts can be used to improve our understanding of this old topic in microbiology.

Taking into account the environmental aspect of antibiotic resistance has advanced our knowledge of the environmental resistome ecology [19, 49, 50]. In contrast to a strict clinical definition of antibiotic resistance, the resistome includes every resistance gene: precursors with low resistance or affinity for antibiotic molecules; cryptic resistance genes with no or low expression of resistance phenotype in their hosts, and clinical resistance genes such as extended spectrum  $\beta$ -lactamases (ESBL) conferring high antibiotic concentration resistance. As a result, all subgroups are more or less overlapping, illustrating importance of host and genetic context for resistance phenotype expression, short term evolution by mutation, recombination and horizontal gene transfer (Figure 1). Most antibiotic molecules target generalist bacterial functions such as cell wall synthesis (e.g.:  $\beta$ -lactams, vancomycin) or protein translation (e.g.: tetracyclines, aminoglycosides), so producers strains need to shield themselves against their own antibiotics in order to avoid suicide [51], which conveniently explains the origin of antibiotic resistance genes in telluric microbes. Yet, the role of antibiotics as inhibitors is not so clear in many cases. From response to antibiotic production to core cell physiological functions, antibiotic resistance genes are playing a number of different roles in soil saprophytic bacteria yet to determine in most cases. Experimental studies trying to address these alternative hypotheses question about the cause and consequences of antibiotic resistance and antibiotic production diversity in soil, in a global view of soil microbial ecology will be reviewed in this chapter in an effort to better understand the different aspects of antibiotic production and resistance genes origins, selection and dissemination in this environment. While antibiotic-resistance dissemination is alarming, these new technologies now enable microbiologists to tap into a considerable resource of new molecules from soil microbes recent results in this topic will also be reviewed.



**Figure 1 – The resistome: potential for antibiotic resistance.**

The resistome encompasses all genes coding antibiotic resistance (AR) related proteins or similar protein likely to evolve toward potent antibiotic resistance. This encompasses problematic AR genes identified in pathogenic bacteria, AR genes from antibiotic producers, cryptic AR genes (i.e. a resistance gene with no obvious association to antibiotic resistance) which may be or not expressed but could confer resistance in another genetic context for example and AR precursor genes coding protein with low level of affinity or resistance to antibiotic molecules. Significant parts of those different genes subset are overlapping due to sequence similarities and are most likely sharing a common evolutionary history. Genes can of course switch from one state to another through either horizontal gene transfer (HGT) in another host or point mutation or recombination. Adapted from Wright et al. Nat. Mic Review. (Wright,G.D, 2010)

### **On the origin of soil biodiversity and the struggle for existence.**

Understanding the specific features of a complex environment such as the soil is critical to determine how all the components of this biota, from cells to molecules, are interacting (Figure 2). Biologically, soil is an incredibly rich environment, both in terms of

microbial abundance and species diversity. A gram of soil may contain from  $10^6$  to  $10^9$  bacterial cells of  $10^3$  to  $10^6$  different bacterial species, depending on its location (i.e. forest or prairie soil, tropical or polar soil) [41, 42]. These numbers may vary when considering the method used [52-54] or how bacterial species are defined [55, 56]. Whatever the definition of bacterial species, that is still technical (e.g.: OTU nucleotide similarity of taxonomic markers or DNA/DNA hybridization experiment [55, 57]), environmental microbiologists acknowledge that soil harbors a tremendous prokaryotic diversity. Soil is formed with aggregates of solid organic matter (SOM) and mineral particles of different sizes maintained together by plant roots, fungi mycelium and bacteria forming sticky exopolysaccharides (EPS) biofilms [58]. Spatially, cell density varies in soil, which is a rather lacunar environment. Most bacterial cells are living in biofilm colonies, packed in small pores formed between mineral particles or in the rhizosphere under influence of plant root exudates, where nutrients are more abundant and the soil biotic capacity is significantly increased (Figure 1) [59]. In addition to local biome climate, a whole range of factors will modify a general model as the one presented on Figure 1. A large range of physicochemical gradients forms in soil through physical diffusion. Environmental gradients are crucial for ecological niche differentiation, which increases the range of different habitats in soil. Ecological theory suggests a model where ecosystem species richness is dependent on the number of different niches available [60]. Thus, the structure of soil composed of heterogeneous aggregates maintains a tremendous bacterial genetic diversity [61-63]. Microbial cells are struggling for existence in this habitat where nutrients distribution is mostly scarce and patchy, trying to colonize the microscale habitat best suited to their ecotype [46]. Specific environmental factors (e.g.: particle size, pH, water availability, vegetation cover etc.) seem correlated to bacterial [64] or fungal [65] species diversity

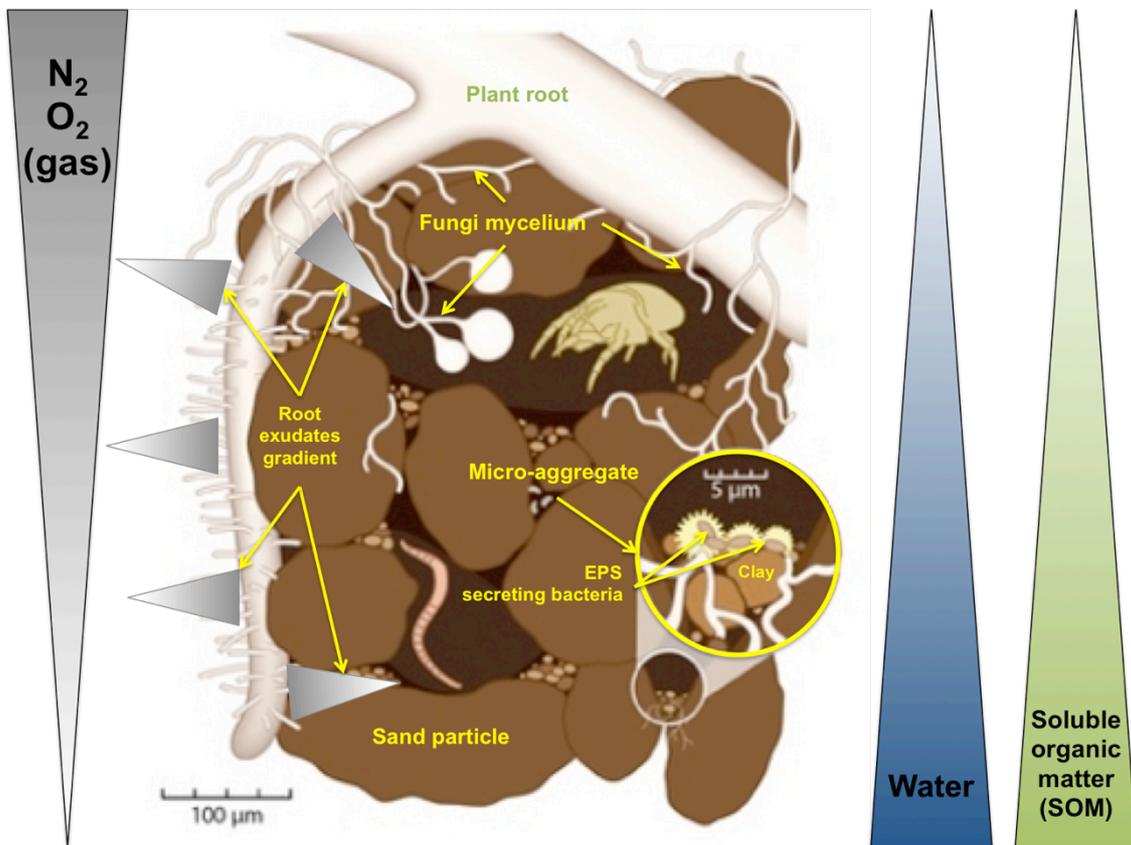
distribution in soil. Can interference competition self-explain origins of all environmental antibiotic producers and resistant strains? Is co-evolution over generations explaining antibiotic resistance and production genes ubiquity and diversity in soil environments? Is this antibiotic resistance genetic diversity a problem in term of healthcare? Soil is an open and connected ecosystem interacting constantly with all compartments of the biosphere, which makes exchange of genetic material possible between ecologically distinct lineages usually found in other ecosystems, such as commensal bacteria inhabiting the human gut, for example [66].

### **The soil parvome: living in the chemical arsenal.**

If antibiotic resistance is a common feature for a large diversity of soil microbes, a proportional diversity of antibiotic resistance genes can be expected in bacterial populations confronted to such pressure in soil. Pathogens' antibiotic resistance genes are sharing striking nucleotide identities with genes coding the same antibiotic-resistance genes in the environment [21, 22, 67]. Soil is likely to be the main reservoir of ARGDs transferred to human pathogens or commensal bacteria. The same resistance enzymes, and even allelic variants, of antibiotic resistance genes are found both in soil and pathogenic bacteria [22, 26, 68, 69]. Genes encoding common clinical resistance enzymes, such as  $\beta$ -lactamases, exhibit more diversity in a soil sample than the same genes in the clinical environment [23]. These results are consistent with the hypothesis of a shared resistome between soil bacteria and pathogens, however, limited in diversity: most gene allelic variations shared between pathogens are far less diverse than those found in environmental samples for the same gene. While limited, this shared resistome has important implications for the future of antibiotics and drug development. Are all environmental resistance genes equally

susceptible to jump into pathogen genomes and thwart present and future antibiotic therapies in a matter of time? Soil is the natural habitat for many microbial antibiotic producers. Actinobacteria and other soil microorganisms provided most of the currently used pharmaceutical drugs, which were developed from the 1940s to 1980s [2]. High-throughput screening of actinomycetes and related microbes has led to isolation of a wide range of antibiotic molecule classes with very different properties and mechanisms of action. Davies (2012) [70] describes these bioactive compounds as the soil “parvome”. The parvome is consisting of all small molecules produced by living organisms as secondary metabolites (from “*parv-*” = small, “*-ome*” = group) [70]. These products encompass, but are not limited to, all secondary metabolites and serve various functions for their producers, from microbes to plants, insects and animals [70, 71]. In the context of soil and antibiotic production, the soil parvome seems to be incredibly abundant and still largely unknown [72]. Judged from what has been currently investigated among the whole tree of life, a major part of bioactive compounds has been identified in microorganisms. Amongst 16,500 different antibiotic molecules produced by microbes: 30% are synthesized by fungi; 53% by actinobacteria and 17% by other bacteria phyla (Figure 3) [5]. These microorganisms secrete an even greater repository of bioactive compounds, such as antibiotics. Compounds isolated from microorganisms during the golden age of antibiotic molecules discovery all come from the “easy to cultivate” part of the overall microbial antibiotic producers’ diversity (Figure 3). Few new natural antimicrobial drugs have been discovered recently and the antibiotic discovery pipeline is undoubtedly drying out and needs to be fixed [73]. The already diverse library of discovered antibiotic compounds soil microbes implies an even more diverse production *in situ*, especially when keeping in mind the low amount of cultured bacteria from environmental samples by traditional laboratory practices. In this respect, new

concepts such as global metagenomic approaches or single-cell sorting, coupled with the development of high-throughput sequencing technologies can now try to circumvent this well-admitted cultivation bias for the study of antibiotic resistance in the environment.



**Figure 2 - Schematic representation of the soil ecosystem.**

Soil is a complex and heterogeneous matrix with varying physicochemical conditions. From high to low nutrient or oxygen concentrations soil is structured in macro- and micro-aggregate structured by biotic and abiotic factors at different scale, forming a complex trophic network. While varying between soils, generally atmospheric gases are forming a diffusion gradient with highest concentrations found in topsoil while water is drained quickly, leaching soluble nutrients in a vertical gradient from top to bottom. Plants roots and fungi mycelium are colonizing spaces between mineral particles of varying size (from clay  $\leq 4\mu\text{m}$  to coarse pebbles  $< 16\text{mm}$ ), giving soil its structuration. Rhizosphere is defined as a soil under influence of plants roots exudates. This is in soil where the highest concentration of bacteria and organic nutrients are found and as a consequence, soil is a rather lacunar ecosystem in term of bacteria population with large areas with a low density of bacterial cells and local hotspots of high cell density and high metabolic activity, often associated with high nutrients concentration. Bacterial populations varying in abundance and all struggling to find a suitable ecological niche are colonizing soil aggregates of different sizes and physico-chemical conditions, favoring a high overall genetic diversity in this environment. As a consequence, soil is considered as a significant reservoir of diverse antibiotic resistance and antibiotic production genes, among a whole range of other bioactive molecules. Adapted from Fortuna, A. (2012) *The Soil Biota*. *Nature Education Knowledge* 3(10):1.

### **The soil resistome diversity: surviving the chemical arsenal.**

From a broad perspective, this large chemical cocktail suggests a reservoir of interacting genes with matching proportions. In terms of microbial ecology, soil microorganisms interact with (and eventually resist) an impressively diverse chemical arsenal, whose function is yet to be determined in most cases but may likely be very old. This emphasizes the ecological significance of antibiotic resistance. Experimental studies also show that potent antibiotic resistance genes can be detected in very old environmental samples. Exploring a 30,000 years old Alaskan permafrost soil sample, D'Costa *et al.* (2011) extracted DNA molecules with a strong nucleotide homology to antibiotic resistance genes isolated from clinical settings. These genes match common resistance genes, such as *Streptomyces* spp.  $\beta$ -lactamases *bla*, several *tetM*-related tetracyclines ribosomal protection proteins, and more surprisingly, a complete *vanHAX* vancomycin resistance operon that is similar to current clinical resistance genes. The "old" *vanA* gene was cloned and expressed in a heterologous *Escherichia coli* host to produce and purify the old VanA protein. Enzymatic and structural studies were performed to assess that ancient VanA protein possess a similar D-alanine-D-lactate ligase activity, as modern VanA protein that confers resistance to gram-positive bacteria in the clinic. Bhullar *et al.* (2012) [25] also found antibiotic-resistant bacteria in secluded environment, in the Lechuguilla Cave (Carlsbad Caverns National Park, USA). These caves are sealed from the outside world for 4 to 7 million years (*i.e.* before *Homo sapiens* appeared about 2.3 million years ago). A plethora of bacterial strains from this human-less environment resisted as many as 14 different antibiotic molecules. Some resistance mechanisms identified in these strains are similar to those for known antibiotic resistance genes, but many of the genes (and their mechanisms of resistance) identified in this study were unknown beforehand [25]. The diversity of

antibiotic resistance genes found in very different bacterial lineages could then be explained as countermeasures to inhibitory concentration of antibiotics secreted by producers. This arms-shield hypothesis could explain the origin of clinical antibiotic resistance genes as corroborated by experimental studies [74, 75], but other hypotheses could also explain the high diversity and abundance of resistance genes in natural environments. These such as modulators of bacterial quorum-sensing [76, 77] or a common origin with essential bacterial components [68, 78], explaining their high prevalence in environmental bacteria genomes. If antibiotic resistance genes are indeed solely due to antibiotic molecule inhibition, these antibiotic molecules are quite potent (*i.e.* often found at inhibitory concentrations) in the soil ecosystem. Thus, resistance genes would significantly increase fitness in order to be effectively selected [79]. For example, producers-resistant and resistant-only strains of *Streptomyces* spp. undergo a coevolution likely to be the result of an intense competition resulting in sympatric speciation by mutual exclusion [80]. While being non-exclusive, the role of antibiotics as inhibitors of microbial growth in soil has been experimentally studied. Soil microbiologists have long since tried to use bio-control agents of phyto-pathogens after observing competitive interactions occurring in the rhizosphere, where competition for resources is intense. Some rhizobacteria, such as *Pseudomonas fluorescens* or *Bacillus subtilis*, produce antibiotic molecules targeting soil-born plant-pathogens also colonizing the rhizosphere. Therefore, antibiotic molecules and their subsequent resistance genes are likely to be found in rhizospheric soil. Observations of antagonistic interactions that often involve plant-pathogens are reviewed by Haas & Défago [81] and demonstrate that antibiotic molecules can be potent in soil and even more so when competition for resources occurs at least in the case of phytopathogens-rhizobacteria interactions. How can samples be representative of the soil structural composition when micro-aggregates range from  $10^{-6}$

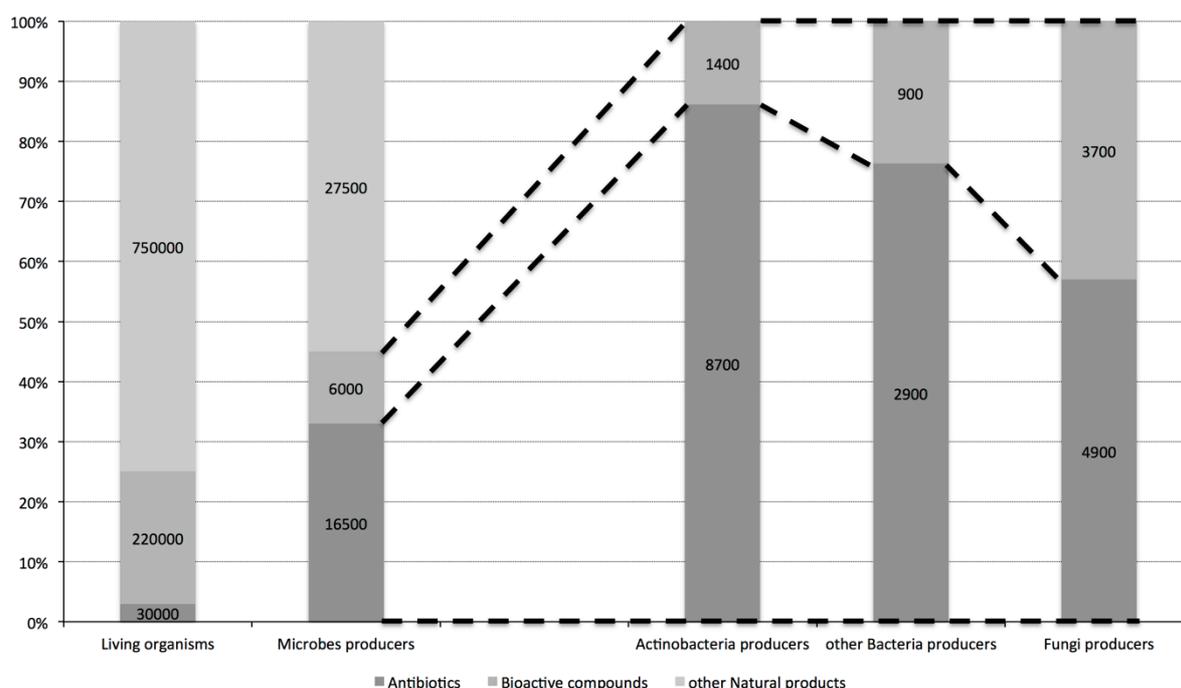
to  $10^{-3}$  meters and sampling is done at the  $10^{-2}$  meter, for DNA extraction, to kilometer scale, for bio-geographic studies? New concepts and methods to carefully model the 3D structure of bacterial population at the micro-aggregate scale are currently under development and will hopefully enhance our understanding of the micrometer-scale interactions occurring in soil [63, 82]. Notably, and considering that soil DNA extraction is a major source of variations [53, 83], new mathematical models are designed for a better and more even sampling between sites, making inter-samples comparisons less subject to errors [84, 85].

#### **Dosing small molecules in soil at the soil micro-aggregate scale.**

Another very important point to consider is the concentration at which antibiotic molecules and microbes interact. Measuring complex and small molecules such as antibiotics in very heterogeneous matrix like soil is subject to technical limitations and we are currently unable to accurately measure trace amounts of antibiotics. A relatively recent compilation of studies measuring *in situ* concentration and degradation kinetic of pharmaceutical compounds, including antibiotics, in soil or water environments are extensively reviewed elsewhere [86-88]. These papers report that the highest concentrations are found either in production animal feces (pig and chicken farms), at antibiotic molecules production sites, or near hospital sewers outlets. In natural soils (*i.e.*: not obviously exposed to antibiotics pollution), antibiotic molecules do not seem to accumulate and concentrations measured are in the  $\mu\text{g}\cdot\text{kg}^{-1}$  range. This is  $10^2$  to  $10^3$  fold less concentrated than the  $\mu\text{g}\cdot\text{g}^{-1}$  range of concentrations used for medical antibiotic therapy but vary a lot depending on molecule properties and soil mineral composition [89-91]. Antibiotics can be either tightly immobilized on soil particles or washed away by water

runoff depending on their chemical properties [91, 92]. In addition to soil sorption, antibiotics may be degraded by biotic or abiotic mechanisms at varying rates [93-95]. All these factors explain why antibiotics are hardly detected in soil and do not seem to accumulate significantly in a bioactive form, except in specific highly polluted environments, such as those described before and where the highest concentrations are measured [96, 97]. The most commonly used analytical technique for antibiotic molecules quantification in soil relies on liquid chromatography separation coupled with mass spectrometry (LC-MS) and tandem mass-spectrometry (LC-MS/MS). This method is rather sensitive and can be used for a large range of compounds. Antibiotics are largely variable in chemical structure and extraction yield from the soil matrix may vary considerably with the same solvent. Therefore, sample preparation prior to analysis is critical and results are highly influenced by this step [98]. Besides solvent choice, antibiotics are showing a broad range of adsorption behaviors on small mineral particles, such as negatively charged clay. Moreover, measuring chemicals may be counter-intuitive in some cases as a detectable compound may not be accessible for bacteria in a bioactive form. Researches now focus on modeling the *in situ* effects of chemical perturbations on bacterial communities with *in vitro* devices, such as microfluidic flow-cells modeling the soil porous environment [99, 100] or exposure to chemical gradients [101, 102] in an attempt to decompose each factor's effects on either pure or mixed microbial cultures. Such devices are still in active development and therefore have not yet been used in thorough environmental microbiology studies. Another way around is to make use of a proxy to measure antibiotic-concentration. With the help of a reporter strain, Bahl et al. (2004) were able to measure that 0.4% of the total tetracycline ingested by rats is bioavailable to their gut microbiota [103]. These studies are still scarce and few have been successfully applied to environmental samples [104]. They have,

however, the clear advantage of measuring bioavailability instead of total antibiotic concentrations, a determinant factor in selection of antibiotic-resistance genes. Quantification of antibiotic concentrations at the soil micro-aggregate scale where antibiotic antagonist interactions occur would be necessary in order to understand at which concentrations, and selective pressure, most bacteria are exposed in soil. The dose-response curve observed in antibiotic therapy, where aim is limitation of microbial growth, may not be the same at low concentration at which bacteria sense antibiotics in situ. The same antibiotic molecule has been shown to act on the same strain both as an activator and an inhibitor of transcription depending on its relative concentration [105, 106].



**Figure 3 - Stacked histogram showing the relative antibiotic and bioactive compounds percentage in known natural products from living organisms.**

Stacked histograms are showing the proportion of known bioactive and antibiotics molecules (i.e.: compounds registered in BNPD; <http://naturalprod.ucsd.edu/>) produced by (from left to right): all living organisms from the three domain of Life; microorganisms only; Actinobacteria; other Bacteria and Fungi. All data are from Bérdy, J. (2005). Bioactive microbial metabolites. *Journal of Antibiotics* 58, 1–26.

### **Antibiotic resistance and production: an ecosystem.**

Despite the fundamental objective of environmental microbiology to describe soil microorganism complex ecology and physiology *in situ*, knowledge gathered will hopefully help us to preserve current and future antibiotic efficiency. Antibiotic resistance is as an ecosystem feature prevalent in the environment and especially in soil based on our current knowledge. Defining a common core of functions conferred by antibiotic resistance genes or the role of antibiotic molecules in nature seems irrelevant when looking at the incredible diversity of mechanisms already described. What we are commonly labeling as antibiotics, however, share common general features such as being small molecules and having interactions with essential bacterial components. Could antibiotic molecules and their corollary antibiotic resistance genes (or rather interacting genes) serve different purposes than interference competition in the environment, such as metabolism regulation or quorum sensing/quenching? The diversity of molecular mechanisms displayed by microorganisms is tremendous and we only know a fraction of it. Several alternative evolutionary scenarios relying on other selection effects beside competition have been suggested as origins for at least part of the antibiotic resistance gene diversity. A well-studied example of secondary adaptation leading to antibiotic resistance is multi-drug resistance (MDR) efflux pumps. MDR efflux pumps form a membrane-anchored protein family whose role is to regulate cell homeostasis by selective diffusion of compounds in or out the cytoplasm. Some pumps have a very specific affinity for antibiotics and can confer a high level of resistance to their hosts, such as tetracycline resistance conferred by specific efflux pumps [107, 108]. Drug-specific and multi-drug resistance efflux pumps are very similar in structure to other efflux pumps and may have, therefore, been selected for cations such as Na<sup>+</sup> or K<sup>+</sup>, used to exchange protons, regulate cell homeostasis, excrete

toxic metabolic compounds and even trigger cell communication. Bacterial metabolism and physiology are quite plastic and a single gene sequence may give rise to a broad spectrum of different functions depending on the cell context. A gene that was previously selected for homeostasis regulation or heavy metal excretion may become well suited in the clinical context and be selected rather for antibiotic resistance rather than its ancestral bacteria physiological feature.

Another example of secondary-adaptation following gene-duplication is the penicillin binding protein and  $\beta$ -lactamases protein families that display respectively a transpeptidase and a  $\beta$ -lactam ring hydrolyzing activity.  $\beta$ -lactams (*i.e.*: penicillin and its derivatives) are the most widely used antibiotic drugs in human medicine. They inhibit bacterial growth by sequestering transpeptidase enzymes known as penicillin-binding proteins (PBP) that prevent the amino-acid crosslinking of peptidoglycan layers and cell-wall formation, and thereby, cause improper division. PBP (the antibiotic target) and  $\beta$ -lactamases (the resistance genes) share sequence similarity at the amino-acid level, same protein fold, same structural motifs and active sites as suggested by structural biologists. Only slight divergences between the two protein families led to considerable diversification of function. Still, PBP and  $\beta$ -lactamases are related from an evolutionary perspective and ancient penicillin-binding proteins are thought to have given rise to those of  $\beta$ -lactamases [78, 109]. After a PBP gene duplication releases the selection pressure on a essential locus,  $\beta$ -lactamases may have been selected for cell-wall related functions, such as peptidoglycan recycling, rather than resistance to  $\beta$ -lactam inhibition [110, 111]. Furthermore, the  $\beta$ -lactamase genetic diversity encountered in the clinic seems quite old, as observed by Song et al. (2005) in 10,000-year-old deep-sea sediments samples from Edison seamount [23]. This imply that the selection for effective antibiotic resistance property preceded the

antibiotic medical era and most genetic innovation in antibiotic resistance genes occurred before Human antibiotic usage. These observations clearly favor a transfer hypothesis of antibiotic resistance genes from soil to commensal and pathogenic bacteria. Antibiotic resistance would then in many cases be more a spandrel (or by-product) of bacterial genetic evolution. Antibiotic resistance is an advantageous trait that may often be a secondary adaptation to very specific environmental conditions where antibiotic concentrations produce stressful conditions, such as clinical environment or a patient's gut under antibiotic therapy. This is most notably true when considering that most widely disseminated antibiotic-resistance genes are not always the most potent (*i.e.*: conferring resistance to the highest minimum inhibitory concentration for a given antibiotic). Antibiotic resistance and production genes are found in environments where degradation and dilution of secreted compounds are likely to occur at a rapid rate, such as in high water content environments. In this context, secreting antibiotics for local competition seems to be a rather costly behavior. Yet, marine specific clades of Actinobacteria can be recovered from ocean samples and are potent antibiotic producers [112, 113], which indicates ubiquity of antibiotic production and resistance genes in all environments. Biosynthesis of antibiotics and many other secondary metabolites is performed by polyketide synthases, which is a system conserved in bacteria, fungi, plants and some animal lineages. Polyketide synthases (PKS) are specific multifunctional enzymes similar to fatty acid biosynthesis pathway enzymes. Genetic modules encoding PKS are organized in clusters where homologous recombination is likely to occur and generate an impressive diversity of possible enzymes, and therefore, bioactive compound synthesis pathways [114, 115]. PKS genetic diversity associated with their multifunctional properties enables soil microorganisms to produce an incredible diversity of different antibiotics molecules [116].

### **Dissemination of antibiotic resistance: from horizontal gene transfer to superbugs.**

As discussed previously, the basis of antibiotic resistance related studies is the catastrophic increase in antibiotic resistance in the clinic that leads to a significant increase in infectious disease mortality. Yet, we have shown with many examples that antibiotic resistance is a common ecological feature in soil bacteria. Experimental observations and ecological arguments indicate that antibiotic resistance genes are ancient and ubiquitous and represent a reservoir of transferable genetic material. From an epidemiology perspective, antibiotic resistance is more and more prevalent in the clinic. Yet, the problematic pathogens involved are restricted to few phyla when compared to the overall prokaryotes diversity, such as the ESKAPE pathogens: *Enterococcus faecium*, *Staphylococcus aureus* belong to Firmicutes; *Enterobacteriaceae* spp. (i.e. *Klebsiella*, *Escherichia* species), *Acinetobacter baumannii* and *Pseudomonas aeruginosa* belong to gamma-Proteobacteria (i.e. Gram negative bacteria) [117, 118]. Resistance to antibiotic treatment in the clinic is much higher than the rate of discovery of new antibiotic resistance genes or variants in pathogen genomes, coherent with a dissemination of selected antibiotic resistance genes (clonal expansion after selection) rather than recruitment of new antibiotic-resistance genes from environmental reservoirs. ARGDs themselves do not seem to be strongly selected in the environment but on the other hand disseminate at an incredibly fast pace in clinical settings and become, such as the New-Dehli metallo- $\beta$ -lactamase *bla*<sub>NDM-1</sub>. This gene was first identified in a New-Delhi hospital in India in 2009 [119] and spread worldwide in very short amount of time [120]. Shortly after this first identification, it was detected worldwide in hospitals, indicating its prevalence in *Enterobacteriaceae* pathogenic strains [121-123]. The *bla*<sub>NDM-1</sub> gene has also been detected in *Acinetobacter baumannii* isolates, which is a soil dwelling bacteria responsible of opportunistic nosocomial infections. Comparative genomics

analysis of *bla<sub>NDM-1</sub>* locus in different genomes suggests that this gene has emerged by chromosomal recombination in *Acinetobacter* strains [124, 125]. This gene is often found on plasmids belonging to different incompatibility groups such as IncN and IncA/C in *Enterobacteriaceae* strains [126, 127]. Such plasmids have a broad-host range and can be replicated in different bacterial lineages. IncN and IncA/C are also conjugative and possess all the genes required for their transfer. Genes carried by these genetic elements can be disseminated at a rapid rate in distant bacterial lineages. Bacteria possess myriads of molecular mechanisms developed for rapid adaptation to their changing environment from accessory genetic elements such as plasmids [40, 128] to transposons [129, 130], integrons [131, 132] or hypermutator phenotypes [133]. Each of these different mechanisms may have implications at different population levels, but have all been shown to mediate antibiotic resistance gene transfer. All these observations are emphasizing on the duality (first and second orders) of natural selection processes involved in antibiotic resistance genes dissemination in pathogens [134]. Gene *bla<sub>NDM-1</sub>* is now spread worldwide due to a succession of distinct HGT events: emergence of the *bla<sub>NDM-1</sub>* gene by genomic recombination within *Acinetobacter* spp. population; transfer on a mobile genetic integrative element, such as a transposon; transposon insertion on a broad-host range plasmid; and dissemination in different bacterial lineages including pathogenic *Enterobacteriaceae*. This evolutionary scenario is more than plausible based on above-cited experimental studies data on this gene determinant of antibiotic resistance, tracking its dissemination from an environmental reservoir (*Acinetobacter* spp.) to efficient dissemination on *Enterobacteriaceae* plasmids, where this gene confers significant fitness advantage. This illustrates a distinction between efficient resistance genes (*i.e.*: conferring resistance above a certain MIC) and efficient dissemination of antibiotic resistance genes,

leading to an increased incidence of multi-drug resistant pathogens [135]. This is by co-selection, for example with heavy-metal resistance genes located in the same genetic structure, that antibiotic resistance genes can be efficiently disseminated without specific antibiotic-resistance selection [136]. Like antibiotic resistance genes, mobile genetic elements are found in most microbial genomes. These molecular mechanisms participate in the generation of genetic diversity and horizontal gene transfer between and within microbe populations. Such events cannot be prevented, but their rate is quite low at the population level because any mutant emerging individually within a population is likely to be outcompeted by neighboring cells or lost due to genetic drift before being fixed in the population.

## **Concluding remarks.**

Most of antibiotic resistance increase in the clinic is more likely to be the result of successful mutant clonal expansion rather than massive exchanges between soil and clinical resistomes. This is the increasing problem of antibiotic in clinical environments that lead microbiologists to question the origin of antibiotic resistance and hypothesized the pre-existence of an environmental resistome. This assumption is supported by the very rapid acquisition of resistance mechanisms by pathogens, previously sensitive, upon introduction of new antibiotics. At first, only exposure of pathogens to antibiotics was considered as a selective factor eliciting antibiotic-resistance genes and increasing their dissemination rate. This only since the 1970's that the environment and its reservoir of antibiotic-producers started to be considered as a source of antibiotic-resistance genes for pathogens and plays a major role in their dissemination, even among pathogens populations. The soil was the first environment to be considered, due to origins of antibiotics molecules, isolated from soil

microorganisms and when it was found that producers themselves encode antibiotic-resistance genes identical to those found in the clinic. Major results concerning the antibiotic-resistance problem were obtained in the field of environmental microbiology with notably with help of molecular biology, and PCR-related applications, directly on DNA extracted from soil samples: such as RISA-PCR, T-RFLP, PCR-DGGE, PCR-AFLP, quantitative PCR and ultimately, sequencing. These methodological and conceptual revolutions unraveled the incredible diversity of soil microorganisms diversity was orders of magnitude higher than previously expected, based on cultivation or microscopic magnification of soil samples results. While very interesting from a fundamental aspect, such discoveries should lead to the re-examination of theories on the origins of antibiotic-resistance, notably considering genetic diversity of antibiotic-resistance genes. Sequence-based analysis show that a whole continuum of antibiotic-resistance genes diversity, coherent with species phylogeny, is present in the environment. Antibiotic-resistance confers resistance, obviously, but antibiotic molecules targets are ubiquitous in bacterial physiology and the molecular biology dogma: replication, transcription, translation or cell partitioning. Interactions with antibiotics-resistance genes seems much more complex than previously thought in a strict arms/shield selection race.

Bacteria evolved a large range of mechanisms to exchange genes, remodel their genome and ultimately adapt to changing environmental conditions. These molecular mechanisms, such as integrons, super-integrons, transposons, plasmids, general transduction via virus particles, etc (i.e.: mobile genetic elements); but also insertion

sequences (IS), homologous recombination or gene duplication & random genetic drift are multiple forms of bacterial "*boorish but permissive sex*<sup>1</sup>", generating genetic diversity.

Bioremediation of soils, discovery of new enzymatic activity with industrial applications and hopefully new pharmaceuticals compounds resources to fight infectious diseases: these thematic will also largely benefits from results obtained by studies on soil adaptive potential and genetic diversity and hopefully help to tackle the problem of antibiotic-resistance in the clinic; while also contributing to considerably enhance our understanding of fundamental microbial and general ecology.

## References

1. Fleming, A. (1929). On the antibacterial action of cultures of a penicillium, with special reference to their use in the isolation of *B. influenzae*. *Br J Exp Pathol* *10*, 226.
2. Lewis, K. (2013). Platforms for antibiotic discovery. *Nat Rev Drug Discov* *12*, 371–387.
3. Quinn, R. (2013). Rethinking antibiotic research and development: World War II and the penicillin collaborative. *Am J Public Health* *103*, 426–434.
4. Waksman, S. A., and Woodruff, H. B. (1940). The Soil as a Source of Microorganisms Antagonistic to Disease-Producing Bacteria. *J Bacteriol* *40*, 581–600.
5. Bérdy, J. (2005). Bioactive microbial metabolites. *J. Antibiot.* *58*, 1–26.
6. Plough, H. H. (1945). Penicillin resistance of *Staphylococcus aur...* [Am J Clin Pathol. 1945] - PubMed - NCBI. *American journal of clinical pathology*.
7. Powers, J. H. (2004). Antimicrobial drug development--the past, the present, and the future. *Clin Microbiol Infect* *10 Suppl 4*, 23–31.
8. Cox, G., and Wright, G. D. (2013). Intrinsic antibiotic resistance: Mechanisms, origins, challenges and solutions. *Int J Med Microbiol* *303*, 287–292.

---

<sup>1</sup> X. Nesme, personal communication: here sex is defined in respect to Ernst Mayr biological species concept (BSC), which is that a species consists of populations of organisms that can reproduce with one another and that are reproductively isolated from other such populations. Sex is "permissive" (HGT allow cross-species genetic recombination) but "boorish" in prokaryotes (these forms of sex transfer only chunks of genome).

9. Lewis, I. M. (1929). Bacterial Antagonism with special reference to the effect of *Pseudomonas fluorescens* on spore forming bacteria of soils. *J Bacteriol* 17, 89–103.
10. Martinez, J. L. (2008). Antibiotics and Antibiotic Resistance Genes in Natural Environments. *Science* 321, 365–367.
11. Shibata, T., Nishinaka, T., Mikawa, T., Aihara, H., Kurumizaka, H., Yokoyama, S., and Ito, Y. (2001). Homologous genetic recombination as an intrinsic dynamic property of a DNA structure induced by RecA/Rad51-family proteins: a possible advantage of DNA over RNA as genomic material. *Proc Natl Acad Sci USA* 98, 8425–8432.
12. Bürger, R. (1999). Evolution of genetic variability and the advantage of sex and recombination in changing environments. *Genetics* 153, 1055–1069.
13. Ochman, H., Lawrence, J. G., and Groisman, E. A. (2000). Lateral gene transfer and the nature of bacterial innovation. *Nature* 405, 299–304. Available at: <http://www.nature.com/nature/journal/v405/n6784/abs/405299a0.html>.
14. Stokes, H. W., and Gillings, M. R. (2011). Gene flow, mobile genetic elements and the recruitment of antibiotic resistance genes into Gram-negative pathogens. *FEMS Microbiol Rev* 35, 790–819.
15. Gaze, W. H., Zhang, L., Abdousslam, N. A., Hawkey, P. M., Calvo-Bado, L., Royle, J., Brown, H., Davis, S., Kay, P., Boxall, A. B. A., et al. (2011). Impacts of anthropogenic activity on the ecology of class 1 integrons and integron-associated genes in the environment. *ISME J* 5, 1253–1261.
16. Nardelli, M., Scalzo, P. M., Ramírez, M. S., Quiroga, M. P., Cassini, M. H., and Centrón, D. (2012). Class 1 integrons in environments with different degrees of urbanization. *PLoS ONE* 7, e39223.
17. Starikova, I., Harms, K., Haugen, P., Lunde, T. T. M., Primicerio, R., Samuelsen, Ø., Nielsen, K. M., and Johnsen, P. J. (2012). A trade-off between the fitness cost of functional integrases and long-term stability of integrons. *PLoS Pathog.* 8, e1003043.
18. Andersson, D. I., and Levin, B. R. (1999). The biological cost of antibiotic resistance. *Curr Opin Microbiol* 2, 489–493.
19. D'Costa, V. M., Griffiths, E., and Wright, G. D. (2007). Expanding the soil antibiotic resistome: exploring environmental diversity. *Curr Opin Microbiol* 10, 481–489.
20. Donato, J. J., Moe, L. A., Converse, B. J., Smart, K. D., Berklein, F. C., McManus, P. S., and Handelsman, J. (2010). Metagenomic analysis of apple orchard soil reveals antibiotic resistance genes encoding predicted bifunctional proteins. *Appl Environ Microbiol* 76, 4396–4401.
21. Allen, H. K., Moe, L. A., Rodbumrer, J., Gaarder, A., and Handelsman, J. (2009). Functional metagenomics reveals diverse beta-lactamases in a remote Alaskan soil.

ISME J 3, 243–251.

22. Demanèche, S., Sanguin, H., Poté, J., Navarro, E., Bernillon, D., Mavingui, P., Wildi, W., Vogel, T. M., and Simonet, P. (2008). Antibiotic-resistant soil bacteria in transgenic plant fields. *Proc Natl Acad Sci USA* 105, 3957–3962.
23. Song, J. S., Jeon, J. H., Lee, J. H., Jeong, S. H., and Jeong, B. C. (2005). Molecular characterization of TEM-type beta-lactamases identified in cold-seep sediments of Edison Seamount (south of Lihir Island, Papua New Guinea). *J Microbiol*.
24. Cordero, O. X., Wildschutte, H., Kirkup, B., Proehl, S., Ngo, L., Hussain, F., Le Roux, F., Mincer, T., and Polz, M. F. (2012). Ecological populations of bacteria act as socially cohesive units of antibiotic production and resistance. *Science* 337, 1228–1231.
25. Bhullar, K., Waglechner, N., Pawlowski, A., Koteva, K., Banks, E. D., Johnston, M. D., Barton, H. A., and Wright, G. D. (2012). Antibiotic resistance is prevalent in an isolated cave microbiome. *PLoS ONE* 7, e34953.
26. D'Costa, V. M., King, C. E., Kalan, L., Morar, M., Sung, W. W. L., Schwarz, C., Froese, D., Zazula, G., Calmels, F., Debruyne, R., et al. (2011). Antibiotic resistance is ancient. *Nature* 477, 457–461.
27. Forsberg, K. J., Reyes, A., Wang, B., Selleck, E. M., Sommer, M. O. A., and Dantas, G. (2012). The Shared Antibiotic Resistome of Soil Bacteria and Human Pathogens. *Science* 337, 1107–1111.
28. Wright, G. D. (2010). Antibiotic resistance in the environment: a link to the clinic? *Curr Opin Microbiol* 13, 589–594.
29. Poirel, L., Rodríguez-Martínez, J. M., Mammeri, H., Liard, A., and Nordmann, P. (2005). Origin of plasmid-mediated quinolone resistance determinant QnrA. *Antimicrob Agents Chemother* 49, 3523–3525.
30. Poirel, L., Héritier, C., and Nordmann, P. (2004). Chromosome-encoded amblar class D beta-lactamase of *Shewanella oneidensis* as a progenitor of carbapenem-hydrolyzing oxacillinase. *Antimicrob Agents Chemother* 48, 348–351.
31. Cantón, R., González-Alba, J. M., and Galán, J. C. (2012). CTX-M Enzymes: Origin and Diffusion. *Front Microbiol* 3, 110.
32. Humeniuk, C., Arlet, G., Gautier, V., Grimont, P., Labia, R., and Philippon, A. (2002). Beta-lactamases of *Kluyvera ascorbata*, probable progenitors of some plasmid-encoded CTX-M types. *Antimicrob Agents Chemother* 46, 3045–3049.
33. Aminov, R., and Mackie, R. (2007). Evolution and ecology of antibiotic resistance genes. *FEMS Microbiol Lett* 271, 147–161.
34. Chee-Sanford, J. C., Aminov, R. I., Krapac, I. J., Garrigues-Jeanjean, N., and Mackie, R. I. (2001). Occurrence and diversity of tetracycline resistance genes in lagoons and

- groundwater underlying two swine production facilities. *Appl Environ Microbiol* 67, 1494–1502.
35. Perry, J. A., and Wright, G. D. (2013). The antibiotic resistance “mobilome”: searching for the link between environment and clinic. *Front Microbiol* 4, 138.
  36. Ghafur, A. K. (2010). An obituary—on the death of antibiotics. *J Assoc Physician India*.
  37. Piddock, L. J. (2011). The crisis of no new antibiotics-what is the way forward? *Lancet Infect Dis*.
  38. Djordjevic, S. P., Stokes, H. W., and Roy Chowdhury, P. (2013). Mobile elements, zoonotic pathogens and commensal bacteria: conduits for the delivery of resistance genes into humans, production animals and soil microbiota. *Front Microbiol* 4, 86.
  39. Fluit, A. C., and Schmitz, F.-J. (2004). Resistance integrons and super-integrons. *Clin Microbiol Infect* 10, 272–288.
  40. Touchon, M., Charpentier, S., Pognard, D., Picard, B., Arlet, G., Rocha, E. P. C., Denamur, E., and Branger, C. (2012). Antibiotic resistance plasmids spread among natural isolates of *Escherichia coli* in spite of CRISPR elements. *Microbiology* 158, 2997–3004.
  41. Torsvik, V. (2011). Prokaryotic Diversity--Magnitude, Dynamics, and Controlling Factors. *Science* 296, 1064–1066.
  42. Whitman, W. B., Coleman, D. C., and Wiebe, W. J. (1998). Prokaryotes: the unseen majority. *Proc Natl Acad Sci USA* 95, 6578–6583.
  43. Angebault, C., and Andremont, A. (2012). Antimicrobial agent exposure and the emergence and spread of resistant microorganisms: issues associated with study design. *Eur. J. Clin. Microbiol. Infect. Dis*.
  44. Cantón, R., and Morosini, M.-I. (2011). Emergence and spread of antibiotic resistance following exposure to antibiotics. *FEMS Microbiol Rev* 35, 977–991.
  45. Hibbing, M. E., Fuqua, C., Parsek, M. R., and Peterson, S. B. (2009). Bacterial competition: surviving and thriving in the microbial jungle. *Nat Rev Microbiol* 8, 15–25.
  46. Cohan, F. M. (2001). Bacterial species and speciation. *Systematic biology*.
  47. Schloss, P. D., and Handelsman, J. (2005). Metagenomics for studying unculturable microorganisms: cutting the Gordian knot. *Genome Biol* 6, 229.
  48. Monier, J.-M., Demanèche, S., Delmont, T. O., Mathieu, A., Vogel, T. M., and Simonet, P. (2011). Metagenomic exploration of antibiotic resistance in soil. *Curr Opin Microbiol* 14, 229–235.

49. Walsh, F., and Duffy, B. (2013). The Culturable Soil Antibiotic Resistome: A Community of Multi-Drug Resistant Bacteria. *PLoS ONE* 8, e65567.
50. D'Costa, V. M., McGrann, K. M., Hughes, D. W., and Wright, G. D. (2006). Sampling the antibiotic resistome. *Science* 311, 374–377.
51. Cundliffe, E., and Demain, A. L. (2010). Avoidance of suicide in antibiotic-producing microbes. *J Ind Microbiol Biotechnol* 37, 643–672.
52. Morgan, J. L., Darling, A. E., and Eisen, J. A. (2010). PLOS ONE: Metagenomic Sequencing of an In Vitro-Simulated Microbial Community. *PLoS ONE*.
53. Delmont, T. O., Robe, P., Cécillon, S., Clark, I. M., Constancias, F., Simonet, P., Hirsch, P. R., and Vogel, T. M. (2011). Accessing the soil metagenome for studies of microbial diversity. *Appl Environ Microbiol* 77, 1315–1324.
54. Delmont, T. O., Robe, P., Clark, I., Simonet, P., and Vogel, T. M. (2011). Metagenomic comparison of direct and indirect soil DNA extraction approaches. *J Microbiol Methods*.
55. Zhang, J., Kapli, P., Pavlidis, P., and Stamatakis, A. (2013). A general species delimitation method with applications to phylogenetic placements. *Bioinformatics*.
56. Sunagawa, S., Mende, D. R., Zeller, G., Izquierdo-Carrasco, F., Berger, S. A., Kultima, J. R., Coelho, L. P., Arumugam, M., Tap, J., Nielsen, H. B., et al. (2013). Metagenomic species profiling using universal phylogenetic marker genes. *Nat Methods* 10, 1196–1199.
57. Stackebrandt, E., Frederiksen, W., Garrity, G. M., Grimont, P. A. D., Kämpfer, P., Maiden, M. C. J., Nesme, X., Rosselló-Mora, R., Swings, J., Trüper, H. G., et al. (2002). Report of the ad hoc committee for the re-evaluation of the species definition in bacteriology. *Int J Syst Evol Microbiol* 52, 1043–1047.
58. Rodríguez, S. J., and Bishop, P. L. (2007). Three-dimensional quantification of soil biofilms using image analysis. *Environmental Eng.*
59. Fortuna, A.-M. (2012). The Soil Biota. *Nature Education Knowledge* 3. Available at: <http://www.nature.com/scitable/knowledge/library/the-soil-biota-84078125> [Accessed September 13, 2013].
60. Prosser, J. I., Bohannon, B. J. M., Curtis, T. P., Ellis, R. J., Firestone, M. K., Freckleton, R. P., Green, J. L., Green, L. E., Killham, K., Lennon, J. J., et al. (2007). The role of ecological theory in microbial ecology. *Nat Rev Microbiol* 5, 384–392.
61. Ranjard, L., Poly, F., Combrisson, J., Richaume, A., Gourbiere, F., Thioulouse, J., and Nazaret, S. (2000). Heterogeneous cell density and genetic structure of bacterial pools associated with various soil microenvironments as determined by enumeration and DNA fingerprinting approach (RISA). *Microb Ecol* 39, 263–272.
62. Youssef, N. H., Couger, M. B., and Elshahed, M. S. (2010). Fine-scale bacterial beta

- diversity within a complex ecosystem (Zodletone Spring, OK, USA): the role of the rare biosphere. *PLoS ONE* 5, e12414.
63. Ranjard, L., and Richaume, A. (2001). Quantitative and qualitative microscale distribution of bacteria in soil. *Res Microbiol* 152, 707–716.
  64. Nunan, N., Wu, K., Young, I. M., Crawford, J. W., and Ritz, K. (2003). Spatial distribution of bacterial communities and their relationships with the micro-architecture of soil. *FEMS Microbiol Ecol* 44, 203–215.
  65. Peay, K. G., Bruns, T. D., Kennedy, P. G., Bergemann, S. E., and Garbelotto, M. (2007). A strong species?area relationship for eukaryotic soil microbes: island size matters for ectomycorrhizal fungi. *Ecol Letters* 10, 470–480.
  66. Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K. S., Manichanh, C., Nielsen, T., Pons, N., Levenez, F., Yamada, T., et al. (2010). A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* 464, 59–65.
  67. Benveniste, R., and Davies, J. (1973). Aminoglycoside antibiotic-inactivating enzymes in actinomycetes similar to those present in clinical isolates of antibiotic-resistant bacteria. *Proc Natl Acad Sci USA* 70, 2276–2280.
  68. Garau, G., Di Guilmi, A. M., and Hall, B. G. (2005). Structure-Based Phylogeny of the Metallo-β-Lactamases. *Antimicrob Agents Chemother* 49, 2778–2784.
  69. Zhu, Y.-G., Johnson, T. A., Su, J.-Q., Qiao, M., Guo, G.-X., Stedtfeld, R. D., Hashsham, S. A., and Tiedje, J. M. (2013). Diverse and abundant antibiotic resistance genes in Chinese swine farms. *Proc Natl Acad Sci USA*.
  70. Davies, J., and Ryan, K. S. (2012). Introducing the parvome: bioactive compounds in the microbial world. *ACS Chem. Biol.* 7, 252–259.
  71. Bernier, S. P., and Surette, M. G. (2013). Concentration-dependent activity of antibiotics in natural environments. *Front Microbiol* 4, 20.
  72. Davies, J. (2011). How to discover new antibiotics: harvesting the parvome. *Curr Opin Chem Biol* 15, 5–10.
  73. Cooper, M. A., and Shlaes, D. (2011). Fix the antibiotics pipeline. *Nature* 472, 32.
  74. Marshall, K. C., and Alexander, M. (1960). Competition between soil bacteria and fusarium. *Plant Soil* 12, 143–153.
  75. Marshall, C. G., Lessard, I. A., Park, I., and Wright, G. D. (1998). Glycopeptide antibiotic resistance genes in glycopeptide-producing organisms. *Antimicrob Agents Chemother* 42, 2215–2220.
  76. Liu, Z., Wang, W., Zhu, Y., Gong, Q., Yu, W., and Lu, X. (2013). Antibiotics at subinhibitory concentrations improve the quorum sensing behavior of *Chromobacterium violaceum*. *FEMS Microbiol Lett* 341, 37–44.

77. Silva, A. J., Benitez, J. A., and Wu, J.-H. (2010). Attenuation of bacterial virulence by quorum sensing-regulated lysis. *J Biotechnol* *150*, 22–30.
78. Meroueh, S. O., Minasov, G., Lee, W., Shoichet, B. K., and Mobashery, S. (2003). Structural aspects for evolution of beta-lactamases from penicillin-binding proteins. *J. Am. Chem. Soc.* *125*, 9612–9618.
79. Hall, A. R. (2013). Genotype-by-environment interactions due to antibiotic resistance and adaptation in *Escherichia coli*. *J Evol Biol* *26*, 1655–1664.
80. Laskaris, P., Tolba, S., Calvo-Bado, L., Wellington, E. M., and Wellington, L. (2010). Coevolution of antibiotic production and counter-resistance in soil bacteria. *Environ Microbiol* *12*, 783–796.
81. Haas, D., and Défago, G. (2005). Biological control of soil-borne pathogens by fluorescent pseudomonads. *Nat Rev Microbiol* *3*, 307–319.
82. Mummey, D., Holben, W., Six, J., and Stahl, P. (2006). Spatial stratification of soil bacterial populations in aggregates of diverse soils. *Microb Ecol* *51*, 404–411.
83. Martin-Laurent, F., Philippot, L., Hallet, S., Chaussod, R., Germon, J. C., Soulas, G., and Catroux, G. (2001). DNA extraction from soils: old bias for new microbial diversity analysis methods. *Appl Environ Microbiol* *67*, 2354–2359.
84. Grundmann, G. L. (2004). Spatial scales of soil bacterial diversity – the size of a clone - Grundmann - 2006 - FEMS Microbiology Ecology - Wiley Online Library. *FEMS Microbiol Ecol*.
85. Dechesne, A., Pallud, C., Debouzie, D., Flandrois, J. P., Vogel, T. M., Gaudet, J. P., and Grundmann, G. L. (2003). A novel method for characterizing the microscale 3D spatial distribution of bacteria in soil. *Soil Biology and Biochemistry* *35*, 1537–1546.
86. Thiele Bruhn, S. (2003). Pharmaceutical antibiotic compounds in soils—a review. *Journal of Plant Nutrition and Soil Science* *166*, 145–167.
87. Kümmerer, K. (2009). Antibiotics in the aquatic environment—a review—part I. *Chemosphere* *75*, 417–434.
88. Kümmerer, K. (2010). Pharmaceuticals in the Environment. *Annu. Rev. Environ. Resourc.* *35*, 57–75.
89. Rabølle, M., and Spliid, N. H. (2000). Sorption and mobility of metronidazole, olaquinox, oxytetracycline and tylosin in soil. *Chemosphere*.
90. Leal, R. M. P., Alleoni, L. R. F., Tornisiello, V. L., and Regitano, J. B. (2013). Sorption of fluoroquinolones and sulfonamides in 13 Brazilian soils. *Chemosphere* *92*, 979–985.
91. Tolls, J. (2001). Sorption of Veterinary Pharmaceuticals in Soils: A Review - Environmental Science & Technology (ACS Publications). *Environ Sci Technol*.

92. Maszkowska, J., Kołodziejska, M., Białk-Bielińska, A., Mrozik, W., Kumirska, J., Stepnowski, P., Palavinskas, R., Krüger, O., and Kalbe, U. (2013). Column and batch tests of sulfonamide leaching from different types of soil. *Journal of Hazardous Materials* 260C, 468–474.
93. Jiang, M., Wang, L., and Ji, R. (2010). Biotic and abiotic degradation of four cephalosporin antibiotics in a lake surface water and sediment. *Chemosphere* 80, 1399–1405.
94. Braschi, I., Blasioli, S., Fellet, C., Lorenzini, R., Garelli, A., Pori, M., and Giacomini, D. (2013). Persistence and degradation of new  $\beta$ -lactam antibiotics in the soil and water environment. *Chemosphere* 93, 152–159.
95. Girardi, C., Greve, J., Lamshöft, M., Fetzer, I., Miltner, A., Schäffer, A., and Kästner, M. (2011). Biodegradation of ciprofloxacin in water and soil and its effects on the microbial communities. *Journal of Hazardous Materials* 198, 22–30.
96. Chee-Sanford, J. C., Mackie, R. I., Koike, S., Krapac, I. G., Lin, Y.-F., Yannarell, A. C., Maxwell, S., and Aminov, R. I. (2009). Fate and transport of antibiotic residues and antibiotic resistance genes following land application of manure waste. *J Environ Qual* 38, 1086–1108.
97. Sun, P., Barmaz, D., Cabrera, M. L., Pavlostathis, S. G., and Huang, C.-H. (2013). Detection and quantification of ionophore antibiotics in runoff, soil and poultry litter. *J Chromatogr A*.
98. Moreno-Bondi, M. C., Marazuela, M. D., Herranz, S., and Rodriguez, E. (2009). An overview of sample preparation procedures for LC-MS multiclass antibiotic determination in environmental and food samples. *Anal Bioanal Chem* 395, 921–946.
99. Willingham, T. W., and Werth, C. J. (2008). Evaluation of the Effects of Porous Media Structure on Mixing-Controlled Reactions Using Pore-Scale Modeling and Micromodel Experiments - Environmental Science & Technology (ACS Publications). *Environmental science & ....*
100. Zhang, C., Kang, Q., Wang, X., Zilles, J. L., Müller, R. H., and Werth, C. J. (2010). Effects of pore-scale heterogeneity and transverse mixing on bacterial growth in porous media. *Environ Sci Technol* 44, 3085–3092.
101. Wang, X., Kang, Y., Luo, C., Zhao, T., Liu, L., Jiang, X., Fu, R., An, S., Chen, J., Jiang, N., et al. (2014). Heteroresistance at the single-cell level: adapting to antibiotic stress through a population-based strategy and growth-controlled interphenotypic coordination. *MBio* 5, e00942–13.
102. Li, B., Qiu, Y., Glidle, A., McIlvenna, D., Luo, Q., Cooper, J., Shi, H.-C., and Yin, H. (2014). Gradient microfluidics enables rapid bacterial growth inhibition testing. *Anal Chem* 86, 3131–3137.

103. Bahl, M. I., Hansen, L. H., Licht, T. R., and Sørensen, S. J. (2004). In Vivo Detection and Quantification of Tetracycline by Use of a Whole-Cell Biosensor in the Rat Intestine. *Antimicrob Agents Chemother* *48*, 1112–1117.
104. Hansen, L. H., Ferrari, B., Sørensen, A. H., Veal, D., and Sørensen, S. J. (2001). Detection of oxytetracycline production by *Streptomyces rimosus* in soil microcosms by combining whole-cell biosensors and flow cytometry. *Appl Environ Microbiol* *67*, 239–244.
105. Slager, J., Kjos, M., Attaiech, L., and Veening, J.-W. (2014). Antibiotic-Induced Replication Stress Triggers Bacterial Competence by Increasing Gene Dosage near the Origin. *Cell* *157*, 395–406.
106. Fajardo, A., and Martínez, J. L. (2008). Antibiotics as signals that trigger specific bacterial responses. *Curr Opin Microbiol* *11*, 161–167.
107. Chuanchuen, R., Karkhoff-Schweizer, R. R., and Schweizer, H. P. (2003). High-level triclosan resistance in *Pseudomonas aeruginosa* is solely a result of efflux. *American Journal of Infection Control* *31*, 124–127.
108. Martinez, J. L., Sánchez, M. B., Martínez-Solano, L., Hernandez, A., Garmendia, L., Fajardo, A., and Alvarez-Ortega, C. (2009). Functional role of bacterial multidrug efflux pumps in microbial natural ecosystems. *FEMS Microbiol Rev* *33*, 430–449.
109. Zapun, A., Contreras-Martel, C., and Vernet, T. (2008). Penicillin-binding proteins and  $\beta$ -lactam resistance. *FEMS Microbiol Rev* *32*, 361–385.
110. Johnson, J. W., Fisher, J. F., and Mobashery, S. (2013). Bacterial cell-wall recycling. *Ann N Y Acad Sci* *1277*, 54–75.
111. Deghorain, M., Fontaine, L., David, B., Mainardi, J.-L., Courtin, P., Daniel, R., Errington, J., Sorokin, A., Bolotin, A., Chapot-Chartier, M.-P., et al. (2010). Functional and morphological adaptation to peptidoglycan precursor alteration in *Lactococcus lactis*. *J Biol Chem* *285*, 24003–24013.
112. Yuan, M., Yu, Y., Li, H. R., Dong, N., and Zhang, X. H. (2014). Phylogenetic Diversity and Biological Activity of Actinobacteria Isolated from the Chukchi Shelf Marine Sediments in the Arctic Ocean. *Marine Drugs*.
113. Fenical, W., and Jensen, P. R. (2006). Developing a new resource for drug discovery: marine actinomycete bacteria. *Nat. Chem. Biol.* *2*, 666–673.
114. Shen, B. (2003). Polyketide biosynthesis beyond the type I, II and III polyketide synthase paradigms. *Curr Opin Chem Biol*.
115. Hutchinson, C. R., and Fujii, I. (1995). Polyketide synthase gene manipulation: a structure-function approach in engineering novel antibiotics. *Annu Rev Microbiol* *49*, 201–238.
116. Ginolhac, A., Jarrin, C., Gillet, B., Robe, P., Pujic, P., Tuphile, K., Bertrand, H., Vogel,

- T. M., Perrière, G., Simonet, P., et al. (2004). Phylogenetic analysis of polyketide synthase I domains from soil metagenomic libraries allows selection of promising clones. *Appl Environ Microbiol* *70*, 5522–5527.
117. Boucher, H. W., Talbot, G. H., Bradley, J. S., Edwards, J. E., Gilbert, D., Rice, L. B., Scheld, M., Spellberg, B., and Bartlett, J. (2009). Bad Bugs, No Drugs: No ESCAPE! An Update from the Infectious Diseases Society of America. *Clin Inf Dis*.
  118. Nordmann, P., Naas, T., and Poirel, L. (2011). Global spread of Carbapenemase-producing Enterobacteriaceae. *Emerging Infect Dis* *17*, 1791–1798.
  119. Yong, D., Toleman, M. A., Giske, C. G., Cho, H. S., Sundman, K., Lee, K., and Walsh, T. R. (2009). Characterization of a New Metallo-β-Lactamase Gene, blaNDM-1, and a Novel Erythromycin Esterase Gene Carried on a Unique Genetic Structure in *Klebsiella pneumoniae* Sequence Type 14 from India. *Antimicrob Agents Chemother* *53*, 5046–5054.
  120. Walsh, T. R., Weeks, J., Livermore, D. M., and Toleman, M. A. (2011). Dissemination of NDM-1 positive bacteria in the New Delhi environment and its implications for human health: an environmental point prevalence study. *Lancet Infect Dis* *11*, 355–362.
  121. Ho, P. L., Li, Z., Lai, E. L., Chiu, S. S., and Cheng, V. C. C. (2012). Emergence of NDM-1-producing Enterobacteriaceae in China. *J Antimicrob Chemother* *67*, 1553–1555.
  122. Dortet, L., Nordmann, P., and Poirel, L. (2012). Association of the emerging carbapenemase NDM-1 with a bleomycin resistance protein in Enterobacteriaceae and *Acinetobacter baumannii*. *Antimicrob Agents Chemother* *56*, 1693–1697.
  123. Ho, P. L., Lo, W. U., Yeung, M. K., Lin, C. H., Chow, K. H., Ang, I., Tong, A. H. Y., Bao, J. Y.-J., Lok, S., and Lo, J. Y. C. (2011). Complete Sequencing of pNDM-HK Encoding NDM-1 Carbapenemase from a Multidrug-Resistant *Escherichia coli* Strain Isolated in Hong Kong. *PLoS ONE* *6*.
  124. Toleman, M. A., Spencer, J., Jones, L., and Walsh, T. R. (2012). blaNDM-1 is a chimera, likely constructed in *Acinetobacter baumannii*. *Antimicrob Agents Chemother*.
  125. Bogaerts, P., Huang, T.-D., Rezende de Castro, R., Bouchahrouf, W., and Glupczynski, Y. (2013). Could *Acinetobacter pittii* act as an NDM-1 reservoir for Enterobacteriaceae? *J Antimicrob Chemother* *68*, 2414–2415.
  126. Villa, L., Poirel, L., Nordmann, P., Carta, C., and Carattoli, A. (2012). Complete sequencing of an IncH plasmid carrying the bla(NDM-1), bla(CTX-M-15) and qnrB1 genes. *J Antimicrob Chemother* *67*, 1645–1650.
  127. Carattoli, A., Villa, L., Poirel, L., Bonnin, R. A., and Nordmann, P. (2012). Evolution of IncA/C bla(CMY-2)-Carrying Plasmids by Acquisition of the bla(NDM-1) Carbapenemase Gene. *Antimicrob Agents Chemother* *56*, 783–786.

128. Heuer, H., and Smalla, K. (2012). Plasmids foster diversification and adaptation of bacterial populations in soil. *FEMS Microbiol Rev* *36*, 1083–1104.
129. Roberts, A. P., and Mullany, P. (2009). A modular master on the move: the Tn916 family of mobile genetic elements. *Trends Microbiol* *17*, 251–258.
130. Liebert, C., Hall, R., and Summers, A. (1999). Transposon Tn21, flagship of the floating genome. *Microbiology and Molecular Biology Reviews* *63*, 507–.
131. Mazel, D. (2006). Integrons: agents of bacterial evolution. *Nat Rev Microbiol* *4*, 608–620.
132. Gillings, M. R., Krishnan, S., Worden, P. J., and Hardwick, S. A. (2008). Recovery of diverse genes for class 1 integron-integrases from environmental DNA samples. *FEMS Microbiol Lett* *287*, 56–62.
133. Jolivet-Gougeon, A., Kovacs, B., Le Gall-David, S., Le Bars, H., Bousarghin, L., Bonnaure-Mallet, M., Lobel, B., Guillé, F., Soussy, C.-J., and Tenke, P. (2011). Bacterial hypermutation: clinical implications. *J Med Microbiol* *60*, 563–573.
134. Tenaillon, O., Taddei, F., Radman, M., and Matic, I. (2001). Second-order selection in bacterial evolution: selection acting on mutation and recombination rates in the course of adaptation. *Res Microbiol* *152*, 11–16.
135. Harbarth, S., and Samore, M. H. (2005). Antimicrobial Resistance Determinants and Future Control.: EBSCOhost. *Emerging Infect Dis.*
136. Baker-Austin, C., Wright, M. S., Stepanauskas, R., and McArthur, J. V. (2006). Co-selection of antibiotic and metal resistance. *Trends Microbiol* *14*, 176–182.

# Chapitre 2.

## ***Étude métagénomique à grande échelle de la résistance antibiotique dans l'environnement.***

### **Résumé.**

L'acquisition et la dissémination des gènes de résistance à des antibiotiques par les bactéries pathogènes est un problème critique de santé publique menaçant sérieusement notre capacité à maîtriser les maladies infectieuses [1-3]. L'accumulation très rapide de phénotypes résistants observés en milieu hospitalier suggère un immense réservoir de gènes de résistance (ARGD) transférables aux bactéries pathogènes et sélectionnés en réponse aux traitements antibiotiques [1, 3-5]. Les ARGD ont déjà été retrouvés dans de nombreux échantillons environnementaux comme le sol [4,5], sur les parois rocheuses de grottes isolées pendant des milliers d'années [6] ou dans le permafrost [7,8]. Les molécules antibiotiques ciblent des fonctions cellulaires essentielles, présentes chez toutes les bactéries [8-11]. L'environnement est un très abondant réservoir de gènes de résistance transférable aux bactéries pathogènes [9-12]. En utilisant une approche métagénomique, il a été observé que des gènes de résistance à des antibiotiques peuvent être détectés dans tous les environnements testés dans cette étude (n=71). Les pompes à efflux représentent le type de résistance le plus communément retrouvé dans l'environnement suivi de gènes impliqués dans la résistance à la vancomycine, aux tétracyclines et aux bêta-lactames. Cette étude permet ainsi de décrire l'abondance et la diversité des gènes de résistance présents dans l'environnement et montre que leur distribution est largement influencée par l'environnement.

## Chapter 2.

# ***Large-scale metagenomic-based study of antibiotic resistance in the environment.***

### **Abstract.**

Antibiotic resistance, including multi-resistance acquisition and dissemination by pathogens, is a critical healthcare issue threatening our management of infectious diseases [1-3]. Rapid accumulation of resistance phenotypes implies a reservoir of transferable antibiotic resistance gene determinants (ARGDs) selected in response to inhibiting antibiotic concentrations as found in hospitals [1-5]. Antibiotic resistance genes were found in environmental isolates, soil DNA [4, 5], secluded caves [6] and permafrost DNA [7]. Antibiotics target essential and ubiquitous cell functions and resistance is a common characteristic of environmental bacteria [8-11]. Environmental ARGDs are an abundant reservoir of potentially transferable resistance for pathogens [9-12]. Using metagenomic sequences, we show that ARGDs can be detected in all (n=71) environments analyzed. Soil metagenomes had the most diverse pool of ARGDs. The most common types of resistances found in environmental metagenomes were efflux pumps and genes conferring resistance to vancomycin, tetracycline or beta-lactam antibiotics used in veterinary and human healthcare. Our study describes the diverse and abundant antibiotic resistance genes in non-clinical environments and shows that these genes are not randomly distributed among different environments (e.g., soil, oceans or human feces).

## Introduction.

The most successful improvement in healthcare during the 20th century was the development of antibiotic therapy that significantly reduced morbidity and mortality of infectious diseases. However, this major success was almost immediately followed by the discovery of pathogenic bacterial strains resisting antibiotic therapies. Horizontal gene transfer (HGT) was suspected as the reason for the quick spread of antibiotic resistance genes among bacterial pathogens in the medical environment [1, 3]. Tracked by the increasing number of reports on the extent of resistance dissemination year after year, the speed at which ARGD were spread among pathogens was considered to be one of the clearest examples of the potential of bacteria to adapt via horizontal gene transfer [8, 13, 14]. However, this fundamental adaptive question with critical sanitary applications led to questions regarding the origin of ARGD in pathogenic bacteria, the role of environmental bacteria as the reservoir of these genes, and the transfer mechanisms between saprophytic environmental bacteria and clinical pathogen bacteria [15-17]. The critical decrease of antibiotic efficiency for treating infectious diseases would be partly due to the horizontal transfer of antibiotic resistance gene determinants from environmental bacteria to pathogenic germs. Environmental microorganisms are considered as a reservoir of adaptive genes for clinical bacteria [18, 19]. Exploration of this reservoir in terms of diversity of resistance mechanisms, their prevalence in bacterial communities from various environments, their potential to be horizontally transferred is critical, on one hand, to decipher their ecological role in the environment and in another hand, to extend lifetime of existing and future antibiotics towards resistance by limiting the dissemination of ARGD and selection of resistant populations in clinical environments [2]. In terms of public health, the

environmental ARGD reservoir can lead to ARGD dissemination in the clinic. Relatively few studies, however, have been conducted on the environmental resistome. This is first due to inherent difficulties to culture 99% of total living bacteria in a given soil sample [20]. This “*unseen majority*” [21] is a large reservoir of genetic diversity, order of magnitudes higher than previously expected by cultivation or direct observation of samples. Use of culture independent approaches and high-throughput sequencing might increase significantly our understanding of environmental resistance in terms of ecology and evolution in order to manage antibiotic use to limit resistance spread [22]. To survey abundance and diversity of antibiotic resistance prevalence and diversity in the environment, at a global scale, a global metagenomic approach was employed. Massive amount of DNA sequence data obtained with next-generation sequencing technologies of total DNA from environmental samples are compared with an up-to-date curated antibiotic resistance sequence database. Combining both the sensitivity of high-throughput sequencing with a dedicated database and relevant stringency as well as a relative sensitivity, due to the high number of sequence reads analyzed (i.e. 8.3Gbp  $\approx$  1500 *Escherichia coli* str. K12 genomes), we characterized the prevalence and diversity of ARGD in a large range of environments (n=71). Results show a global prevalence of ARGD in the environment. This is strengthened by results using either functional-screening of metagenomic DNA or isolation of resistant bacteria accessible in the literature.

## **Experimental procedures.**

### **Metagenomes selection.**

Environmental shotgun sequencing (EES, [12]) metagenomic datasets used in this study were downloaded from two public repositories: the MG-RAST server

(<http://metagenomics.anl.gov/>, [23]) and the MetaSoil project website (<http://www.genomenviron.org/Projects/METASOIL.html>) [24]). The environmental metagenomes selected consisted of: soil (Rothamsted Park Grass 3d plot, Puerto-Rico forest soil and Waseca County farm soil), human feces (Twin Study, 29 Healthy Human Gut project), chicken gut (Chicken Cecum Microbiome), cow rumen, chicken gut, mouse gut (Obese & Lean Mouse Gut), oceans (Global Ocean Survey, Bacterioplankton, Whale Fall), activated sludge from Phosphorous removing wastewater treatment plant, Singapore Indoor Air Filter, microbial fuel cell, hyper-saline sediments, Arctic snow and Antarctic lakes. The selection of the metagenomes was based on the availability of at least two metagenomes from the same type of ecosystem. Accession numbers on the MG-RASTv3 (<http://metagenomics.anl.gov/>) platform and references for each dataset are available as supplementary information (Table S1). Rothamsted Park Grass Experiment untreated plot (3d) sequence datasets can be downloaded on the MetaSoil website (<http://www.genomenviron.org/Projects/METASOIL.html>). Sequence reads submitted to the two repositories were downloaded locally and artificial duplicates, known to occur especially with the Roche 454 FLX technology, were removed for each dataset using cd-hit-454 program [25]. All metagenomic datasets were obtained either with Sanger or Roche FLX 454 sequencing technology, by different research groups, laboratories and sequencing centers and has been made publicly available on MG-RAST (Table S1). All datasets after being curated were compiled into one single multi-FASTA file representing the query for the BLAST analysis. The query file totalized about 8.3 Gbp of sequences, which corresponds to roughly 1500 *Escherichia coli* genome equivalent.

#### **ARGD annotation**

The Antibiotic Resistance DataBase (ARDB) is a specific database containing only proteins known to confer antibiotic resistance to their bacterial host developed by Liu et al. [26]. The ARDB was retrieved from <http://ardb.cbcb.umd.edu/> to build a non-redundant protein database. In order to identify sequence reads a homology search between metagenomic sequence reads and antibiotic resistance proteins was performed using BLASTx [27]. Only the lowest e-value hit per read hit-list was kept as a “best-hit per read” annotation strategy. As described in Looft et al. [28], a stringent cut-off suited to short-reads dataset analysis (bit-score  $\geq 60$  and percent-identity  $\geq 35\%$ ) has been chosen to filter the BLAST output file and generate a hit-table per metagenome. To normalize the differences in average sequence read length and read number per metagenome in order to perform inter-environmental comparisons, the abundance counts for each accession number in the database has been divided by the total number of annotated reads in the dataset. The global hit-table is available as supplementary material (Excel Table S4). In order to check if annotation based on ARDB is not giving significant amount of false-positives hits ARDB annotations are processed with a second BLAST step. Every read with a match in ARDB (i.e.: ARGD-annotated reads) are extracted from our initial query file. Matches are then searched in a generalist database (NCBI\_nr, November 2013 release). Results are available in Table S5. NCBI accession number annotation is providing minimal function information and need to be supplemented by a curated annotation subsystem. SEED annotation system is used to provide comprehensive annotations of our reads. SEED convert NCBI accession number to “subsystems”, providing a hierarchical functional annotation system (Table S5) [29]. Most antibiotic resistance related functions in SEED are found under the Virulence subsystem and subsequently classified into more specific functions such as: multi-drug efflux pumps,  $\beta$ -lactamases, etc.

## **Data analyses.**

Dataset properties such as read number, base pair volume and average read lengths influence greatly the abundance count of such analysis method. Therefore, to compare the dissimilarity of distributions between metagenomes, only datasets obtained from the same sequencing technology (Roche 454 pyrosequencing) were used. Sanger datasets are less homogeneous and the number of observations is too low for some environments to perform such analysis (data not shown). This subset represents three distinct environments: human feces, soil and ocean from various projects. All statistical computation was performed with R-software (<http://www.r-project.org/>) using various third-party packages. Distance matrix between each ARGD distribution has been calculated using Mountford index using Vegan R-software package on the raw ARGD hits table (Table S2) [30]. Hierarchical-clustering of the distance matrix obtained has been performed with the *hclust* function from the R software base-package using the Ward method. All calculations are made with the raw hits (*i.e.* not normalized by annotated reads in each metagenomic dataset) from all accession numbers of our database (Table S2).

## **Results.**

A global environmental survey was performed with seventy-one environmental shotgun metagenomic DNA datasets obtained from public repositories websites. These datasets were derived from sequencing DNA extracted from various environmental samples (Environmental Shotgun Sequencing, EES [12]). Two different sequencing technologies were used to generate these datasets (Roche 454 and Sanger sequencing). These datasets equal a total of 8Gbp (equivalent to about 1550 *Escherichia coli* genomes). Details concerning these datasets are available in supplemental information Table S1. The environmental

metagenomic reads from these datasets were compared to a antibiotic resistance database (ARDB) containing 2999 different amino-acid sequences known to confer antibiotic resistance phenotypes to their host (<http://ardb.cbcb.umd.edu>) [26]. Every metagenomic dataset contained sequences with potential antibiotic resistance based on their similarity to this antibiotic resistance database (ARDB) at the cut-off threshold used (i.e. BLASTx bit-score > 60 and alignment identity > 35%; see experimental procedures). Sequences matching the ARDB were then re-annotated with a generalist database (NCBI nr, release Nov. 2013, SEED), which confirmed specificity of the antibiotic resistance annotation when using SEED annotation (Figure S2 and S3). The SEED annotation system was designed to annotate 2133 different functions at the subsystem level [29], but lacks the resolution for more specific functions, such as antibiotic resistance. Therefore, 94% (n=121583) of the reads that were annotated as resistance genes based on the ARDB were placed in the “Not Assigned” category in SEED (Figure S2). The same general trends were observed with both annotation procedures, but results obtained with ARDB annotations were more consistent regarding antibiotic-resistance specific annotations as expected for a dedicated database. This was essentially due to a function-specific, curated, database. Results obtained with SEED using NCBI\_nr BLAST results are available as supplemental data (Figure S2, S3 and Table S1) and the ARDB annotations are described below.

### **Total antibiotic resistance gene determinants (ARGDs) abundance and diversity in each environment.**

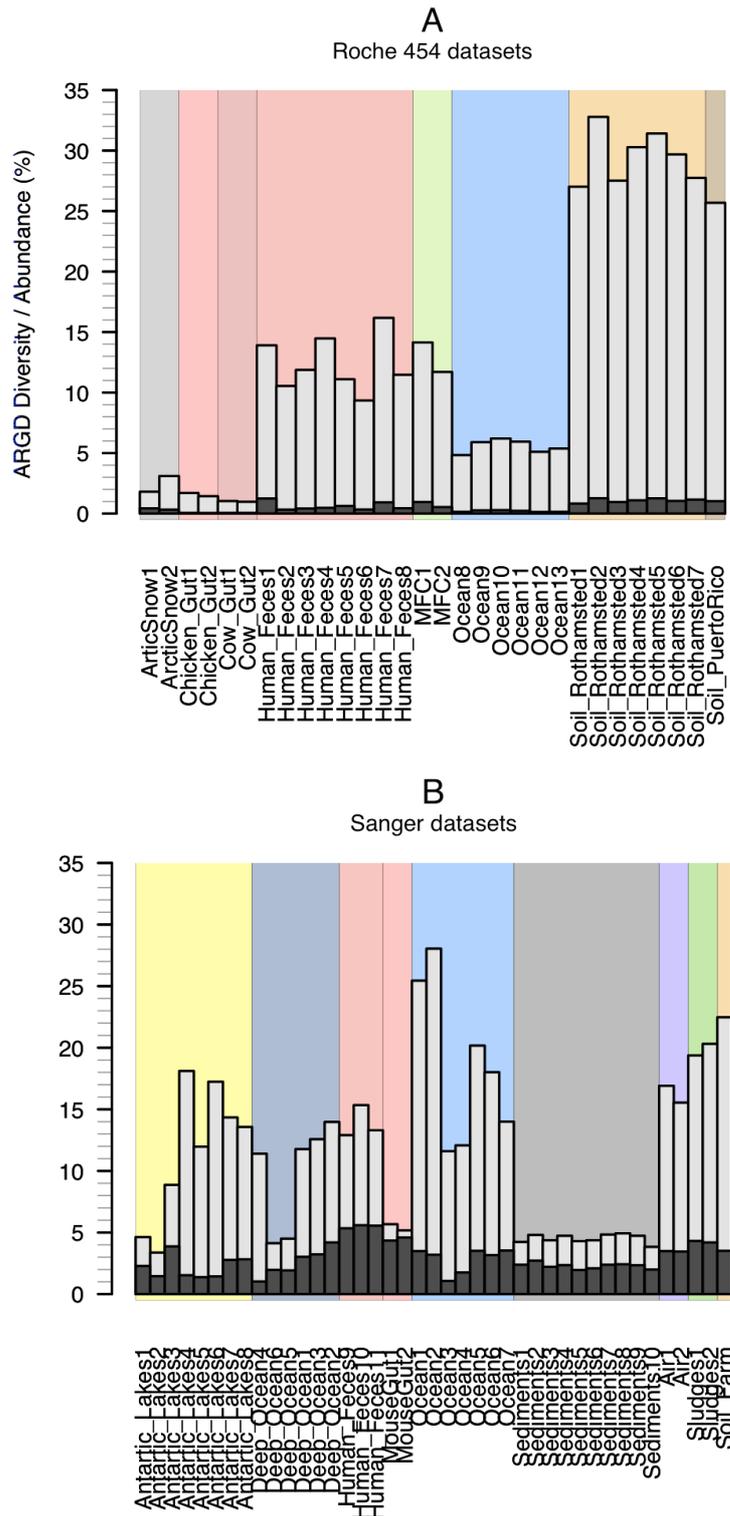
Relative antibiotic resistance gene determinant (ARGD) abundance ranged from 0.05% in chicken gut (metagenome accession number 4440284.3) to 5.6% in human feces (4440942.3) and differed significantly (at 95% CI) among the different environmental datasets (Figure 1, see also Table S1). Datasets from environments with the highest

proportion of ARGD-annotated sequence reads were the three human feces datasets from Japan 29 Healthy Human Gut project (5.6% to 5.3% of all reads; Sanger sequencing), mouse gut (4.6% and 4.3% of reads; Sanger sequencing) and activated sludge (4.3% to 4.2% of reads; Sanger sequencing) (Figure 1). Human feces metagenomes obtained with Roche 454 pyrosequencing technology had less than 1% ARGD relative abundance. Sequencing technology can significantly influence annotation and need to be taken into consideration when looking at metagenomic data. The longer reads obtained by Sanger were easier to annotate, thereby, increasing their relative annotated read frequency, but such datasets have lower read counts compared to high-throughput next generation sequencing technologies, which are harder to annotate due to their relatively shorter reads such as those obtained from Roche 454 [31]. Still, Sanger-sequenced human feces datasets contained relatively more ARGD-annotated reads than environmental datasets from oceans or farm soil also obtained with Sanger. All three Sanger-sequenced datasets had higher relative numbers of ARGD-annotated reads than cow gut, chicken gut and six ocean datasets, which were obtained by Roche 454. These Roche 454-sequenced datasets are also those with the shortest mean read length (about 100bp) among the studied metagenomes. The ARGD relative diversity found in each environment was estimated as a percentage of the 2999 different sequences from the ARDB detected in the corresponding sequence datasets (Figure 1). Overall, the percentage of different ARGD sequences detected ranged from 0.97% in cow gut (4441680.3) to 32.78% in Rothamsted soil E41. Our results showed that soil had the highest level of ARGD diversity with an average of  $28.8\% \pm 3.14\%$  of the database sequences found in the soil datasets (max. 32.8% in Rothamsted E41 and min. 22.5% in Waseca Farm Soil) (Figure 1). Datasets from cow rumen, chicken cecum and arctic snow had the lowest number of different ARGD sequences with only 1.3% of the ARDB

database sequences detected in the cow rumen, chicken cecum and arctic snow metagenomic datasets. Metagenomic datasets from human feces from different origins and sequencing technologies had an average of  $12.7\% \pm 2.1\%$  of the different ARDB database sequences (max. 16.17% in Twin Study 4440616.3 and min. 9.3% in Twin Study 4440614.3). This observation is consistent with the general trend of a reduced diversity level of antibiotic resistance genes in non-soil environments compared to soil (Figure 1).

### **Most prevalent resistance types in the environment.**

The macab and bcr efflux pumps accounted for 18.9% and 8.5% of the total ARGD hits in all environments, respectively (Table 1, details available as supplementary Table S2 and Figure S1). Most ARGD-annotated reads detected in metagenomic datasets using our stringent annotation cut-off were related to various types of efflux-pump components: membrane fusion protein from resistance-nodulation cell division efflux pumps (RND) and ATP-binding cassette (ABC) domains related to bacitracin efflux with 27.9% and 20.8% of the total hits of ARGD annotated sequences in all metagenomes, respectively (Table 1, see also supplemental Table S2). Tetracycline ribosomal protection protein encoding genes (tet\_rpp ; 5.7%) and vancomycin resistance related genes (van a : 3.6% ; van b : 2.9% ; van d : 3.5% ; van e 2.3% ; van g 2.8% ; van c 1.8%) were also found. Penicillin binding protein (Pbp) accounted for 5.4% of annotated reads.  $\beta$ -lactamases were less frequently detected (class A  $\beta$ -lactamases, bla\_a : 3.4%; class B  $\beta$ -lactamases, bla\_b: 0.2% ; class C  $\beta$ -lactamases, bla\_c: 0.42% ; class D  $\beta$ -lactamases, bla\_d: 0.04%) (Table S2).



**Figure 1 – Relative diversity and abundance of ARGD between metagenomes.**

Barplot showing the relative abundance of reads annotated as ARGD (gray bars), expressed as number of ARGD hits divided by total number of reads in the dataset and ARGD diversity (white bars), expressed as the percentage of the 2999 unique sequences in ARDB database found in each dataset. Shaded boxes indicate a different environmental origin with from left to right: Antarctic lakes; Arctic snow; chicken gut; cow gut; deep oceans; human feces; microbial fuel-cell; mouse gut; ocean; halophile sediment; activated sludge and soil. Black dots indicate datasets obtained by Sanger sequencing method, others are obtained by Roche 454 sequencing.

<b>Resistance Class</b>	<b>Mechanism of resistance</b>	<b>Antibiotic specificity</b>
mexef, ceo, mexvw, acr, mexhi, mexcd, mexab, mdtnop, amr, adeabc, smeabc, smedef, mdtef, mexxy, mdtk	RND class transporter	Multi-drug resistance efflux
macab	RND class transporter : Macrolide	
bcr, bcr_mfs	ABC class transporter system : Bacitracin	
mls_abc	ABC class transporter : Macrolide	
mls_mfs, mls_hdr	MFS class transporter : Macrolide	
cml	MFS class transporter : Chloramphenicol	
rosab	Potassium antiporter system	
mepa, norm	MATE transporter	
tcma, mdr, qac	MFS transporter	
vana, vanb, vanc, vand, vane, vang	Vancomycin resistance operon genes (vanH, vanS, vanR, vanX, vanY) for each vancomycin resistance operons : VanA- ;VanB- ;VanC- ;VanD- ;VanE- and VanG-types	Vancomycin
tet_rpp	Tetracycline ribosomal protection protein	Tetracycline
tet_efflux	Tetracycline-specific efflux pump	
tet_flavo	Flavoproteins resistance to Tetracycline	
bla_a, bla_b	Class A and Class B $\beta$ -lactamases	$\beta$ -lactams
pbp	Penicillin-binding protein	
baca	Bacitracin resistance	Bacitracin
cata	Chloramphenicol acetyltransferase	Chloramphenicol
ksga	Kasugamycin resistance	Kasugamycin
arna	Polymixin resistance	Polymixin
pur8	Puromycin resistance	Puromycin
vat	Virginiamycin resistance	Streptogramin
sul	Sulfonamide resistance	Sulfonamide
dfra	Trimethoprim resistance	Trimethoprim

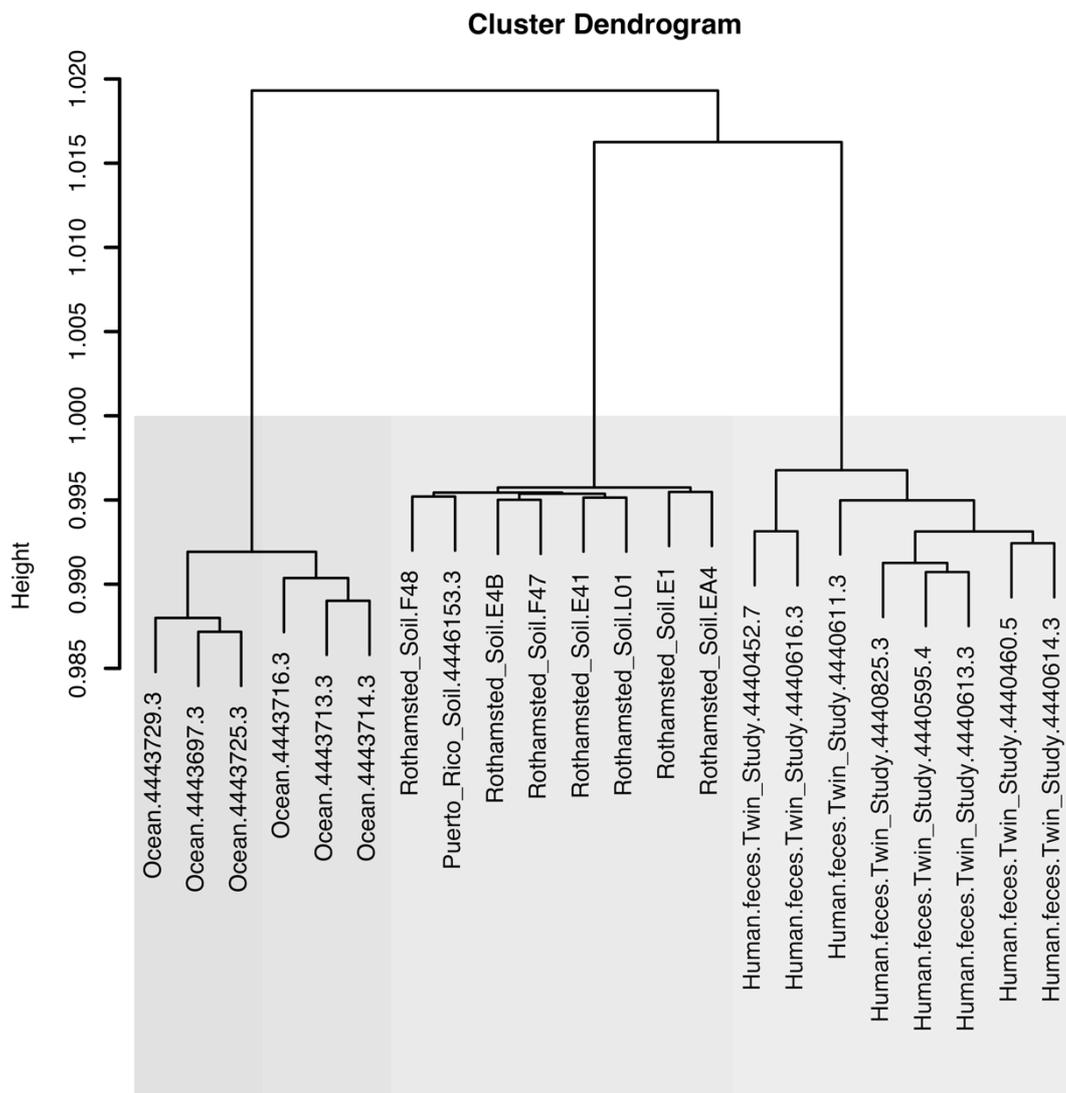
**Table 1. – Details for the 50 most prevalent resistance types found in all metagenomes.**

Most prevalent resistance type found in our analysis are MDR efflux pumps, vancomycin resistance, tetracycline resistance,  $\beta$ -lactams resistance and resistance to other antibiotic molecules: Polymixin ; Bacitracin ; Sulfonamide ; Kasugamycin ; Puromycin ; Trimethoprim ; Streptogramin ; Chloramphenicol. Resistance mechanisms involved and antibiotic molecule targets are specified for each group of the 50 resistance classes found and their respective abundance is indicated in supplemental data table S2 and figure S1.

### **ARGD distribution shows distinct environmental clusters.**

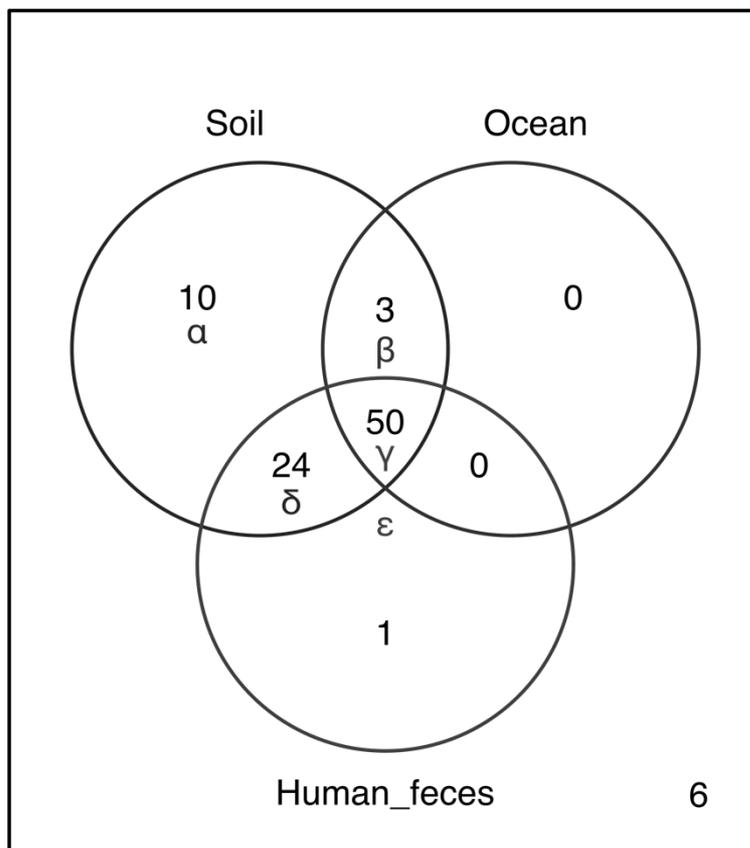
In order to overcome the bias related to the sequencing method for ecosystem comparison, an additional ARGD distribution analysis based on a clustering method was carried out on a restricted subset of metagenomic sequences obtained by Roche 454 pyrosequencing. Sanger datasets were insufficiently homogeneous, and therefore, were discarded from the analysis. The dendrogram obtained was plotted using Ward hierarchical classification of the datasets Mountford index (see supplemental Experimental Procedures). The Mountford index reflects similarities between samples based on ARGD sequence presence-absence for each metagenome datasets (i.e. ARGD  $\beta$ -diversity is defined as a function of alpha- and gamma- diversities). The different datasets used in the analysis were clustered in three different groups that matched their environmental origin: ocean group, soil group and human feces group (Figure 2). Soil datasets were similar regardless of the soil geographical origin or usage (a prairie soil in England or a forest soil in Puerto Rico) whereas two separate clusters were identified for ocean datasets. Open ocean datasets grouped separately from coastal ocean datasets. The Venn diagram of resistance class level annotations showed a common set of antibiotic resistance genes shared between all three environmental clusters (Figure 3, see also Table S3). Among all different ARGD types (n=94), a majority (n=50) was common to all environments and less than 10% (n=6) were not found anywhere. The six types of ARGD not detected in any cluster were: *dfrb* (group B trimethoprim resistant dihydrofolate reductase), *fusb* (fusaric acid resistance), *mph* (macrolide phosphotransferase); *rif* (viral rifampin resistance), *vph* (viomycin phosphotransferase) and *tet\_other* (tetracycline resistance genes that are not efflux pumps or ribosomal binding proteins). Ocean and human feces clusters did not share any ARGD type. Multidrug efflux pump class *Imrp* was found only in the human feces cluster. ARGDs

shared between soil and ocean clusters (n=3) were genes conferring resistance to fosfomycin, puromycin and quinolones. A large subset (n=24) of ARGD classes was shared between human feces and soil clusters. This subset includes all classes of  $\beta$ -lactamases, resistances to aminoglycoside through antibiotic molecules modification (phosphotransferases, acetyltransferases or nucleotidyltransferases), various multidrug efflux pumps and flavoproteins conferring resistance to tetracycline.



**Figure 2 - Hierarchical clustering based on Mountford diversity index of ARGD among soil, ocean and human feces Roche 454 metagenomic datasets.**

Mountford index has been computed from the distribution of ARGD hits in each metagenomic dataset considering each different ARGD as a different species. Hierarchical clustering based on ward method has been performed on Mountford index matrix.



**Figure 3 - Common ARGD class between four different environmental clusters based on Roche 454 Datasets.**

Numbers inside each circles intersections are indicating number of common resistance classes between environmental datasets. **Only in soil ( $\alpha$ ):** mls\_hdr, pmra, qacab, qnr, tmr, ble, bmr, emea, ere, fush. **Common to Soil and Ocean only  $\beta$ :** fos (fosfomycin resistance) ; pac (puromycin resistance) ; mfpa (quinolones resistance). **Common to Soil, Ocean and Human Feces ( $\gamma$ ):** acr, adeabc, amr, macab, mdtef, mdtk, mdtm, mdtnop, mexab, mexcd, mexef, mexhi, mexvw, mexxy, smeabc, smedef (RND efflux pumps), arna (polymixin resistance), baca, bcr, bcr\_mfs (bacitracin resistance), catb, ceo, cml (chloramphenicol resistance), dfra (group A trimethoprim resistant dihydrofolate reductase), emrd, emre, mdr, nora, norm, rosab, tcma (MFS transporters), erm, lsa (macrolide resistance), ksga (kasugamycin resistance), meca, pbp ( $\beta$ -lactams resistance), mls\_abc (ABC family transporter), pur8 (puromycin resistance), qac (quaternary ammonium compound efflux pump), sul (sulfonamide-resistant dihydropteroate synthase), tet\_efflux, tet\_xprt (tetracycline efflux pumps), tet\_rpp (tetracycline ribosomal protection protein), vana, vanb, vanc, vand, vane, vang (vancomycin resistance), vat (virginiamycin A acetyltransferase). **Common to Soil and Human Feces ( $\delta$ ):** aac, ant, aph (aminoglycoside transferases), bla\_a, bla\_b, bla\_c, bla\_d ( $\beta$ -lactamases), blt, cata (chloramphenicol acetyltransferase), lnu (lincomycin nucleotidyltransferase) , mdfa, mdtg, mdth, mdtl, mepa, mls\_mfs, ykk (Multidrug efflux pump), mecrl (methicillin resistance), sta, str (streptomycin resistance), tet\_flavo, tet\_mod (resistance to tetracycline), vanz (vancomycin resistance) and tsnr (thiosprepton resistance). **Only in Human Feces ( $\epsilon$ ):** lmrp (MFS transporter).

## Discussion.

The objective of our study was to evaluate the quantity and diversity of antibiotic resistance genes from different environments using the metagenomic datasets from various biomes (soils, ocean, human gut, etc.). Metagenomic sequences were annotated as ARGD if their sequence matches a set of known antibiotic resistance genetic determinants (ARDB). Our results show that soil environments have the largest apparent diversity of ARGDs while gut-associated microbiota have a higher proportion of ARGD-annotated reads, which may be due to exposure to antibiotic molecules that increased the overall antibiotic resistance of the commensal flora [32, 33].

Efflux pumps dominated the best-hit-only annotation (*i.e.*, not removing best-hits below a bit-score and percent identity threshold of respectively above 60% and 35%). Antibiotic-specific efflux pumps share homologies at the amino-acid level with other pump genes and share functional domains with every efflux pump protein (e.g., membrane-fusion protein (MFP) domain). Efflux pumps are present in both antibiotic-susceptible and antibiotic-resistant bacteria and can be poor markers of resistance phenotype. Efflux-pump-related resistance is often conferred by point mutations in their regulatory genes [34, 35]. However, in our analysis, most detected efflux-pump reads were related to efflux of antibiotic or toxic compounds even when re-examined with NCBI\_nr (Figure S3). This implied that the antibiotic resistance hit was not superseded by a hit with generalist database. Some ARGD protein families with many different variants, such as  $\beta$ -lactamases or efflux pumps, were more abundantly referenced in our database and that might have increased detection potential compared to ARGD with few representatives in the ARDB. ARGDs, such as Vancomycin resistance operons,  $\beta$ -lactamases or tetracycline ribosomal

protection proteins, usually confer the resistance phenotype directly when translated. This makes these ARGD-annotated reads better markers of potential antibiotic resistance phenotypes. ARGDs were systematically detected in every environment screened despite all the described limitations. ARGDs are probably even more abundant in all of the ecosystems than what has been observed with sequence-based analysis. Thus, we believe our method retrieved many bona fide ARGDs from the environmental resistome.

Metagenomes obtained with the same technology from three distinct environmental groups (ocean, soil and human feces) on the basis of presence/absence of annotated ARGD sequences (Figure 2). While many ARGDs were shared between all environments (Figure 3), soil and human feces shared more (24 out of 94) ARGD classes (*i.e.* groups of sequences with shared antibiotic resistance-conferring mechanisms as defined in ARDB) than any other pairs of ecosystems. This heightened degree of shared ARGDs between soil and human feces datasets has been previously explored by functional metagenomics and led to the conclusion that gene flow between these environments is likely [1, 3, 17]. Soil is probably the major source of antibiotic-producing soil microorganisms (e.g., *Streptomyces chrysogenum* producing penicillin) and they live in apparently one of the most diverse environments in terms of ARGDs (Figure 1). Considering this diversity, the environmental and more specifically the soil reservoir of antibiotic resistance genes could be major contributors to antibiotic resistance mechanisms acquisition, dissemination and diversification in pathogenic bacteria populations, by means of horizontal gene transfer.

The potential ecological pressure that low natural concentrations of antibiotics might exert has important implications concerning the role of antibiotics molecules and ARGDs in bacterial ecology. In environments without anthropogenic amendment of antibiotics molecules, such as the Rothamsted soil or the Alaskan permafrost, antibiotic molecule

concentrations were not detected, yet functional antibiotic resistance genes were always detected [4, 6, 7]. In addition to their antibacterial activity, antibiotic molecules have been found to act as signal molecules between bacterial cells and many antibiotic resistance genes are themselves involved in diverse bacterial physiological functions, such as gene transcription and stress-response regulation [8, 36-40]. Their possible multifunctional role associated to a low fitness costs (and eventually fitness benefits independent of antibiotic resistance [41]) might explain the prevalence of most ARGDs in the environment [42]. Indeed, maintenance in their host may not require a constant selective pressure, making environmental bacteria a persistent source of ARGDs for pathogens [14]. Human feces datasets showed that antibiotic resistance genes were present in human gut resident bacteria with as much as 5.6% ARGD-annotated reads. In human gut, other mechanisms of antibiotic resistance gene selection besides inhibition resistance could be involved. The most abundant ARGD-annotated reads detected in human datasets were vancomycin resistance genes even though vancomycin is a strongly regulated and rarely used "last resort" antibiotic in medicine. In contrast, Amoxicillin related resistance genes ( $\beta$ -lactamases) were much less abundant in human feces, although this drug is commonly used.

The ecological role of every ARGD in the environment has yet to be elucidated, but some ARGDs have been shown to play central roles, such as inter-cellular signaling for their original (*i.e.* not clinical) saprophyte host [37]. Antibiotic molecules and antibiotic resistance are an inherent part of common bacterial eco-physiology in the environment. This ARGD resource is selected in pathogens when they acquire this trait and enhance their fitness in the case of antibiotherapy. Yet, not all common pathogens are resistant to antibiotics. With the efficiency of bacteria to adapt to the antibiotic challenge via mutations, DNA recombination, plasmid- and integron- mediated gene transfer; they are outpacing the

health industry in terms of development and innovation. Further studies on the mechanism of antibiotic resistance dispersion are critical for enhancing antibiotic clinical lifetimes. What role specific mobile genetic elements play? What do they carry? These are crucial questions for future of antibiotic therapy development. Antibiotic resistance genes in the environment (and clinical antibiotic resistance) are not likely to disappear in the near future, although selection mechanisms seem to limit dissemination of antibiotic resistance. Deciphering these mechanisms may provide the tools for enhancing the long-term efficiency of currently available and future antibiotics.

## **Acknowledgment.**

We thank Dr. Annie Chaboud for critical reading of the revised manuscript. We thank Penny Hirsch and Ian Clark from Rothamsted Research (UK) for the soil used in the ANR project METASOIL. This work was financed by grants from the French Agency for the Environment and Energy Management (ADEME – project “Générique”) and from French Agency for Food, Environmental and Occupational Health & Safety (ANSES - project “Impactance”, convention n° EST-2012/2/28). The French Ministry of Higher Education and Research funded JN. TOD was supported by a doctoral fellowship from the Rhône Alpes Région. This work is accepted for publication in *Current Biology* in the form of a report.

## **References.**

1. Ochman, H., Lawrence, J. G., and Groisman, E. A. (2000). Lateral gene transfer and the nature of bacterial innovation. *Nature* *405*, 299–304.
2. Cantón, R., and Morosini, M.-I. (2011). Emergence and spread of antibiotic resistance following exposure to antibiotics. *FEMS Microbiol Rev* *35*, 977–991.
3. Lawrence, J. G., and Roth, J. R. (1996). Selfish operons: horizontal transfer may drive the evolution of gene clusters. *Genetics* *143*, 1843–1860.
4. Allen, H. K., Moe, L. A., Rodbumrer, J., Gaarder, A., and Handelsman, J. (2009). Functional metagenomics reveals diverse beta-lactamases in a remote Alaskan soil. *ISME J* *3*, 243–251.
5. Demanèche, S., Sanguin, H., Poté, J., Navarro, E., Bernillon, D., Mavingui, P., Wildi, W., Vogel, T. M., and Simonet, P. (2008). Antibiotic-resistant soil bacteria in transgenic plant fields. *Proc Natl Acad Sci USA* *105*, 3957–3962.

6. Bhullar, K., Waglechner, N., Pawlowski, A., Koteva, K., Banks, E. D., Johnston, M. D., Barton, H. A., and Wright, G. D. (2012). Antibiotic resistance is prevalent in an isolated cave microbiome. *PLoS ONE* 7, e34953.
7. D'Costa, V. M., King, C. E., Kalan, L., Morar, M., Sung, W. W. L., Schwarz, C., Froese, D., Zazula, G., Calmels, F., Debruyne, R., et al. (2011). Antibiotic resistance is ancient. *Nature* 477, 457–461.
8. Martinez, J. L., Fajardo, A., Garmendia, L., Hernandez, A., Linares, J. F., Martínez-Solano, L., and Sánchez, M. B. (2009). A global view of antibiotic resistance. *FEMS Microbiol Rev* 33, 44–65.
9. Humeniuk, C., Arlet, G., Gautier, V., Grimont, P., Labia, R., and Philippon, A. (2002). Beta-lactamases of *Kluyvera ascorbata*, probable progenitors of some plasmid-encoded CTX-M types. *Antimicrob Agents Chemother* 46, 3045–3049.
10. Toleman, M. A., Spencer, J., Jones, L., and Walsh, T. R. (2012). BlaNDM-1 is a chimera, likely constructed in *Acinetobacter baumannii*. *Antimicrob Agents Chemother*.
11. Aminov, R., and Mackie, R. (2007). Evolution and ecology of antibiotic resistance genes. *FEMS Microbiol Lett* 271, 147–161.
12. Eisen, J. A. (2007). Environmental Shotgun Sequencing: Its Potential and Challenges for Studying the Hidden World of Microbes. *PLoS Biol* 5, e82.
13. Martínez, J. L. (2013). Bacterial pathogens: from natural ecosystems to human hosts. *Environ Microbiol* 15, 325–333.
14. Perry, J. A., and Wright, G. D. (2013). The antibiotic resistance “mobilome”: searching for the link between environment and clinic. *Front Microbiol* 4, 138.
15. Davies, J., and Davies, D. (2010). Origins and Evolution of Antibiotic Resistance. *Microbiology and Molecular Biology Reviews* 74, 417–433.
16. D'Costa, V. M., Griffiths, E., and Wright, G. D. (2007). Expanding the soil antibiotic resistome: exploring environmental diversity. *Curr Opin Microbiol* 10, 481–489.
17. Forsberg, K. J., Reyes, A., Wang, B., Selleck, E. M., Sommer, M. O. A., and Dantas, G. (2012). The Shared Antibiotic Resistome of Soil Bacteria and Human Pathogens. *Science* 337, 1107–1111.
18. Walsh, F., and Duffy, B. (2013). The Culturable Soil Antibiotic Resistome: A Community of Multi-Drug Resistant Bacteria. *PLoS ONE* 8, e65567.
19. D'Costa, V. M., McGrann, K. M., Hughes, D. W., and Wright, G. D. (2006). Sampling the antibiotic resistome. *Science* 311, 374–377.
20. Schloss, P. D., and Handelsman, J. (2005). Metagenomics for studying unculturable microorganisms: cutting the Gordian knot. *Genome Biol* 6, 229.

21. Whitman, W. B., Coleman, D. C., and Wiebe, W. J. (1998). Prokaryotes: the unseen majority. *Proc Natl Acad Sci USA* *95*, 6578–6583.
22. Monier, J.-M., Demanèche, S., Delmont, T. O., Mathieu, A., Vogel, T. M., and Simonet, P. (2011). Metagenomic exploration of antibiotic resistance in soil. *Curr Opin Microbiol* *14*, 229–235.
23. Meyer, F., Paarmann, D., D'Souza, M., Olson, R., Glass, E. M., Kubal, M., Paczian, T., Rodriguez, A., Stevens, R., Wilke, A., et al. (2008). The metagenomics RAST server - a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics* *9*, 386.
24. Delmont, T. O., Prestat, E., Keegan, K. P., Faubladièr, M., Robe, P., Clark, I. M., Pelletier, E., Hirsch, P. R., Meyer, F., Gilbert, J. A., et al. (2012). Structure, fluctuation and magnitude of a natural grassland soil metagenome. *ISME J* *6*, 1677–1687.
25. Niu, B., Fu, L., Sun, S., and Li, W. (2010). Artificial and natural duplicates in pyrosequencing reads of metagenomic data. *BMC Bioinformatics* *11*, 187.
26. Liu, B., and Pop, M. (2009). ARDB--Antibiotic Resistance Genes Database. *Nucleic Acids Res* *37*, D443–7.
27. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J Mol Biol* *215*, 403–410.
28. Looft, T., Johnson, T. A., Allen, H. K., Bayles, D. O., Alt, D. P., Stedtfeld, R. D., Sul, W. J., Stedtfeld, T. M., Chai, B., Cole, J. R., et al. (2012). In-feed antibiotic effects on the swine intestinal microbiome. *Proc Natl Acad Sci USA* *109*, 1691–1696.
29. Overbeek, R. A., Begley, T., Butler, R. M., Choudhuri, J. V., Chuang, H.-Y., Cohoon, M., de Crécy-Lagard, V., Diaz, N., Disz, T., Edwards, R. A., et al. (2005). The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. *Nucleic Acids Res* *33*, 5691–5702.
30. Dixon, P. (2003). VEGAN, a package of R functions for community ecology. *J Veg Sci* *14*, 927–930.
31. Wooley, J. C., Godzik, A., and Friedberg, I. (2010). A Primer on Metagenomics. *PLoS Comput Biol* *6*, –.
32. Woerther, P.-L., Angebault, C., Jacquier, H., Clermont, O., Mnai, El, A., Moreau, B., Djossou, F., Peroz, G., Catzeflis, F., Denamur, E., et al. (2013). Characterization of fecal extended-spectrum- $\beta$ -lactamase-producing *Escherichia coli* in a remote community during a long time period. *Antimicrob Agents Chemother* *57*, 5060–5066.
33. Forslund, K., Sunagawa, S., Roat Kultima, J., Mende, D., Arumugam, M., Typas, A., and Bork, P. (2013). Country-specific antibiotic use practices impact the human gut resistome. *Genome Res*.

34. Martinez, J. L., Sánchez, M. B., Martínez-Solano, L., Hernandez, A., Garmendia, L., Fajardo, A., and Alvarez-Ortega, C. (2009). Functional role of bacterial multidrug efflux pumps in microbial natural ecosystems. *FEMS Microbiol Rev* *33*, 430–449.
35. Piddock, L. J. V. (2006). Clinically relevant chromosomally encoded multidrug resistance efflux pumps in bacteria. *Clin Microbiol Rev* *19*, 382–402.
36. Aminov, R. I. (2009). The role of antibiotics and antibiotic resistance in nature. *Environ Microbiol* *11*, 2970–2988.
37. Fajardo, A., and Martínez, J. L. (2008). Antibiotics as signals that trigger specific bacterial responses. *Curr Opin Microbiol* *11*, 161–167. Available at: <http://www.sciencedirect.com/science/article/pii/S1369527408000143>.
38. Wellington, E. M. H., Boxall, A. B., Cross, P., Feil, E. J., Gaze, W. H., Hawkey, P. M., Johnson-Rollings, A. S., Jones, D. L., Lee, N. M., Otten, W., et al. (2013). The role of the natural environment in the emergence of antibiotic resistance in gram-negative bacteria. *Lancet Infect Dis* *13*, 155–165.
39. Alonso, A., Sanchez, P., and Martinez, J. L. (2001). Environmental selection of antibiotic resistance genes. *Environ Microbiol* *3*, 1–9.
40. Allen, H. K., Donato, J., Wang, H. H., Cloud-Hansen, K. A., Davies, J., and Handelsman, J. (2010). Call of the wild: antibiotic resistance genes in natural environments. *Nat Rev Microbiol* *8*, 251–259.
41. Andersson, D. I., and Hughes, D. (2010). Antibiotic resistance and its cost: is it possible to reverse resistance? *Nat Rev Microbiol* *8*, 260–271.
42. Martinez, J. L. (2008). Antibiotics and Antibiotic Resistance Genes in Natural Environments. *Science* *321*, 365–367.

# Chapitre 3

## ***Expérimentation en microcosmes de sol de Rothamsted, Park Grass amendé en molécules antibiotique.***

### **Résumé.**

Dans le contexte de l'étude de la résistance à des antibiotiques dans l'environnement, le sol présente un intérêt particulier car de nombreuses observations de bactéries ont déjà été rapportées. En outre, il représente le compartiment environnemental duquel la majeure part des antibiotiques pharmaceutiques a été isolée et de nombreuses bactéries saprophytes du sol sont très résistantes à certaines molécules. Dans ce contexte, le sol semble être un réservoir considérable de gènes de résistance pour les pathogènes. La question de la dissémination des gènes de résistance depuis le réservoir environnemental vers les bactéries pathogènes est, en effet, un enjeu crucial de la recherche en écologie microbienne et en santé. Les bactéries pathogènes résistantes sont sélectionnées en réponse à la pression antibiotique exercée en milieu clinique; mais ces molécules et en particulier les antibiotiques ont un très fort taux d'excrétion. D'importantes quantités d'antibiotiques pharmaceutiques sont donc très introduites dans l'environnement par les activités humaines: santé, traitement des eaux usées, épandage de fumier provenant d'animaux traités par des antibiotiques par exemple. Afin d'évaluer l'impact de ce type de pollution dans le sol, une série de microcosmes amendés soit par des fumiers, soit par des molécules antibiotiques a été réalisée. Ces microcosmes ont été analysés pendant 6 mois d'incubation et nos résultats montrent que, outre les modifications de communauté liées à l'incubation en microcosme, les effets des différents amendements sont significativement différents. La communauté microbienne du sol est aussi drastiquement affectée par l'introduction de grandes quantités de bactéries d'origine fécales qui modifient non seulement la communauté initiale, mais aussi le développement de cette communauté en microcosme.

# Chapter 3

## ***Antibiotic-amended Rothamsted, Park Grass soil microcosms experiment.***

### **Abstract.**

Most pharmaceuticals antibiotics used nowadays were extracted from soil bacteria or fungi and soil harbors an abundant share of highly resistant saprophytic bacteria. In this context, soil seems to be a major reservoir of antibiotic-resistance genes for pathogens. The question of antibiotic-resistance gene dissemination from the environmental reservoir to pathogens is still an open question in microbial ecology with crucial implications in the sustainable management of our still currently active anti-infectious drugs. Pathogens are selected for antibiotic-resistance when a selective-pressure is applied, such as in the clinic. However, antibiotics are designed especially for easy excretion by patient's or animal's organism. Large quantities of antibiotics and other pharmaceuticals are introduced by human activities in the environment, notably through soil fertilization using manure from treated animals or hospital sewage systems for example. It is in this context of the utmost importance to properly evaluate the consequences of such pollution on the natural soil microbial community. In this respect, a series of Rothamsted, Park Grass soil microcosms amended with either pure-antibiotic or manure from treated animals were setup and incubated for 6 months and sampled at T0, T1, T3 and T6 months. Metagenomic DNA prokaryotes community profile is assessed using RISA-PCR and antibiotic-resistance genes abundance is estimated using quantitative PCR in each microcosms condition over the 6 months incubation period. Results show that the soil community is drastically impacted by the microcosm incubation and in lesser extent by antibiotic-molecules pollution. However, most important changes in bacterial community were due to massive addition of living fecal bacteria introduced in soil in microcosms, drastically changing the initial microbial community found in soil and its evolution in microcosms amended with manure.

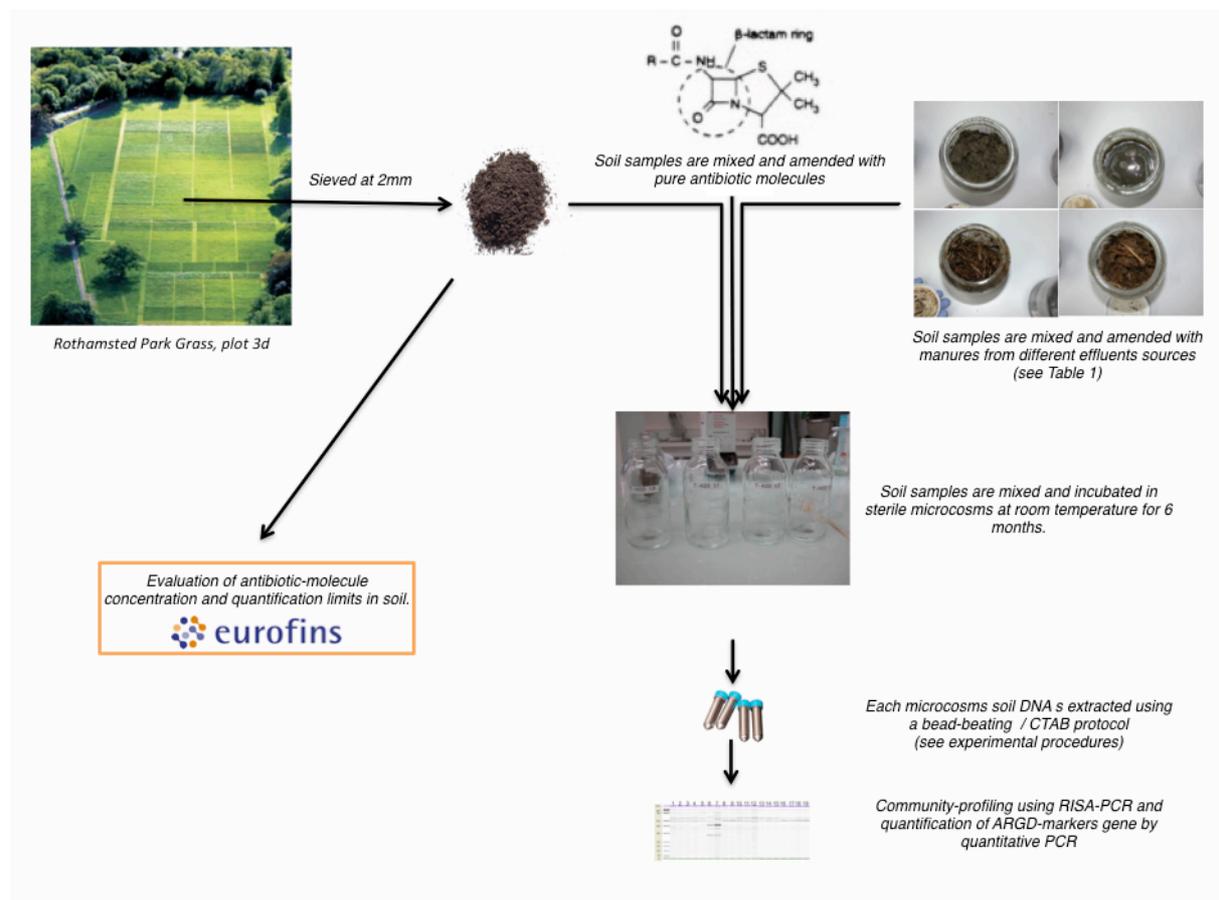
## Introduction.

Most of antibiotic molecules currently used are derived from bioactive molecules produced by soil microorganisms [1, 2]. This chemical arsenal is for example selected as a way for producers to increase their competitive advantage in a local competition by inhibiting the growth of neighboring cells [3]. This selection for toxic molecule secretion is inevitably leading to countermeasures such as antibiotic-resistance genes. These are selected in soil producer bacteria, to resist their own arsenal [4], and in all local bacteria to overcome this toxic competition. Genomes from environmental isolates of antibiotic-producing and antibiotic-resistant bacteria show remarkable signs of intense rearrangements and rapid mutations through incorporation of mobile genetic elements, such as transposons, integrons or plasmids [5]. These observations are coherent with an intense arms-shield race in the soil environment and claim for an abundant reservoir of antibiotic-resistant bacteria in soil, proportional to its share of antibiotic-producers [2, 6-8]. Antibiotic-resistant bacteria are indeed common in soil, as previously shown by culture [9, 10] or heterologous host expression of DNA [11, 12] from environmental soil samples. More importantly for the problematic clinical aspect of antibiotic-resistance, leading to antibiotic treatment failure: many genes isolated from soil bacteria are also found in the human-gut microbiome or clinical-isolates [13-16]. These observations question the selection and recruitment of antibiotic-resistance genes in pathogen genomes and the implication of the environmental resistome [17]. This subject is even more important when sub-inhibitory trace amounts of antibiotic molecules are shown to significantly increase the selection of antibiotic-resistance, but also affects multiple aspects of bacterial physiology such as quorum-sensing behavior (*i.e.*: cell-cell communication) or transcription activity [18, 19]. All

these effects *in fine* considerably affects the soil microbial community in many different aspects [20]. This is problematic for proper risks assessments of antibiotic molecules dispersal in this environment, as our understanding of the complex soil ecology is still very limited considering the extent of its diversity [21, 22]. Antibiotic-molecules have been measured in many environmental matrices, either aquatic [23] or more complex such as wastewater treatment plants (WWTP) activated-sludge [24], animal manure [25-27] or soil [28]. These different samples are indeed representative of the major part of antibiotic molecules entry routes in soil [25, 29, 30]. In these reviews, authors compiled results from a number of studies and showed that increases of measured pharmaceuticals antibiotics in soil are mostly due to soil fertilization using antibiotic-contaminated manure from urban-wastewater treatment plants or manure from animals treated with veterinary antibiotics. Consequently, a number of studies investigated the augmentation of antibiotic-resistance gene abundance by quantitative PCR assays in the vicinity of impacted areas, in relation to antibiotic pollution: tetracycline ARGD in industrial poultry [31] or swine farms [32], diverse resistance genes in densely populated WWTP [33-35] urban [36] or amoxicillin resistance genes in pig manure [37]. These studies all measured significant amounts of antibiotic-resistance gene copies in soil correlated that is in some case correlated with antibiotic-concentrations measured in the same sample. All these studies reveal, however, reveal the intrinsic soil complexity that should be further investigated in controlled conditions. A crucial aspect of the antibiotic-resistance problem is the dissemination, or recruitment, of antibiotic resistance genes in pathogenic bacteria. Considering the already measured impact of antibiotic-resistance in soil, it is now needed to model more precisely the dynamics of such impacts. Indeed, the complex soil ecology makes it very difficult to prove, and therefore assert bluntly, that antibiotic-resistance molecules are selecting antibiotic-

resistance genes in soil following its pollution. More importantly, entry of antibiotic-molecules in soils is mainly due to soil fertilization using manure from effluents highly concentrated in pharmaceutical antibiotics [28]. In this case, the pollution is also consisting of large amounts of commensal bacteria associated with fecal matter and may have also significant impact on the dissemination of antibiotic resistance genes. In order to verify the differential impacts on the soil microbial community of either pure molecules or manure, a microcosm incubation strategy was setup. Microcosm experiments were coupled with quantitative PCR assays and community fingerprinting (RISA-PCR) performed on soil sampled from each microcosm metagenomic DNA extracted at different incubation time-points (T0; T1: one month; T3: 3 months; T6: six months). The pasture soil used in this study is a reference soil monitored by a long-term experiment since 1879 (i.e.: 72 years before penicillin discovery in 1928), previously used in a deep-sequencing study [38]. This soil was shown to be abundant in antibiotic-resistance genes by means of direct sequencing of its metagenomic DNA [39]. In addition, each antibiotic molecule concentration used as amendment in this study was measured during incubation, to follow antibiotics degradation in soil in respect to microbial community modifications. Our observations are consistent with an abundant antibiotic-resistance reservoir hypothesis, even in soils not impacted by antibiotics. This justify Rothamsted soil usage as a typical naturally antibiotic-resistance containing soil. Our objective is to determine if antibiotic molecules or manure amendments interfere with the natural dynamic of antibiotic-resistance gene abundance and evolution of the microbial community. Can our results correlate increase in antibiotic-resistance genes and mobile-genetic elements such as *int1* and *trfA*, known to be implied in resistance transfer [40, 41]? The dynamic of two different mobile-genetic element markers (*int1* for Class 1 integrons and *trfA* for IncP broad-host plasmid) were also examined to find

significant differences between microcosms, underlying their potential role in rapid bacterial adaptation following drastic environmental perturbations. Our objective in this study was to simplify the soil ecosystem (all things being equal), in order to rigorously assess the different effects of different molecules. Is this objective however realist when considering that microcosm incubation itself is shown to significantly alter the microbial community in the absence of any amendment?



**Figure 1 - Experimental design used in this study to assess impacts of a controlled antibiotic molecule and manure amendment on a reference soil: Rothamsted, Park Grass soil.**

Soil samples collected from the Rothamsted, Park Grass long-term experiment control plot (3d) are amended with either pure antibiotic molecules (ampicillin, tetracycline, trimethoprim, sulfamethoxazole, enrofloxacin) or farm effluents (from treated and non treated pigs, treated cows and treated chickens with antibiotics). All microcosms were setup in 250ml sterile glass bottles using 180g of soil; incubated for 6 months at room-temperature and sampled at several time points: T0, T1, T3 and T6, with each number corresponding to incubation duration in months.

## Experimental procedures.

### Antibiotics molecules and farm effluents tested in this study.

Molecules used for soil microcosm amendments in this study represent a large spectrum of different antibiotic classes widely used either in human or veterinary medicine: tetracycline, sulfamethoxazol, sulfamethazin, trimethoprim, tylosin enrofloxacin and ampicillin. In addition, microcosms were also amended with a selection of farm effluents with varying characteristics: liquid manure from pigs treated with tetracycline and tylosin (sample PT2; 200 animals treated); feces from pigs (sample PNT), manure from cows treated with tetracycline (sample B1T1; mixed with straw; 20 animals out of 150 treated); feces from chicken treated with enrofloxacin (23000 animals treated for 5 days).

### Microcosms setup.

Soil samples are provided by the courtesy of Ian M. Clark and Penny R. Hirsh, Rothamsted Research, Hertfordshire, UK. A total of 15kg of Rothamsted soil collected in March 2011, stored at 4°C for shipping and sieved at 2mm upon reception. For the sake of clarity, Figure 1 displays a general schematic of the experimental procedures described below. Soils are amended at varying molecule or effluent concentrations and simulate concentration recorded in the literature and based on antibiotic concentration found in farm effluents samples, antibiotic molecule concentrations in such samples and standard soil manure fertilization practices [28] (Table 1, Figure 1). This led us to use a 2% (w/w) amendment concentration for a final 180g total weight microcosm. Two controls conditions are prepared. One is a microcosm condition control and setup by simply introducing sieved Rothamsted soil in a sealed bottle (Control) while the other one is supplemented with 4ml of sterile water (Control H<sub>2</sub>O). This volume corresponds to the amount of liquid added in

Microcosm	Total mass	Amendment (%: w/w; else [µg/kg])	[C] <sub>ATB</sub> (mg/mL)	Amendment (mg)	V <sub>AB</sub> (ml)	Soil (g)	T0	T1	T3	T6	[Amp.] µg/kg	[Enro.] µg/kg	[Tet.] µg/kg	[Tylo.] µg/kg	[Trim.] µg/kg	[Sulf.] µg/kg
T-T0	180g	0.0%	NA	0	NA	180.0	X				<LQ	<LQ	<LQ	absent	<LQ	<LQ
T-T1	180g	0.0%	NA	0	NA	180.0		X			n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
T-T3	180g	0.0%	NA	0	NA	180.0			X		<LQ	absent	absent	<LQ	<LQ	<LQ
T-T6	180g	0.0%	NA	0	NA	180.0				X	<LQ	absent	absent	<LQ	absent	<LQ
T-H2O-T0	180g	0.0%	NA	4ml H2O	NA	176.0	X				<LQ	<LQ	<LQ	absent	<LQ	<LQ
T-H2O-T1	180g	0.0%	NA	4ml H2O	NA	176.0		X			n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
T-H2O-T3	180g	0.0%	NA	4ml H2O	NA	176.0			X		<LQ	absent	absent	trace	<LQ	<LQ
T-H2O-T6	180g	0.0%	NA	4ml H2O	NA	176.0				X	<LQ	absent	absent	<LQ	absent	<LQ
PorcsT2-T0	180g	2.20%	NA	4000	NA	176.0	X				<LQ	<LQ	<LQ	absent	<LQ	<LQ
PorcsT2-T1	180g	2.20%	NA	4000	NA	176.0		X			n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
PorcsT2-T3	180g	2.20%	NA	4000	NA	176.0			X		<LQ	absent	absent	<LQ	<LQ	<LQ
PorcsT2-T6	180g	2.20%	NA	4000	NA	176.0				X	<LQ	absent	absent	<LQ	absent	<LQ
PorcsNT-T0	180g	2.20%	NA	4000	NA	176.0	X				<LQ	<LQ	<LQ	absent	<LQ	<LQ
PorcsNT-T1	180g	2.20%	NA	4000	NA	176.0		X			n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
PorcsNT-T3	180g	2.20%	NA	4000	NA	176.0			X		<LQ	absent	absent	<LQ	<LQ	<LQ
PorcsNT-T6	180g	2.20%	NA	4000	NA	176.0				X	<LQ	absent	absent	<LQ	absent	<LQ
Bovins1T1-T0	180g	2.20%	NA	4000	NA	176.0	X				<LQ	<LQ	<LQ	<LQ	<LQ	<LQ
Bovins1T1-T1	180g	2.20%	NA	4000	NA	176.0		X			n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
Bovins1T1-T3	180g	2.20%	NA	4000	NA	176.0			X		<LQ	absent	absent	<LQ	<LQ	<LQ
Bovins1T1-T6	180g	2.20%	NA	4000	NA	176.0				X	<LQ	absent	absent	<LQ	absent	<LQ
Aviaires2T1-2-T0	180g	2.20%	NA	4000	NA	176.0	X				<LQ	<LQ	<LQ	absent	<LQ	<LQ
Aviaires2T1-2-T1	180g	2.20%	NA	4000	NA	176.0		X			n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
Aviaires2T1-2-T3	180g	2.20%	NA	4000	NA	176.0			X		<LQ	absent	absent	<LQ	<LQ	<LQ
Aviaires2T1-2-T6	180g	2.20%	NA	4000	NA	176.0				X	<LQ	absent	absent	<LQ	absent	<LQ
Aviaires2T1-40-T0	180g	40.0%	NA	72000	NA	108.0	X				<LQ	+	<LQ	absent	<LQ	<LQ
Aviaires2T1-40-T1	180g	40.0%	NA	72000	NA	108.0		X			n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
Aviaires2T1-40-T3	180g	40.0%	NA	72000	NA	108.0			X		<LQ	absent	absent	<LQ	<LQ	<LQ
Aviaires2T1-40-T6	180g	40.0%	NA	72000	NA	108.0				X	<LQ	+	absent	<LQ	absent	<LQ
Amp-HC-T0	180g	222.00	1.025	0.040	0.039	176.0	X				<LQ	<LQ	<LQ	absent	<LQ	<LQ
Amp-HC-T1	180g	222.00	1.025	0.040	0.039	176.0		X			n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
Amp-HC-T3	180g	222.00	1.025	0.040	0.039	176.0			X		<LQ	absent	absent	<LQ	<LQ	<LQ
Amp-HC-T6	180g	222.00	1.025	0.040	0.039	176.0				X	<LQ	absent	absent	<LQ	absent	<LQ
Amp-LC-T0	180g	67.00	1.025	0.012	0.012	176.0	X				<LQ	<LQ	<LQ	absent	<LQ	<LQ
Amp-LC-T1	180g	67.00	1.025	0.012	0.012	176.0		X			n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
Amp-LC-T3	180g	67.00	1.025	0.012	0.012	176.0			X		<LQ	<LQ	absent	<LQ	<LQ	<LQ
Amp-LC-T6	180g	67.00	1.025	0.012	0.012	176.0				X	<LQ	<LQ	absent	<LQ	absent	<LQ
Enro-HC-T0	180g	1110.00	0.8	0.200	0.250	176.0	X				<LQ	+++	<LQ	<LQ	<LQ	<LQ
Enro-HC-T1	180g	1110.00	0.8	0.200	0.250	176.0		X			n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
Enro-HC-T3	180g	1110.00	0.8	0.200	0.250	176.0			X		<LQ	+	absent	<LQ	<LQ	<LQ
Enro-HC-T6	180g	1110.00	0.8	0.200	0.250	176.0				X	<LQ	absent	absent	<LQ	absent	<LQ
Enro-LC-T0	180g	330.00	0.8	0.060	0.075	176.0	X				<LQ	++	<LQ	<LQ	<LQ	<LQ
Enro-LC-T1	180g	330.00	0.8	0.060	0.075	176.0		X			n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
Enro-LC-T3	180g	330.00	0.8	0.060	0.075	176.0			X		<LQ	+	absent	<LQ	<LQ	<LQ
Enro-LC-T6	180g	330.00	0.8	0.060	0.075	176.0				X	<LQ	absent	absent	<LQ	absent	<LQ
Tet-HC-T0	180g	1110.00	1.051	0.200	0.190	176.0	X				<LQ	<LQ	<LQ	absent	<LQ	<LQ
Tet-HC-T1	180g	1110.00	1.051	0.200	0.190	176.0		X			n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
Tet-HC-T3	180g	1110.00	1.051	0.200	0.190	176.0			X		<LQ	absent	absent	<LQ	<LQ	<LQ
Tet-HC-T6	180g	1110.00	1.051	0.200	0.190	176.0				X	<LQ	absent	absent	<LQ	absent	<LQ
Tet-LC-T0	180g	220.00	1.051	0.040	0.038	176.0	X				<LQ	<LQ	<LQ	absent	<LQ	<LQ
Tet-LC-T1	180g	220.00	1.051	0.040	0.038	176.0		X			n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
Tet-LC-T3	180g	220.00	1.051	0.040	0.038	176.0			X		<LQ	absent	absent	<LQ	<LQ	<LQ
Tet-LC-T6	180g	220.00	1.051	0.040	0.038	176.0				X	<LQ	absent	absent	<LQ	absent	<LQ
Tylo-HC-T0	180g	555.00	0.956	0.100	0.105	176.0	X				<LQ	<LQ	<LQ	368	<LQ	<LQ
Tylo-HC-T1	180g	555.00	0.956	0.100	0.105	176.0		X			n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
Tylo-HC-T3	180g	555.00	0.956	0.100	0.105	176.0			X		<LQ	absent	absent	306	<LQ	<LQ
Tylo-HC-T6	180g	555.00	0.956	0.100	0.105	176.0				X	<LQ	absent	absent	23.4	absent	<LQ
Tylo-LC-T0	180g	110.00	0.956	0.020	0.021	176.0	X				<LQ	<LQ	<LQ	69.7	<LQ	<LQ
Tylo-LC-T1	180g	110.00	0.956	0.020	0.021	176.0		X			n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
Tylo-LC-T3	180g	110.00	0.956	0.020	0.021	176.0			X		<LQ	absent	absent	trace	<LQ	<LQ
Tylo-LC-T6	180g	110.00	0.956	0.020	0.021	176.0				X	<LQ	absent	absent	<LQ	absent	<LQ
Trimethoprim-HC-T0	180g	220.00	1.152	0.040	0.035	176.0	X				<LQ	<LQ	<LQ	absent	278.5	<LQ
Trimethoprim-HC-T1	180g	220.00	1.152	0.040	0.035	176.0		X			n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
Trimethoprim-HC-T3	180g	220.00	1.152	0.040	0.035	176.0			X		<LQ	absent	absent	<LQ	121.6	<LQ
Trimethoprim-HC-T6	180g	220.00	1.152	0.040	0.035	176.0				X	<LQ	absent	absent	<LQ	++	<LQ
Trimethoprim-LC-T0	180g	22.00	1.152	0.004	0.003	176.0	X				<LQ	<LQ	<LQ	absent	24.7	<LQ
Trimethoprim-LC-T1	180g	22.00	1.152	0.004	0.003	176.0		X			n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
Trimethoprim-LC-T3	180g	22.00	1.152	0.004	0.003	176.0			X		<LQ	absent	absent	<LQ	15.1	<LQ
Trimethoprim-LC-T6	180g	22.00	1.152	0.004	0.003	176.0				X	<LQ	absent	absent	<LQ	trace	<LQ
Sulfamethoxazole-HC-T0	180g	110.00	1.015	0.020	0.020	176.0	X				<LQ	<LQ	<LQ	absent	<LQ	40.9
Sulfamethoxazole-HC-T1	180g	110.00	1.015	0.020	0.020	176.0		X			n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
Sulfamethoxazole-HC-T3	180g	110.00	1.015	0.020	0.020	176.0			X		<LQ	absent	absent	<LQ	<LQ	trace
Sulfamethoxazole-HC-T6	180g	110.00	1.015	0.020	0.020	176.0				X	<LQ	absent	absent	<LQ	absent	<LQ
Sulfamethoxazole-LC-T0	180g	22.00	1.015	0.004	0.004	176.0	X				<LQ	<LQ	<LQ	absent	<LQ	trace
Sulfamethoxazole-LC-T1	180g	22.00	1.015	0.004	0.004	176.0		X			n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
Sulfamethoxazole-LC-T3	180g	22.00	1.015	0.004	0.004	176.0			X		<LQ	absent	absent	<LQ	<LQ	trace
Sulfamethoxazole-LC-T6	180g	22.00	1.015	0.004	0.004	176.0				X	<LQ	absent	absent	<LQ	absent	<LQ

n.d.: not determined; <LQ: below reliable quantification limits (2< [C] < 6 µg/kg dry matter); presence (+), (++) , (+++): molecule is present above LQ but can not be quantified reliably and is reported qualitatively; trace: molecule is detected but can not be quantified reliably (\*: low extraction efficiency);

**Table 1 - Table presenting amendment details for each microcosm setup used in this study.**

Microcosms are setup with three replicates for each condition. Due to the quantity of material required for chemical analysis one out of three microcosm for each condition is sacrificed for chemical quantification at each time (and is indicated by a crossmark). Quantification results are obtained by Eurofins, Maxéville for each antibiotic molecule using LC/MS-MS.

microcosms amended with antibiotic molecules (4ml final) and a 4g in 180g final (2% w/w) amendment used for farm effluents microcosms conditions. All antibiotic molecules amended microcosms are prepared in two conditions: high-concentration (HC) or low-concentration of antibiotics (LC = reported concentration values in manure containing soils). High concentrations (HC) were set for each microcosm to mimic a 2X concentration difference with 4g LC amendment concentrations for the same molecule. Effluent-amended microcosms are prepared at two concentrations for chicken manure (2% or 40% w/w; respectively: A2T1-2 and A2T1-40); pig liquid manure and pig feces (2.5% w/w for both) and cow manure (2% w/w); Table 1. For each microcosm condition reported in Table 1, soil is mixed with amendments either in the form of farm effluents or 4ml of pure antibiotic-molecule solutions diluted in water at the desired concentration (Table 1) and thoroughly mixed with soil for a final volume of 180g for each microcosm condition. Each microcosm condition is prepared in batch with the respective amendment, using 800g of soil for 4x180g final weight microcosms. After a thorough mixing between amendment (antibiotic molecule in 4ml H<sub>2</sub>O or manure in required proportion, Table 1), 180g of soil is dispatched per incubation bottles corresponding to each sampling time point for chemical analysis or metagenomic DNA extraction (T0, T1, T3, T6).

#### **Antibiotic molecule quantification in soil and manure samples.**

Antibiotic molecules are extracted from farm effluents and soil samples (dried and grinded at 250 $\mu$ M) after a solvent extraction (methanol/EDTA/citric acid) and under agitation. After concentration by evaporation, samples are injected in a liquid-chromatography coupled with tandem mass-spectrometry (LC/MS-MS) to detect and quantify each antibiotic molecule in each sample. Our partner at Eurofins Environmentals Expertises-Maxéville, Dr. Emile Gardeur-Algros performed method developments and

quantification of soil and manure (Table 2); and microcosms incubated for T0, T1 and T6 months; upon provision of microcosm 150g of soil for each analysis. Results are reported in Table 1 as quantification in  $\mu\text{g}$  of each antibiotic-molecule per kg of soil dry matter ( $\mu\text{g}/\text{kg}$  of DM). When results are reliable only when above machine quantification limits thresholds defined using previous results on antibiotic-molecule solvent extraction efficiency and quantification calibration range previously obtained on Rothamsted soil for each molecule (Table 2). Else, results are expressed qualitatively as follow: "+", "++", or "+++" denote presence above  $6\mu\text{g}/\text{kg}$  but unreliable absolute quantification; "absence" means  $<2\mu\text{g}/\text{kg}$  of DM (hardware detection level) and "<QL": detection below quantification limits but above  $2\mu\text{g}/\text{kg}$  of dry matter). Antibiotics are quantified only at T0, T1 and T6 time-point.

#### **Metagenomic DNA extraction from soil microcosms.**

Soil is sampled (20g) from each corresponding microcosms incubation bottles at different time-points: T0, T1, T3 and T6 months (Table 1). Metagenomic DNA from three independent extraction using 500mg of soil from each microcosm samples are obtained with a modified protocol from Griffiths et al. [42] (j=228 independent metagenomic DNA extractions). Briefly: 500mg of soil is introduced in a tube containing a lysis matrix composed of glass and silica beads of varying sizes (Lysis Matrix E, MP Biomedicals) with 500 $\mu\text{L}$  of CTAB (hexadecyltrimethylammonium bromide; mild-detergent) and 500 $\mu\text{L}$  of phenol-chloroform and vigorously (bead-beating lysis). DNA is further purified by cold ethanol-NaCl precipitation and silica-based column purification (Illustra PCR & Band Gel Purification kit, GE Healthcare) following standard molecular biology procedures. DNA concentration is estimated using a fluorescent based assay (QuBit BR dsDNA assay, Invitrogen) and diluted at a final concentration of 10ng/ $\mu\text{L}$  in ultrapure water.

Antibiotic molecule	Relative quantification bias (%)					Calibration model	Valid calibration range (µg/L)
	0.5 µg/L	10 µg/L	20 µg/L	50 µg/L	200 µg/L		
Sulfamethoxazol	-4.8	2	8.1	-7.04	-7.3	Linear	0.5 to 100
Sulfaméthazine	-5.2	14.2	7.1	-4.96	17.5	Linear	0.5 to 200
Trimethoprim	-1.08	10.74	4.4	-4.2	-8.68	Linear	0.5 to 200
Ampicillin	2.4	-11.14	-0.88	3.3	-0.92	Quadratic	10 to 500
Enrofloxacin	2.4	-7.5	-8.9	1.76	-3.5	Quadratic	0.5 to 500
Tetracycline	1.4	9.8	-2.76	-1.58	1.4	Quadratic	0.5 to 200
Tylosin	3.2	-5.12	-3.4	5.7	-2.12	Quadratic	0.5 to 500

Table 2A

Antibiotic molecule	Relative extraction efficiency measure in soil		
	Rothamsted	Liquid manure	Manure + straw
Sulfamethoxazol	97%	122%	112%
Sulfaméthazine	111%	96%	69%
Triméthoprime	82%	85%	81%
Ampicilline	79%	100%	107%
Enrofloxacine	71%	101%	81%
Tétracycline	117%	94%	99%
Tylosine	64%	63%	81%

Table 2B

Antibiotic molecule	Quantification limits in different environmental matrices		
	Soil + Straw manure/Liquid manure	Straw manure	Liquid manure
(units)	µg/kg dry mat.	µg/kg dry mat.	µg/L
Triméthoprime	1	6	6
Sulfaméthazine	1	6	6
Sulfaméthoxazole	1	6	6
Tétracycline	10	60	60
Tylosine	5	30	30
Enrofloxacine	20	120	120
Ampicilline	5	30	30

Table 2C

**Table 2 - Calibration results details for antibiotic molecules quantification methods in soil.**

Tables giving details on the results of calibration experiments for robust chemical-quantification of antibiotic molecules methods in soil and environmental effluents used as amendment in this study: **A) Calibration range** for each antibiotic-molecule; **B) Solvent-extraction efficiency** in the different environmental matrices determined using 480µg of each molecule per kg of dry matter; **C) Quantification limits** for each molecule considering solvent-extraction efficiency in different matrices: straw manure; liquid manure and soil fertilized with manure.

	Sample Type	Sulfamethoxazole	Ampicillin	Trimethoprim	Enrofloxacin	Tetracycline	Tylosine
		µg/kg dry matter					
A2T1	Chicken feces	<LQ	<LQ	16.5	810	<LQ	<LQ
PT2	Pigs liquid manure	<LQ	<LQ	<LQ	<LQ	++	<LQ
PNT	Pigs feces	<LQ	<LQ	<LQ	<LQ	trace	<LQ
B1T1	Cow feces	<LQ	<LQ	<LQ	<LQ	15192	<LQ

**Table 3 - Quantification results in farm effluents.**

Antibiotic molecule quantification in farm effluents used in this study. Results are expressed either quantitatively in µg/kg of sample dry matter or qualitatively when results are not reliable enough for absolute quantification: <LQ, quantification limits for each molecule. (+) or (++) : concentration above quantification limits but extraction-efficiency in sample is not satisfactory.

### **Microbial taxonomic community fingerprinting.**

Community fingerprinting is assessed by ribosomal 16S-23S intergenic region spacer analysis (RISA) for each soil metagenomic DNA using primers S-D-Bact-1522-b-S-20 (eubacterial rRNA small subunit, 5'-TGCGGCTGGATCCCCTCCTT- 3') and L-D-Bact-132-a-A-18 (eubacterial rRNA large subunit, 5'- CCGGGTTTCCCCATTCGG-3') [43, 44]. PCR products are analyzed by capillary electrophoresis (BioAnalyzer 2100, Agilent) and electropherogram results are extracted as fluorescence intensity (increased PCR products concentration) measured each 0.05s retention time step (product size), in all 76 DNA samples microcosm/time couples. The resulting table consists then in a matrix of variables spanning PCR products time ranging from +70s to +110s retention-time (PCR products size: 150-1000 base-pairs equivalent) discretized in 840 time-classes where fluorescence intensity is recorded for each sample. Data are analyzed by principal component analysis (PCA) of the 840 time-classes variables (Figure 3 & 4) and between-groups analysis of the aforementioned PCA analysis (Figure 6 & 7).

### **Antibiotic-resistance genes quantification in soil microcosms.**

Soil microcosms are incubated at room temperature in 250 ml sterilized glass bottles. A 20g soil sample is collected from microcosm bottles at T0 (incubation start), T1 (one month incubation), T3 (3 months) and T6 (6 months), for metagenomic DNA extractions. Samples are sent for antibiotic molecule quantification at times T0, T3 and T6. A selection of 10 antibiotic-resistance (*aadE*, *aphA3*: aminoglycoside resistance; *bla<sub>TEM-1</sub>*, *bla<sub>SHV</sub>*: extended beta-lactams resistance; *cfiA*: metallo-beta-lactamase; *ermA*, *ermB*: erythromycin resistance; *sul1*: sulfonamides resistance; *tetA(G)*, *tetA(H)*: tetracycline-specific efflux pump genes and 2 mobile genetic elements markers (*int1*: class 1 integrons; *trfA*-IncP: broad-host plasmid IncP) quantitative PCR assays are used in this study and were previously

used in environmental surveys of antibiotic resistance gene abundance (Table 4). A total number of  $n=4332$  qPCR reactions ( $j=228$  independent metagenomic DNA extraction = [3 independent extractions for 76 microcosm/time couples];  $i=19$  quantitative PCR assays;  $n=i \cdot j = 4332$  independent quantitative PCR reactions excluding calibration curves for each assay). Relative quantification between samples is normalized using qPCR results obtained with primers 16S rRNA gene universal primers for each 76 DNA tubes. Calibration curves for 16S rRNA gene quantification are performed using a known copy-numbers serial dilution of *Sphingobium francense* type-strain Sp+ 16S rRNA gene, a soil borne Proteobacteria [NCBI TaxId: 332057]. Reactions are performed in a final volume of 20  $\mu\text{L}$  using 10  $\mu\text{L}$  SensiFast SYBR No-Rox (Bioline) reaction buffer containing Taq polymerase, 0.8  $\mu\text{L}$  of each primer at 100 nM and 2  $\mu\text{L}$  of template DNA at  $10 \text{ ng} \cdot \mu\text{L}^{-1}$  and run with the following program in a qPCR-thermocycler RotorGen 3000 (Quiagen, Courtaboeuf, France): 10 mins at 95°C and 40 cycles of 30secs at 95°C; 30secs at specific  $T_a$  (see Table S1); 30 secs at 72°C. A final melting analysis is performed recording fluorescence while temperature ramps from 50°C to 99°C at 1°C per second to ensure product specificity. Serial dilutions of  $10^8$  to  $10^2$  copies of positive-control DNA template for each gene are used to serve as calibration curve for determining cycle threshold and assess efficiencies for each reaction. Only reactions with a calibration curve goodness-of-fit  $R^2$ -value above 0.99 are considered valid for quantification. For each reaction, reaction tube CT is measured using Corbett Research RotorGene software version 2.1.0 (Build 9) built-in threshold-calibration tool. Relative quantification between samples is possible after normalizing quantification data with an internal calibrator (16S rRNA gene). Results are expressed either relative to each sample quantification at T0 (Figure 3&6) or relative to quantification results in microcosm control condition T-H<sub>2</sub>O (water added) at the same time-point (Figure 5&7) using equation 1. Raw quantification results table, including

specific cycle-threshold extraction, Ct differences with copy number ratios in reference to samples T0 or Control H<sub>2</sub>O for each target in each DNA sample are available as online supplementary information for convenience (Figure S1 - <http://bit.ly/1oA30SB>).

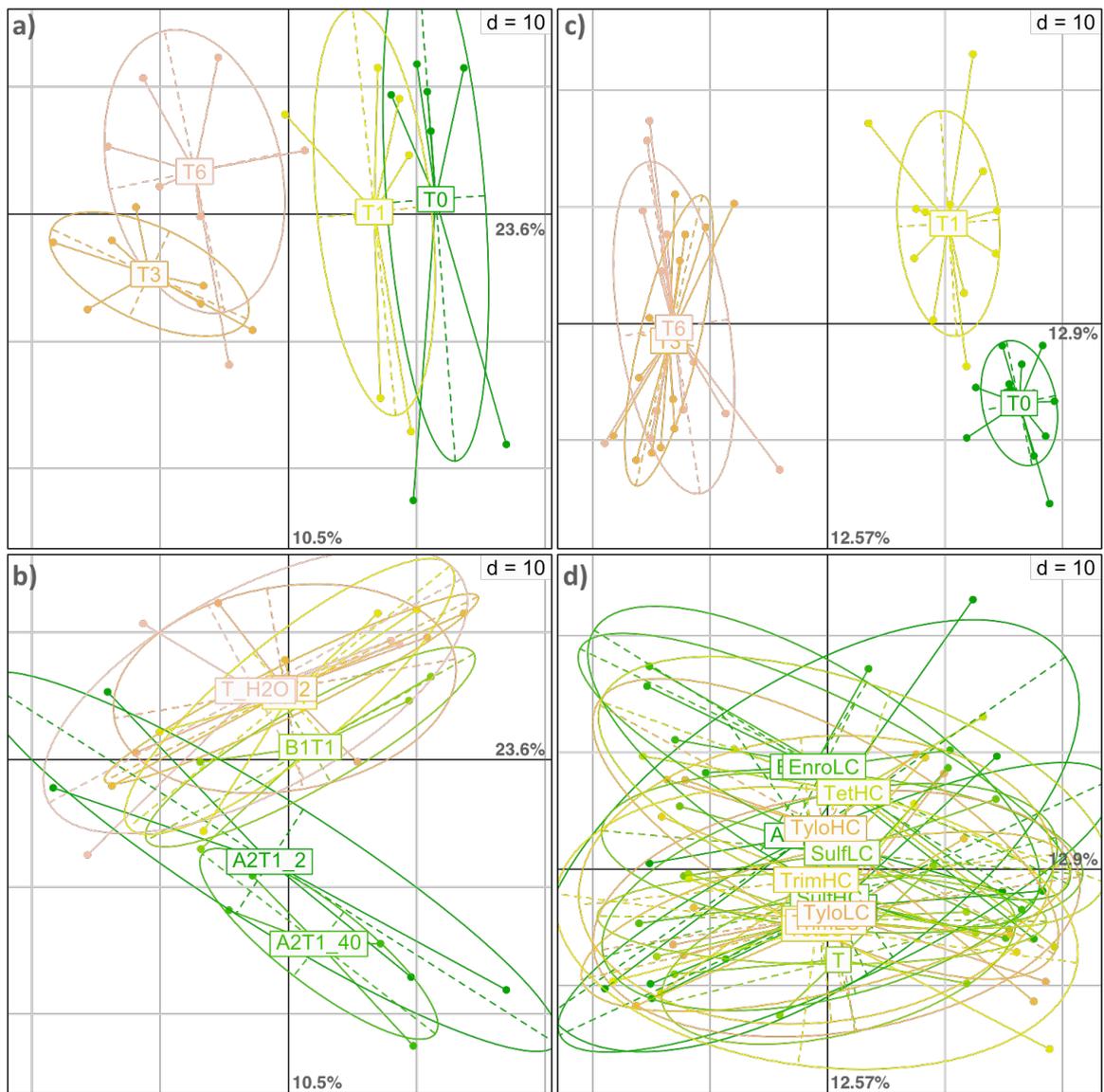
## Results.

### **Abundance of antibiotics molecules in soil microcosm samples measured by LC/MS-MS.**

Measuring small concentrations of complex molecules such as antibiotics in matrices such as soil or animal manures is a real technical challenge, even though a very careful attention is drawn on methods development as in this study. Antibiotic-molecules quantification was achieved by an external third-party company specialized in environmental contaminant detection: Eurofins Environmentals Expertises-Maxéville. Specific antibiotic molecules extraction efficiencies, calibration ranges and quantification limits are established in the Rothamsted soil and manure similar to those used in this study: straw manure and liquid manure (Table 2) and extraction efficiency was always verified for each quantification reaction with standards of known concentration. When quantification results are flawed by molecule extraction efficiency or concentrations below quantification limits, but above detection level, results are given qualitatively. It was not possible to quantify any of the tested antibiotic molecules in Rothamsted Park Grass soil prior to microcosm amendment. In farm effluents the following antibiotics were detected: trimethoprim was detected in chicken feces A2T1 (16.5µg/kg); tetracycline was detected in treated-pig liquid manure PT2 (++), but also in non-treated pig feces PNT, at a significantly lower concentration (+) and in cow feces B1T1 in relatively high concentration (15192µg/kg). In microcosms amended with pure antibiotics molecules, sulfamethoxazole,

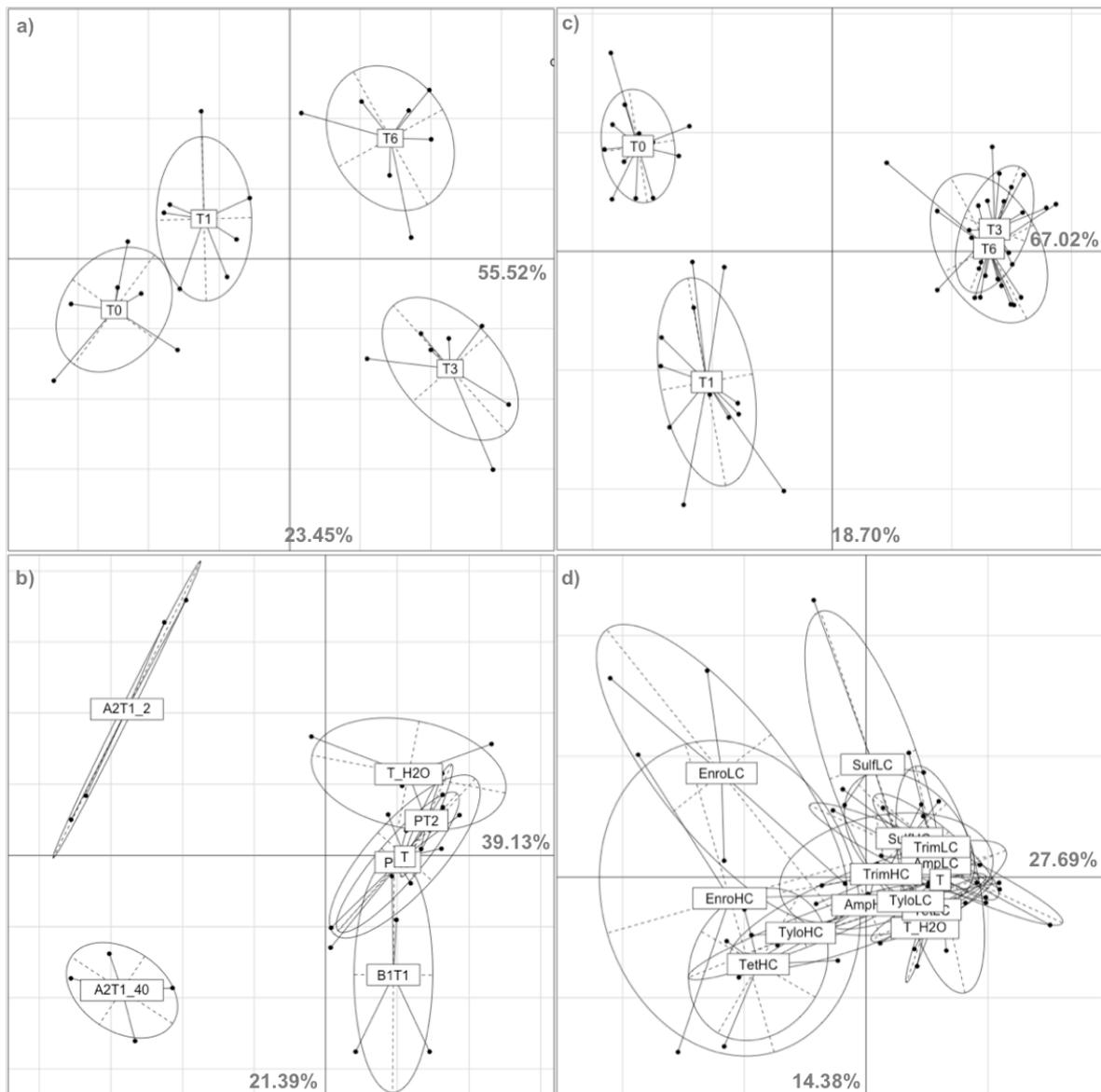
trimethoprim, tylosin and enrofloxacin are quantified in their respective microcosm amendment conditions. Quantification results are compared with concentration expected by theory (i.e.: concentration measured for each amendment condition at T0; Table 1).

Sulfamethoxazole is quantified in conditions sulfamethoxazole HC and LC at T0 (HC T0=41µg/kg; theory: 111µg/kg and LC T0=traces; theory: 22µg/kg) but only trace amount is found at T1 and T6 in both microcosms conditions and are below quantification limits after 6 months, indicating a decay of the molecule overtime in soil. Similarly, trimethoprim is quantified at 278µg/kg and 25µg/kg respectively in conditions Trimethoprim HC and LC at T0, theoretical concentrations: 220µg/kg and 22µg/kg. After one month: 122µg/kg and 15µg/kg are found in conditions HC and LC and after 6 months (T6), trimethoprim is detected above quantification limits (>6µg/kg) in condition HC and only in trace amount in condition LC. Tylosin is quantified at 368 µg/kg in sample HC (theory: 555 µg/kg) at T0, 306µg/kg at T1 and 23.4 µg/kg after 6 months (T6). In sample LC, tylosin is quantified at 70µg/kg at T0 (theory: 110 µg/kg) at T0, in traces amounts at T1 and is below quantification limits at T6. In microcosms amended with enrofloxacin, the molecule is detected in both microcosms during the whole incubation period (T1 to T6), but results are showing a number of interferences (molecule instability) and absolute quantification cannot be assessed. No antibiotics can be detected in microcosms amended with tetracycline or ampicillin. In microcosms amended with farm effluents, quantifications results are more difficult to interpret since no antibiotics can be quantified in any microcosm conditions above quantification limits, despite high concentrations found in effluents themselves (Table 3). However, enrofloxacin is detected in microcosms amended with 40%(w/w) chicken feces (A2T1-40) and were previously detected in the corresponding effluent (A2T1: 16.5 µg/kg trimethoprim and 810 µg/kg tetracycline).



**Figure 1 - Principal component analysis of RISA-PCR fingerprinting results obtained for each microcosm condition and time.**

PCA plots are calculated using results obtained by micro-capillary electrophoresis of RISA-PCR products from each microcosm and time point (T0, T1, T3 and T6 months). Drastic changes in the community profile are expected in microcosms amended with farm effluents. To minimize confounding effects of the effluents input, microcosms are analyzed separately: amended with effluents (panels **a** and **b**) and amended with antibiotic molecules (panels **c** and **d**). To evaluate the most significant factor (Time or Treatment), influencing samples clustering displayed on the ACP, samples are colored by respective factors in each panel: time in months in panels **a** and **c** and microcosm treatment in panels **b** and **d**.



**Figure 2 - PCA-plots between groups analysis of RISA-PCR taxonomic fingerprinting profiles for each microcosm condition at T0, T1, T3 and T6 months.**

BGA plots calculated using principal-component analysis of microcosms: amended with effluents (panels **a** and **b**) and amended with antibiotic molecules (panels **c** and **d**). PCA plots from Figure 2 are constrained in  $k$  groups defined by each sample factor: sampling time (panels **a** and **c**) and microcosm treatments (**b** and **d**). Monte-Carlo test based on 1000 random permutations and estimated  $p$ -values for  $k$ -groups in as displayed in plots **a**, **b**, **c** and **d** are respectively: **a**:  $k=4$ ;  $p=0.000999001$ ; **b**:  $k=7$ ;  $p=0.05894106$ ; **c**:  $k=4$ ;  $p=0.000999001$ ; **d**:  $k=14$ ;  $p=0.99900$ ; alternative hypothesis: number of groups is greater than  $k$ ).

### **Amendments impacts on soil prokaryotes community profiles assessed by RISA-PCR.**

RISA profiles obtained by micro-capillary electrophoresis (BioAnalyzer 2100, Agilent) are compared by principal component analysis to reduce the variable table dimensions from  $k=840$  to  $k=2$  dimensions (principal components) that can therefore be displayed in a two-dimensions scatter plot (Figure 1). Microcosms amended with farm effluents are analyzed separately (Figure 1; panels a & b) from those amended with pure antibiotic-molecules (Figure 1; panels c & d). The two microcosm control conditions (T & T H<sub>2</sub>O) are included in both analyses. Samples are either grouped by time-point (a & c) or microcosm amendment condition (b & d) using the same data-points. Groups formed by the microcosm incubation time seem more coherent than groups formed by microcosms condition factor; indicating a major effect of soil incubation in glass bottles microcosms independently of amendment condition. This observation is also true when viewing plots along principal component axis 2 and 3 (data not shown). In order to test significance of these grouping factors a between group analysis is performed. The previously calculated PCA plot is constrained by groups based either on sampling-time or amendment condition (Figure 2). Between groups analysis shows that soil sample taxonomic communities are mainly impacted by soil incubation time rather than microcosm amendment conditions. Samples amended with pure antibiotic-molecules are clearly not well separated for most conditions in Figure 5d; with the exception of conditions: Tylo HC, Tet HC and both Enrofloxacin concentrations (Enro HC, Enro LC), where enrofloxacin is detected by chemical-analysis (Table 1). When considering incubation-time as grouping factor, samples at after a one-month incubation are clustered separately from T<sub>0</sub>, emphasizing on the sole microcosm incubation effect. After three and six months incubation, all samples are grouped together, indicating similar microbial communities and resilience potential of soil after three months post-perturbation (i.e.: soil

microcosm incubation; BGA p-value <0.001). In contrast, soil microcosms amended with farm effluents communities, after one, three and six month's incubation are clearly different (BGA p-value <0.001) and do not suggest resilience of the community. Farm effluent conditions are considerably modifying the indigenous soil community and strong differences between microcosm community fingerprinting amended with chicken feces (A2T1-2 and A2T1-40; Figure 1). These differences are significant when considering BGA plots: samples from microcosm conditions A2T1-2 and A2T1-40 form two distinct groups, away from all other samples. While less evident, the combined effect of time and amendment makes samples from B1T1 microcosms separating from other samples after 3 and 6 months incubation and therefore do not suggest soil resilience after 3 months, as in microcosms amended only with antibiotic molecules (Figure 2b; BGA p-value<0.06).

#### **Antibiotic-resistance gene abundance in soil microcosms measured by qPCR.**

Antibiotic resistance gene quantifications are measured in metagenomic DNA samples extracted from soil microcosms by quantitative PCR using a selection of primers set targeting: *aadE*, *aphA-3* (aminoglycosides resistances) *bla<sub>SHV</sub>*, *bla<sub>TEM</sub>* (extended-spectrum beta-lactamase); *cfia* (metalo-beta-lactamase); *ermA*, *ermB* (macrolides resistances); *sul1* (sulfonamides resistance) and *tetA(G)*, *tetA(H)* (tetracycline resistance). Quantitative PCR targeting two mobile genetic element markers are also quantified in all soil samples (*int1* for integrons and *trfA*-IncP for broad-host plasmid IncP group). Quantification for each qPCR assay targeting antibiotic resistance-genes in each microcosm condition results are expressed either relative to T0 for the same condition or relative to the same time-point (T0, T1, T3 or T6) in control H<sub>2</sub>O microcosm condition.

Target gene	Gene function	Resistance to	qPCR efficiency	Calibration curve R2	Accession Number	Ref.
16S rRNA	Ribosomal RNA	<i>n.a.</i>	0.91	0.9918	NR_042944.1	Takai et al. (2000)
aadE	Streptomycin adenylyltransferase	Streptomycin, spectinomycin	0.93	0.9956	X02340	Delgado et al. (2011)
aphA-3	Aminoglycoside 3'-phosphotransferase	Kanamycin	0.92	0.9975	V01547	Delgado et al. (2011)
blaSHV	SHV-11 type extended spectrum $\beta$ -lactamase	Penicillins, cephalosporins	0.86	0.9958	KF585137	Hammond et al. (2005)
blaTEM	TEM-1 type extended spectrum $\beta$ -lactamase	Penicillins, cephalosporins	0.95	0.9985	AY425988.1	Demanèche et al. (2008)
cfiA	cfiA metallo- $\beta$ -lactamase	Penicillins, cephalosporins, cephamycin, imipenem	1.08	0.9969	JF795577.1	Pumbwe et al. (2007)
ermA	rRNA adenine N6-methyltransferase 1	Macrolides, lincosamides, type B streptogramins	1.01	0.9957	NC_002952.2	Merino-Diaz et al. (2007)
ermB	rRNA adenine N6-methyltransferases	Macrolides, lincosamides, type B streptogramins	0.82	0.9936	AB539726	Chen et al. (2007)
sulI	Sulfonamid-resistant dihydropteroate synthase	Sulfonamides	1.02	0.9999	AM261282	Heuer and Smalla (2007)
tetA(G)	Class G TetA tetracycline-specific efflux pump	Tetracyclines	0.95	0.9968	S52437	Fan et al. (2007)
tetA(H)	Class H TetA tetracycline-specific efflux pump	Tetracyclines	0.93	0.9997	U00792	Fan et al. (2007)
int1	Class 1 integrons tyrosine recombinase	<i>n.a.</i>	0.92	0.9997	AM261282	Hoyle et al. (2006)
trfA	Broad-host plasmid IncP trfA replication initiation protein	<i>n.a.</i>	1	0.9911	AM261282	Bahl et al. (2009)

**Table 4 – Characteristics of antibiotic resistance and mobile genetic element marker genes used in this study and respective quantitative PCR reactions efficiency and calibration-curve goodness-of-fit.**

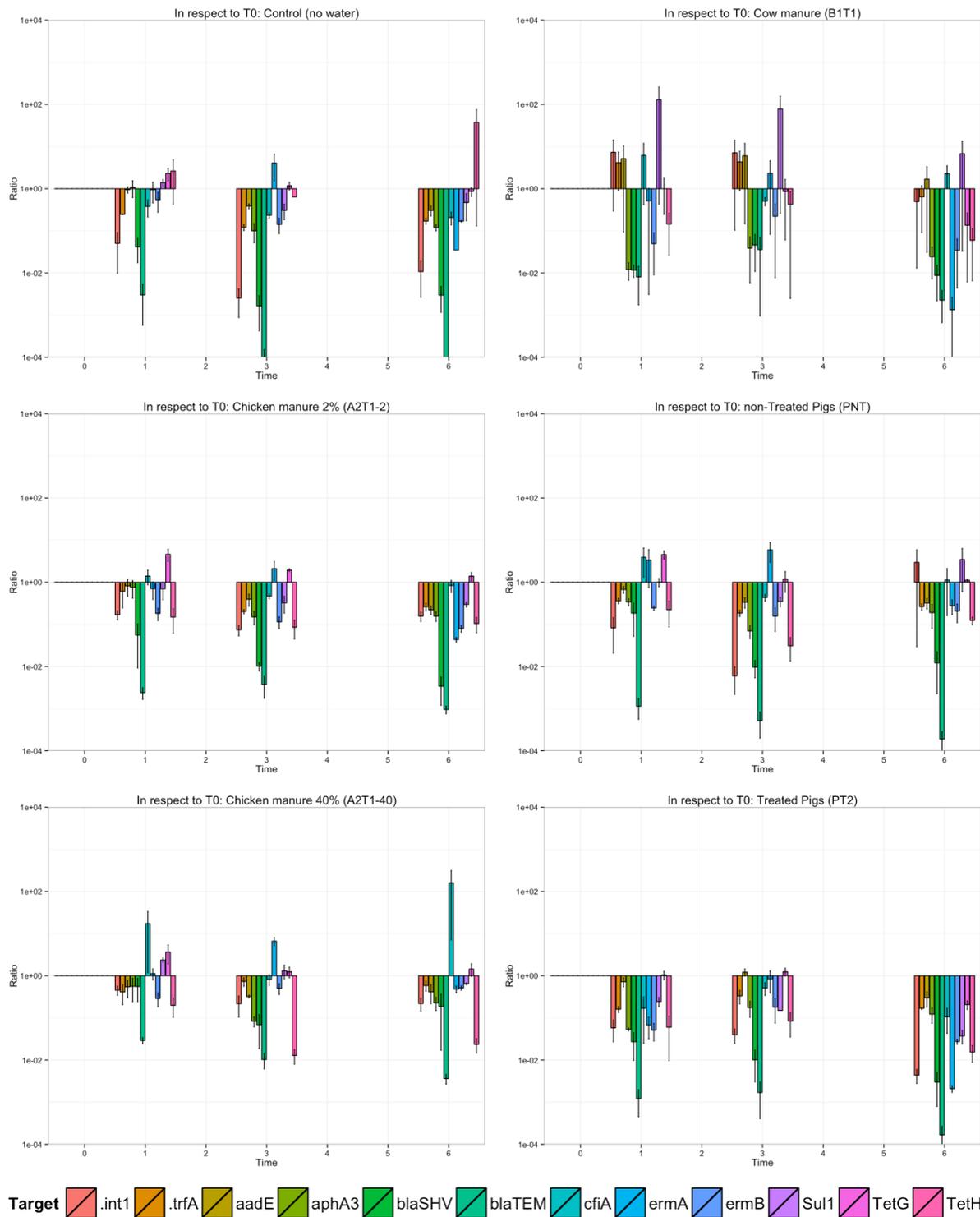
Accession Number indicate the respective sequence template used as positive control for each reaction on NCBI sequence database.

$$\text{ratio} = \frac{(E_{\text{target}})^{\Delta C_p \text{ target (control - sample)}}}{(E_{\text{Ref}})^{\Delta C_p \text{ Ref (control - sample)}}$$

**Equation 1 - Ratio formula.**

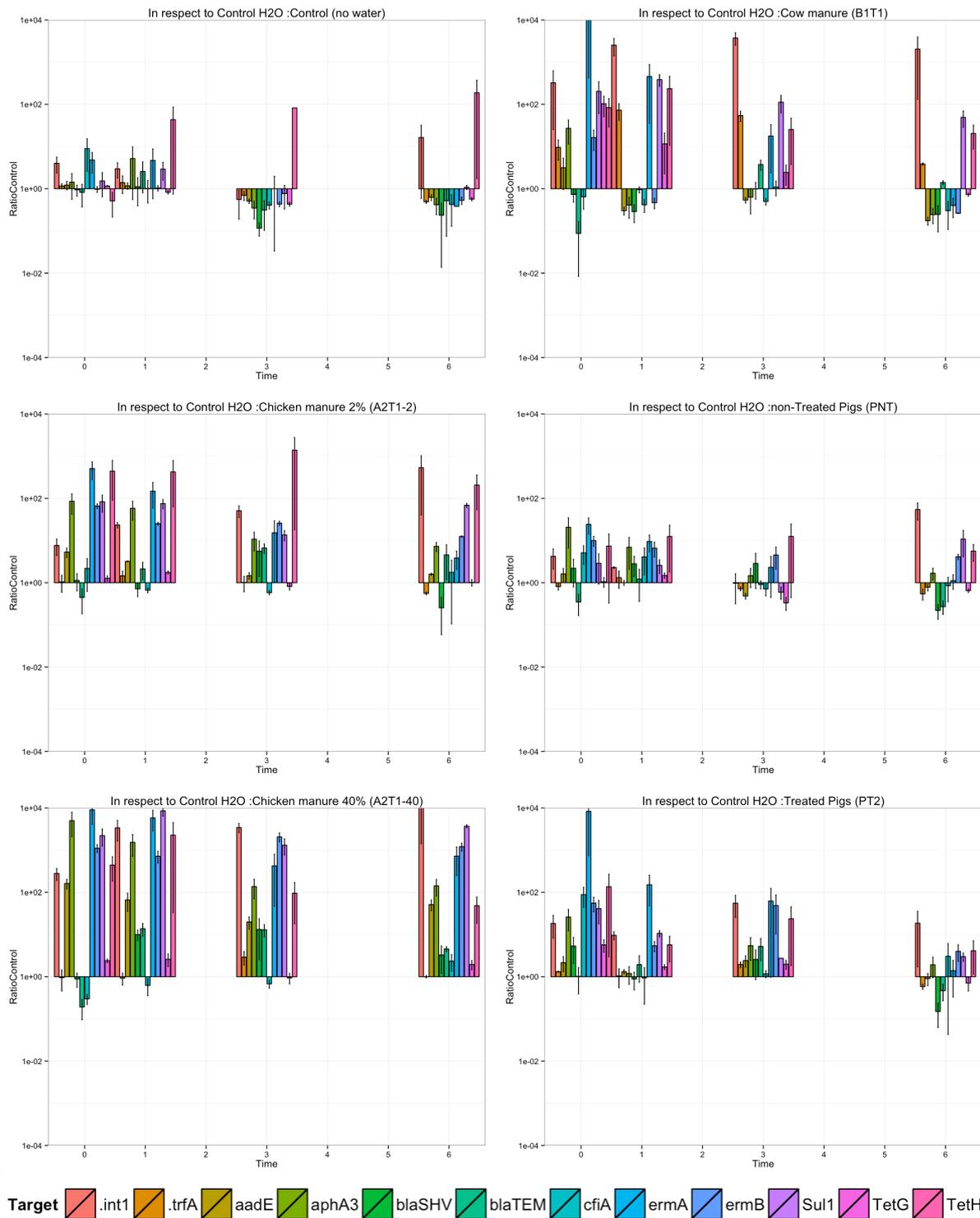
Formula used to calculate a gene quantification ratio (target ratio) between two samples using a reference gene (16S rRNA gene) as sample calibrator, in reference to a control sample (Rothamsted, Park Grass soil). Pfaffl, M. W. Et al. (2002). Nucleic Acids Res 30, e36

This double-presentation displays both dynamics over time of antibiotic resistance copy (relative to each sample T0; Figures 3&5) and different impacts of each amendment in respect to control condition (relative to control H<sub>2</sub>O at the same incubation time; Figure 4&6). Relative quantification ratios are plotted on a logarithmic y-axis an increase (ratio < 1; upper bars) is therefore displayed as upper-bars and decrease (ratio > 1) as lower bars) in measured gene-copy, relative to reference sample (relative T0 or control H<sub>2</sub>O) for each qPCR target. All reactions are standardized using 16S rRNA gene copy number to average differences in prokaryotic genomes number between samples. When considering soil amended with farm effluents (Figure 3&5), the most significant important changes in ARGD and MGE copy abundance over time and in respect to Control condition (no water; upper-left panel), concern soil microcosms amended with 40% (w/w) chicken-feces and cow manure (B1T1). In sample A2T1-40, a significant increase in *ermA* at T3 and *bla*<sub>TEM</sub> gene copies at T6 is observed together with a decrease overtime in *bla*<sub>SHV</sub> and *tetA*(H) genes relative to T0 abundance (Figure 3). This dynamic is however not constant: *bla*<sub>TEM</sub> is detected in increased abundance at T1 but not at T3 and the same trend is not observed in microcosms amended with only 2% (w/w) chicken feces. In sample B1T1, genes mobile-genetic elements markers *int1* and *trfA*-IncP are increased between T0 and T1 and drop at T6. This dynamic is also found with *sul1* copies-ratio dynamic but high variations in quantification results are observed in samples from this microcosm condition and results are not significant enough statistically wise. Effects in antibiotic resistance gene relative abundance dynamic over time can be deduced from results obtained for soil microcosm control condition with no-added water (i.e.: raw soil incubation). Most measured gene relative abundance to T0 in no-water control decrease between T0 and T3 months and stabilize at T6; with the exception of *ermA*, being more abundant at T3 than in T0.



**Figure 3 - Barplots showing quantification for each qPCR assay relative to gene copy quantification at T0 for each effluent-amended microcosm.**

Quantitative PCR results are expressed as quantification ratio relative to each sample quantification at T0, using the number of 16S rRNA gene copy as internal sample calibrator. Barplots are ordered on the x-axis in respect to each incubation time, in months. Quantification assays were performed on soil samples metagenomic DNA extracted from three individual soil samples of each microcosm at each sample time. Variations in quantification results are expressed as error-bars (standard-error for each sample mean quantification ratio).



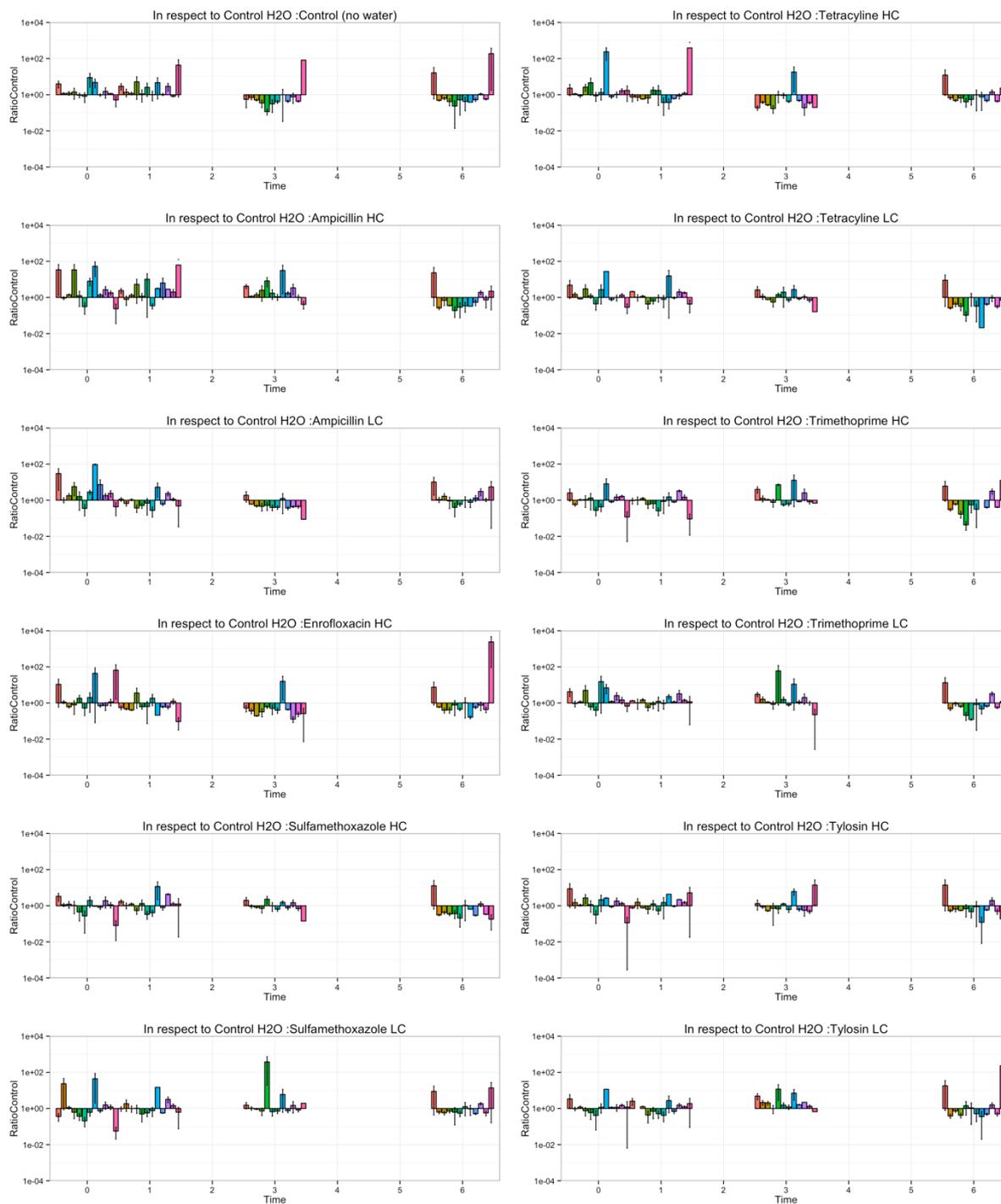
**Figure 4 - Barplots showing quantification for each qPCR assay relative to gene copy quantification in water-amended control for each effluent amended microcosm.**

Quantitative PCR results are expressed as quantification ratio relative to each sample quantification at T0, using the number of 16S rRNA gene copy as internal sample calibrator. Barplots are ordered on the x-axis in respect to each incubation time, in months. Quantification assays were performed on soil sample metagenomic DNA extracted from three individual soil samples of each microcosm at each sample time. Variations in quantification results are expressed as error-bars (standard-error for each sample mean quantification ratio).



**Figure 5 - Barplots showing quantification for each qPCR assay relative to gene copy quantification at T0 for each antibiotic molecule amended microcosms.**

Quantitative PCR results are expressed as quantification ratio relative to each target quantification in water-amended control microcosm, using the number of 16S rRNA gene copy as internal sample calibrator. Barplots are ordered on the x-axis in respect to each incubation time, in months. Quantification assays were performed on soil samples metagenomic DNA extracted from three individual soil samples of each microcosm at each sample time. Variations in quantification results are expressed as error-bars (standard-error for each sample mean quantification ratio).



**Figure 6 - Barplots showing quantification for each qPCR assay relative to gene copy quantification in water-amended control for each antibiotic molecule amended microcosms.**

Quantitative PCR results are expressed as quantification ratio relative to each target quantification in water-amended control microcosm, using the number of 16S rRNA gene copy as internal sample calibrator. Barplots are ordered on the x-axis in respect to each incubation time, in months. Quantification assays were performed on soil sample metagenomic DNA extracted from three individual soil samples of each microcosm at each sample time. Variations in quantification results are expressed as error-bars (standard-error for each sample mean quantification ratio).

When comparing samples with quantification ratio relative to control H<sub>2</sub>O, microcosms amended with farm effluents (Figure 4) contains significantly more ARGD than in control at T0 with the notable exception of *bla*<sub>TEM</sub> and *bla*<sub>SHV</sub> found at the same ratio. Similarly, the relative number of mobile genetic elements markers is significantly higher in microcosms A2T1-2 and A2T1-40 (chicken feces at 2% and 40% w/w), with a higher increase in A2T1-40. In contrast, genes *aadE* and *aphA-3* are increased relative to control H<sub>2</sub>O at T0 but decreased in their relative abundance over-time from T1 to T6 in both A2T1-40 and A2T1-2 conditions. A similar trend can be observed for genes *aphA-3*, *ermA*, *ermB*, *sul1* in condition PT2 and with less changes but a similar trend in condition PNT, for the same genes. In both microcosms amended with chicken feces, the relative number of *trfA* gene is increased over-time with stronger changes in sample A2T1-40. In condition B1T1 amended with cow feces treated with antibiotics, mobile genetic elements are also increased in T0 and T1, and T3. At 6 months incubation, *trfA* abundance is similar to T3 but *int1* abundance is reduced, while increased relative to control H<sub>2</sub>O. In microcosms amended with pure antibiotic-molecules, changes in relative quantification between microcosm conditions are very limited when compared with quantifications in those amended with farm effluents. Due to technical variations in quantification, most differences observed are not significant. Still, *bla*<sub>SHV</sub> relative abundance in reference to control H<sub>2</sub>O (Figure 6) is higher in condition Ampicillin HC microcosm at T1 and T3 but is equivalent to control after 6 months incubation.

## Discussion.

Large quantities of pharmaceuticals including antibiotic molecules are dispersed in the environment by human-activities (*i.e.*: urban-wastewater systems, production-sites [28]). In addition, traditional agricultural practices such as use of animal manure for soil fertilization (*i.e.*: mix of animal litter straw with feces or liquid manure) are recognized as a major source of antibiotic molecules pollution in soils [25, 28, 45]. A number of studies were already conducted on the abundance of antibiotic-resistance genes, or antibiotic resistant bacteria, in various environments with history of antibiotic molecule pollution: wastewater samples [46], edible crops [47] or manure [48]. However, these studies are mainly trying to assess quantification of a specific gene, in respect to antibiotic pollution history of the sampling site. While providing invaluable data in the frame of environmental antibiotic-resistance knowledge, such studies are mostly observatory. Obviously, if an antibiotic-resistance gene is present in soil bacteria and this soil is impacted by large amounts of antibiotics (*i.e.*: quantified above inhibition concentration), these bacteria will be strongly selected in response. Sensitive methods such as quantitative PCR will most likely detect a correlation between the two. Similarly, if only a fraction of diversity is considered, for example when examining resistance profile of specific taxa such as *Enterobacteriaceae*, results can hardly be extrapolated to the rest of the soil community affected by the same perturbation. Proper risk assessments of an exogenous source of antibiotic molecules on dissemination of antibiotic resistance genes can be performed with simple and sensitive methods, such as quantitative PCR or RISA; it however requires focusing on a single soil to decipher the different factors driving observed changes in results as presented here. This soil has the advantage of being monitored since before the medical antibiotic era (1874;

penicillin discovery: 1928) and have therefore never been directly exposed to pharmaceutical antibiotic molecules but was shown in a previous shotgun metagenomic study to harbor a considerable diversity of antibiotic-resistance genes (Chapter 2, [39]). Environmental antibiotic pollutions are likely to favor the recruitment of environmental resistance genes in pathogens genomes or exchange of genetic material between soil and commensal bacteria [16, 49]. A fitness-gain from antibiotic-resistance genes is true only in pathogens or bacteria confronted to drastic antibiotic selection. The role and selection of antibiotic resistance genes in nature is however more obscure than in clinical environments. It is in this respect that microcosms prepared from a single soil source never polluted by antibiotics are setup for 6 months incubation period. Microcosm experiment is setup similarly as described previously by Jacquioud et al. studying how the Rothamsted soil bacterial community responded to chitin enrichment in a microcosm experiment [50]. Microcosms for a same condition are prepared in batch with 800g of soil, enough for 4x 180g (*i.e.*: T0, T1, T3 and T6) but are incubated in different bottles for each corresponding sampling time, due to the amounts required for chemical analysis (150g); an amount that can not be sampled without considerably altering the soil structure. The soil community is therefore physically different in microcosm bottles and may have developed differently in different bottles. This is a convenient explanation for some weird results, such as variations in qPCR quantifications for *bla<sub>TEM</sub>* in sample A2T1-2: increased abundance at T1, not different from T0 at 3 months incubation and again increased abundance at T6 and was also observed in a previous microcosm lab study where taxa-specific 16S rRNA gene abundance variation in chitin-amended microcosms was compared to infer quantitative changes in microbial community following chitin enrichment [50]. Such variations are not seen at the whole community level, indicating that incubation in different bottles is not sufficient to

explain observed variations between microcosm conditions and comforts our assumption that these samples can be compared over-time, as presented in quantitative PCR results, providing insights on the dynamics of ARGD in response to antibiotic perturbation. Another problem in environmental survey concerns comparison of samples from different origins or nature. Metagenomic DNA extractions from manure, sediments or soil of different composition will not be affected by the same extraction biases and those can considerably change results interpretations and/or significance [51, 52]. All soil microcosm samples were incubated for 0, 1, 3, and 6 months (respectively: T0, T1, T3 & T6). Samples were analyzed by RISA-PCR (taxonomic profiling) and quantitative PCR (antibiotic-resistance concentration). Antibiotics molecule were also quantified when possible: despite a careful method development, it is still a true challenge to assess absolute quantification of chemicals in complex environments such as soil with reliability. Indeed, in our soil samples, most antibiotic molecules were not detected at theoretical concentrations previously calculated (Table 1). However, when quantified or detected, at incubation start (T0), antibiotics are always found at lower concentration in the following months. With the exception of microcosm Trimethoprim HC where traces amounts are still found after 6 months, all other antibiotics are either absent or below quantification limits ( $>6\mu\text{g}/\text{kg}$  of soil dry matter) in all other microcosms. This can be explained by the different chemical nature of antibiotics molecules conferring distinct water solubility,  $\text{pK}_\text{A}$  (pH equilibrium-constant) or molar-mass. These differences are drastically affecting molecule behavior in soil, such as degradation potential (either biotic or abiotic) or quantification reliability, as shown in our results (Table 1). Antibiotics in manure samples can range between 450 to 900 $\mu\text{g}/\text{kg}$  for 199 $\mu\text{g}/\text{kg}$  tetracyclines, 13 to 67  $\mu\text{g}/\text{kg}$  for macrolides and 6 to 52 $\mu\text{g}/\text{kg}$  for fluoroquinolones<sup>1</sup>. Concentration of the same molecules in soils are however found in much lower quantities:

up to 199 $\mu\text{g}/\text{kg}$  tetracyclines, 11 $\mu\text{g}/\text{kg}$  sulfonamides and 7 $\mu\text{g}/\text{kg}$  chlortetracycline are found in soils after two year of conventional land-farming using animal-manure fertilization, for example [28]. These results are consistent with no antibiotics detectable in the Rothamsted soil that was never exposed to antibiotics and antibiotics do not accumulate in soil to reach the same concentrations as those found directly in animal feces, coherent with respective amended antibiotic molecules decay observed in microcosms when quantification was possible (Table 1) (*i.e.*: with tylosin sulfamethoxazole and trimethoprim in respective microcosm amendment condition). Changes in prokaryote taxonomic community were measured in each sample during incubation in microcosms with RISA-PCR and evolution of antibiotic resistance abundance by quantitative PCR (qPCR). Microcosm samples taxonomic community profiles were assessed by RISA-PCR (Figure 1&2). Microcosms amendments by pure antibiotic molecules are hardly separated and have very low effects on the soil community when analyzed RISA-PCR. This is not the case for incubation time that is strongly separating groups, based on the duration of incubation and community fingerprints at 3 and 6 months are clustered together, suggesting stabilization of the community. A similar incubation time effect is also observed in farm effluent microcosms but microbial communities are significantly different in these microcosms and these differences increased overtime. Such effects are very strong in microcosms amended with chicken and cow feces, probably due to the addition of large quantities of living microbes with each microcosm amendment condition rather than any antibiotic molecule selective effect. Abundances for a selection of antibiotic resistance genes were measured by quantitative PCR assays and similarly, most significant increases in relative ARGD quantification are due to the introduction of an exogenous bacterial community in soil by farm effluents. Also, the relative bacterial biomass in these microcosms will considerably increase with introduction

of farm effluents, when compared with control microcosms conditions or those amended with pure antibiotic molecules. These differences in prokaryotic biomass are averaged between samples using 16S rRNA gene copy number detected in each sample. As a consequence, differences in ARGD absolute gene copy number in these soils are orders of magnitude more important than reported in our results, only reflecting changes in relative abundance. Absolute gene-copy quantification in environmental samples is problematic considering aforementioned biases in metagenomic DNA extraction and sample heterogeneity and should always be considered very carefully, especially for comparisons between samples from different origins [53]. Results obtained with farm-effluents amended microcosms are however striking and confirm results previously observed in field studies following manure applications. In a similar quantitative PCR approach, Heuer and Smalla (2012) [40] reported that soil fertilization with sulfadiazine-containing manure increases significantly corresponding antibiotic-resistance gene *sul2* in the two months following manure application and that this increase is proportional to sulfadiazine concentration added in soils. More importantly, this increase in *sul2* gene quantification is well correlated ( $R^2 = 0.72$ ) with increase in *traN* gene copy, a gene specific to low G-C% plasmids that are common sulfonamide resistance genes vectors. Recently, Musovic *et al.* (2014) [54] retrieved microbial communities from soil and assessed their plasmid-uptake rate (permissiveness) by filter-mating conjugation and show in their results that plasmid-uptake rate is significantly higher for bacterial communities extracted from manure-fertilized soils. These studies, together with our qPCR quantification results are again reaffirming the duality of the antibiotic-resistance dissemination problem in the clinic: selection on beneficial traits (antibiotic-resistance phenotype) and acquisition-rate of such genes by pathogens. From our results and the growing literature on the soil resistome transferability

(i.e. the "resitome-mobilome") it becomes more and more clear that horizontal-gene transfer plays a major role in dissemination of antibiotic-resistance gene, selection of convenient expression vectors for these traits.

## References.

1. Waksman, S. A., and Woodruff, H. B. (1940). The Soil as a Source of Microorganisms Antagonistic to Disease-Producing Bacteria. *J Bacteriol* *40*, 581–600.
2. Bérdy, J. (2005). Bioactive microbial metabolites. *J. Antibiot.* *58*, 1–26.
3. Hibbing, M. E., Fuqua, C., Parsek, M. R., and Peterson, S. B. (2009). Bacterial competition: surviving and thriving in the microbial jungle. *Nat Rev Microbiol* *8*, 15–25.
4. Benveniste, R., and Davies, J. (1973). Aminoglycoside antibiotic-inactivating enzymes in actinomycetes similar to those present in clinical isolates of antibiotic-resistant bacteria. *Proc Natl Acad Sci USA* *70*, 2276–2280.
5. Perry, J. A., and Wright, G. D. (2013). The antibiotic resistance “mobilome”: searching for the link between environment and clinic. *Front Microbiol* *4*, 138.
6. Davies, J. (2011). How to discover new antibiotics: harvesting the parvome. *Curr Opin Chem Biol* *15*, 5–10.
7. Smith, T. (2000). Antibiotics from soil bacteria. *Nat Struct Biol* *7*, 189–190.
8. Laskaris, P., Tolba, S., Calvo-Bado, L., Wellington, E. M., and Wellington, L. (2010). Coevolution of antibiotic production and counter-resistance in soil bacteria. *Environ Microbiol* *12*, 783–796.
9. Demanèche, S., Sanguin, H., Poté, J., Navarro, E., Bernillon, D., Mavingui, P., Wildi, W., Vogel, T. M., and Simonet, P. (2008). Antibiotic-resistant soil bacteria in transgenic plant fields. *Proc Natl Acad Sci USA* *105*, 3957–3962.
10. Bhullar, K., Waglechner, N., Pawlowski, A., Koteva, K., Banks, E. D., Johnston, M. D., Barton, H. A., and Wright, G. D. (2012). Antibiotic resistance is prevalent in an isolated cave microbiome. *PLoS ONE* *7*, e34953.
11. Allen, H. K., Moe, L. A., Rodbumrer, J., Gaarder, A., and Handelsman, J. (2009). Functional metagenomics reveals diverse beta-lactamases in a remote Alaskan soil. *ISME J* *3*, 243–251.

12. Shade, A., Klimowicz, A. K., Spear, R. N., Linske, M., Donato, J. J., Hogan, C. S., McManus, P. S., and Handelsman, J. (2013). Streptomycin application has no detectable effect on bacterial community structure in apple orchard soil. *Appl Environ Microbiol*.
13. Jones, M. E., Draghi, D. C., Thornsberry, C., Karlowsky, J. A., Sahm, D. F., and Wenzel, R. P. (2004). Emerging resistance among bacterial pathogens in the intensive care unit--a European and North American Surveillance study (2000-2002). *Ann Clin Microbiol Antimicrob* 3, 14.
14. Forslund, K., Sunagawa, S., Roat Kultima, J., Mende, D., Arumugam, M., Typas, A., and Bork, P. (2013). Country-specific antibiotic use practices impact the human gut resistome. *Genome Res*.
15. Forslund, K., Sunagawa, S., Coelho, L. P., and Bork, P. (2014). Metagenomic insights into the human gut resistome and the forces that shape it. *Bioessays*.
16. Forsberg, K. J., Reyes, A., Wang, B., Selleck, E. M., Sommer, M. O. A., and Dantas, G. (2012). The Shared Antibiotic Resistome of Soil Bacteria and Human Pathogens. *Science* 337, 1107–1111.
17. Allen, H. K., Donato, J., Wang, H. H., Cloud-Hansen, K. A., Davies, J., and Handelsman, J. (2010). Call of the wild: antibiotic resistance genes in natural environments. *Nat Rev Microbiol* 8, 251–259.
18. Liu, Z., Wang, W., Zhu, Y., Gong, Q., Yu, W., and Lu, X. (2013). Antibiotics at subinhibitory concentrations improve the quorum sensing behavior of *Chromobacterium violaceum*. *FEMS Microbiol Lett* 341, 37–44.
19. Fajardo, A., and Martínez, J. L. (2008). Antibiotics as signals that trigger specific bacterial responses. *Curr Opin Microbiol* 11, 161–167. Available at: <http://www.sciencedirect.com/science/article/pii/S1369527408000143>.
20. Davies, J. (2006). Are antibiotics naturally antibiotics? *J Ind Microbiol Biotechnol* 33, 496–499.
21. Torsvik, V. (2011). Prokaryotic Diversity--Magnitude, Dynamics, and Controlling Factors. *Science* 296, 1064–1066.
22. Whitman, W. B., Coleman, D. C., and Wiebe, W. J. (1998). Prokaryotes: the unseen majority. *Proc Natl Acad Sci USA* 95, 6578–6583.
23. Kümmerer, K. (2009). Antibiotics in the aquatic environment--a review--part I. *Chemosphere* 75, 417–434.
24. Li, B., and Zhang, T. (2010). Biodegradation and Adsorption of Antibiotics in the Activated Sludge Process. *Environ Sci Technol* 44, 3468–3473.
25. Chee-Sanford, J. C., Mackie, R. I., Koike, S., Krapac, I. G., Lin, Y.-F., Yannarell, A. C., Maxwell, S., and Aminov, R. I. (2009). Fate and transport of antibiotic residues and

- antibiotic resistance genes following land application of manure waste. *J Environ Qual* **38**, 1086–1108.
26. Sun, P., Barmaz, D., Cabrera, M. L., Pavlostathis, S. G., and Huang, C.-H. (2013). Detection and quantification of ionophore antibiotics in runoff, soil and poultry litter. *J Chromatogr A*.
  27. Heuer, H., Focks, A., Lamshöft, M., Smalla, K., Matthies, M., and Spiteller, M. (2008). Fate of sulfadiazine administered to pigs and its quantitative effect on the dynamics of bacterial resistance genes in manure and manured soil. *Soil Biology and Biochemistry* **40**, 1892–1900.
  28. Thiele Bruhn, S. (2003). Pharmaceutical antibiotic compounds in soils—a review. *Journal of Plant Nutrition and Soil Science* **166**, 145–167.
  29. Jørgensen, S. E., and Halling-Sørensen, B. (2000). Drugs in the environment. *Chemosphere* **40**, 691–699.
  30. Kümmerer, K. (2001). Drugs in the environment: emission of drugs, diagnostic aids and disinfectants into wastewater by hospitals in relation to other sources – a review. *Chemosphere*.
  31. You, Y., Hilpert, M., and Ward, M. J. (2012). Detection of a common and persistent tet(L)-carrying plasmid in chicken-waste-impacted farm soil. *Appl Environ Microbiol* **78**, 3203–3213.
  32. Zhu, Y.-G., Johnson, T. A., Su, J.-Q., Qiao, M., Guo, G.-X., Stedtfeld, R. D., Hashsham, S. A., and Tiedje, J. M. (2013). Diverse and abundant antibiotic resistance genes in Chinese swine farms. *Proc Natl Acad Sci USA*.
  33. Auerbach, E. A., Seyfried, E. E., and McMahon, K. D. (2007). Tetracycline resistance genes in activated sludge wastewater treatment plants. *Water Res* **41**, 1143–1151.
  34. Szczepanowski, R., Linke, B., Krahn, I., Gartemann, K.-H., Gützkow, T., Eichler, W., Pühler, A., and Schlüter, A. (2009). Detection of 140 clinically relevant antibiotic-resistance genes in the plasmid metagenome of wastewater treatment plant bacteria showing reduced susceptibility to selected antibiotics. *Microbiology (Reading, Engl)* **155**, 2306–2319.
  35. Yang, Y., Li, B., Ju, F., and Zhang, T. (2013). Exploring variation of antibiotic resistance genes in activated sludge over a four-year period through a metagenomic approach. *Environ Sci Technol* **47**, 10197–10205.
  36. Yang, H., Byelashov, O. A., Geornaras, I., Goodridge, L. D., Nightingale, K. K., Belk, K. E., Smith, G. C., and Sofos, J. N. (2010). Presence of antibiotic-resistant commensal bacteria in samples from agricultural, city, and national park environments evaluated by standard culture and real-time PCR methods. *Can J Microbiol.* **56**, 761–770.

37. Binh, C. T. T., Heuer, H., Kaupenjohann, M., and Smalla, K. (2008). Piggery manure used for soil fertilization is a reservoir for transferable antibiotic resistance plasmids. *FEMS Microbiol Ecol* 66, 25–37.
38. Delmont, T. O., Prestat, E., Keegan, K. P., Faubladiere, M., Robe, P., Clark, I. M., Pelletier, E., Hirsch, P. R., Meyer, F., Gilbert, J. A., et al. (2012). Structure, fluctuation and magnitude of a natural grassland soil metagenome. *ISME J* 6, 1677–1687.
39. Nesme, J., Cécillon, S., Delmont, T. O., Monier, J.-M., Vogel, T. M., and Simonet, P. (2014). Large-Scale Metagenomic-Based Study of Antibiotic Resistance in the Environment. *Curr Biol*.
40. Heuer, H., Binh, C. T. T., Jechalke, S., Kopmann, C., Zimmerling, U., Krögerrecklenfort, E., Ledger, T., González, B., Top, E., and Smalla, K. (2012). IncP-1 $\epsilon$  Plasmids are Important Vectors of Antibiotic Resistance Genes in Agricultural Systems: Diversification Driven by Class 1 Integron Gene Cassettes. *Front Microbiol* 3, 2.
41. Laroche, E., Pawlak, B., Berthe, T., Skurnik, D., and Petit, F. (2009). Occurrence of antibiotic resistance and class 1, 2 and 3 integrons in *Escherichia coli* isolated from a densely populated estuary (Seine, France). *FEMS Microbiol Ecol* 68, 118–130.
42. Griffiths, R. I., Whiteley, A. S., O'Donnell, A. G., and Bailey, M. J. (2000). Rapid method for coextraction of DNA and RNA from natural environments for analysis of ribosomal DNA- and rRNA-based microbial community composition. *Appl Environ Microbiol* 66, 5488–5491.
43. Normand, P., Ponsonnet, C., Nesme, X., and Neyra, M. (1996). ITS analysis of prokaryotes. *Molecular Microbial Ecology Manual*.
44. Ranjard, L., Poly, F., Combrisson, J., Richaume, A., Gourbiere, F., Thioulouse, J., and Nazaret, S. (2000). Heterogeneous cell density and genetic structure of bacterial pools associated with various soil microenvironments as determined by enumeration and DNA fingerprinting approach (RISA). *Microb Ecol* 39, 263–272.
45. Heuer, H., Solehati, Q., Zimmerling, U., Kleineidam, K., Schloter, M., Müller, T., Focks, A., Thiele-Bruhn, S., and Smalla, K. (2011). Accumulation of sulfonamide resistance genes in arable soils due to repeated application of manure containing sulfadiazine. *Appl Environ Microbiol* 77, 2527–2530.
46. Thevenon, F., Adatte, T., Wildi, W., and Poté, J. (2011). Antibiotic resistant bacteria/genes dissemination in lacustrine sediments highly increased following cultural eutrophication of Lake Geneva (Switzerland). *Chemosphere*.
47. Marti, R., Scott, A., Tien, Y.-C., Murray, R., Sabourin, L., Zhang, Y., and Topp, E. (2013). Impact of manure fertilization on the abundance of antibiotic-resistant bacteria and frequency of detection of antibiotic resistance genes in soil and on vegetables at harvest. *Appl Environ Microbiol* 79, 5701–5709.

48. Chen, J., Yu, Z., Michel, F. C., Wittum, T., and Morrison, M. (2007). Development and Application of Real-Time PCR Assays for Quantification of *erm* Genes Conferring Resistance to Macrolides-Lincosamides-Streptogramin B in Livestock Manure and Manure Management Systems. *Appl Environ Microbiol* 73, 4407–4416.
49. Martinez, J. L. (2008). Antibiotics and Antibiotic Resistance Genes in Natural Environments. *Science* 321, 365–367.
50. Jacquiod, S., Franqueville, L., Cécillon, S., Vogel, T. M., and Simonet, P. (2013). Soil bacterial community shifts after chitin enrichment: an integrative metagenomic approach. *PLoS ONE* 8, e79699.
51. Delmont, T. O., Robe, P., Clark, I., Simonet, P., and Vogel, T. M. (2011). Metagenomic comparison of direct and indirect soil DNA extraction approaches. *J Microbiol Methods*.
52. Delmont, T. O., Simonet, P., and Vogel, T. M. (2013). Mastering methodological pitfalls for surviving the metagenomic jungle. *Bioessays* 35, 744–754.
53. Smith, C. J., and Osborn, A. M. (2009). Advantages and limitations of quantitative PCR (Q-PCR)-based approaches in microbial ecology. *FEMS Microbiol Ecol* 67, 6–20.
54. Musovic, S., Klümper, U., Dechesne, A., Magid, J., and Smets, B. F. (2014). Long-term manure exposure increases soil bacterial community potential for plasmid uptake. *Env Mic Reports* 6, 125–130.

## Chapitre 4.

# ***Le résistome amazonien: évaluation de l'abondance et de la diversité des gènes de résistances à des antibiotiques le long d'un transect de sol en Guyane Française.***

### Résumé.

Un aspect clé des études sur la résistance bactérienne à des antibiotiques dans l'environnement concerne l'évaluation de l'impact anthropique sur les milieux naturels lié à la consommation d'antibiotiques. Trois- Sauts est un village situé dans le Parc National Amazonien de Guyane et isolé de toute source d'antibiotiques mis à part ceux fournis par le poste de santé local qui enregistre chaque prescription distribuée. Un transect de sol a été prélevé trois fois tous les 600 m sur une distance de 3 km; soit 18 échantillons de sol allant du village à la forêt primaire. Nous émettons l'hypothèse qu'une source d'antibiotiques peut affecter la communauté bactérienne du sol en termes d'abondance de gènes de résistance à des antibiotiques et/ou de gènes liés à des éléments génétique mobile (MGE). De l'ADN métagénomique du sol a été extrait à partir de tous les échantillons récoltés et est soumis directement au à une séquençage massif aléatoire (shotgun; Roche 454 pyroséquençage), résultant en 18 jeux de données de lectures de séquences d'ADN métagénomique. Chaque jeu de données a été annoté indépendamment et été criblé à la recherche des gènes de résistance aux antibiotiques par homologie de séquence avec une base de données de référence. Nous avons également procédé à l'évaluation plus précise de l'abondance de certains gènes par PCR quantitative sur les mêmes échantillons d'ADN, en se concentrant sur certains gènes marqueurs typique de la résistance (*bla<sub>SHV</sub>*, *bla<sub>TEM</sub>*, *sul(I)*, *ermA*, *tetA(G)*, *tetA(H)*, *cfiA*) et déjà retrouvés dans des échantillons environnementaux. Des gènes marqueurs d'éléments génétiques mobiles ont également été quantifiés: *int1* pour les integrons et *trfA* pour les plasmides conugatifs. Des éléments génétiques mobiles et des gènes de résistance aux antibiotiques sont présents dans tous les échantillons de sol testés. Cependant, nous n'avons pas observé une augmentation significative de la charge antibiotique des gènes de résistance entre distale (3000m) et le village (0m) échantillons. Ces résultats tendent à confirmer l'incroyable réservoir de déterminants génétiques de résistance aux antibiotiques hébergés par les bactéries du sol et pose la question de la dissémination des gènes de résistance aux antibiotiques dans l'environnement suite à la consommation d'antibiotiques humaine.

## Chapter 4.

# ***Amazonian resistome: evaluating antibiotic resistance abundance and diversity across a French Guiana forest soil gradient through metagenomic approaches.***

### **Abstract.**

A key aspect of environmental resistance studies is to evaluate anthropogenic impact on natural environments due to antibiotic consumption. Trois-Sauts is a French Guiana village located in Amazonian National Park. Trois-Sauts is isolated from any sources of antibiotics but those provided by the local health-post registering every prescription. A soil transect was sampled three times every 600m for 3km resulting in 18 soil samples from the village to the forest. We hypothesized that a single antibiotic source could affect the soil bacterial community in terms of antibiotic resistance genes and mobile genetic element (MGEs) abundance. Soil metagenomic DNA was extracted from all samples and submitted directly to Roche 454 pyrosequencing, resulting in 18 datasets of metagenomic DNA sequence reads. Each dataset was annotated independently and antibiotic resistance genes were screened by sequence homology to a reference database. We also performed quantitative PCR on the same DNA samples, focusing on some antibiotic resistance representative marker genes (*blaSHV*, *blaTEM*, *sul(I)*, *ermA*, *tetA(G)*, *tetA(H)*, *cfiA*) already found in environmental samples. Mobile genetic element markers (such as *int1* for integrons and *trfA* for conjugative plasmids) were also quantified. Mobile genetic elements and antibiotic resistance genes are present in every soil sample. However, we did not observe a significant increase in antibiotic resistance gene load between distal (3000m) and village (0m) samples. These results tend to confirm the already incredible reservoir of antibiotic resistance gene determinants harbored by soil bacteria and question dissemination of antibiotic resistance genes in the environment following human antibiotic consumption.

## Introduction.

Study of the environmental antibiotic-resistance has already proven that many microorganisms from environmental origins are highly resistant to clinical antibiotics. These observations of an abundant reservoir of antibiotic-resistant microorganisms in the environment may have serious implications for the clinical aspect of antibiotic resistance [1-3]. Considering that antibiotic-resistance is a ubiquitous environmental feature, it is of critical concern to conduct risk assessment studies on the impact of human-made activities on the abundance and diversity of environmental antibiotic-resistance genes. It has already been shown that the human-gut microbiome and soil bacteria share identical antibiotic-resistance genes [2]; questioning the genetic permeability of these environments. The soil, or pedosphere, is an extremely complex and heterogeneous environment linked it with all Earth's biosphere compartments and require special attention in the frame of antibiotic resistance studies [4, 5]. Most of our currently used antibiotic pharmacopeia has been isolated from soil microbes such as *Actinobacteria* or fungi [6, 7]. Culture-based [8], functional metagenomic [9, 10] or sequence-based [1] studies showed that antibiotic-resistance gene diversity in soil is order of magnitudes higher than the clinical resistome but also an incredible share of yet-to-discover antibiotic-producers [11, 12].

The soil microbial biomass is incredible and its species diversity is orders of magnitude more important than the most widely described groups of macro organisms, such as arthropods. Estimating the soil microorganism species diversity is a whole field of microbial and theoretical ecology and its estimation can fluctuate a lot, depending on definitions of a bacterial species for example [13, 14]. All microbiologists acknowledge that soil is one of the most diversified ecosystems and that its microbial biomass is by far above

any other [15, 16]. This diversity makes soil microbial ecologists carving for new technologies and conceptual framework to better apprehend this tremendous complexity. In this respect, advances in molecular biology enabled microbiologists to work directly with soil-extracted DNA, processed with more standard techniques, such as PCR, gene cloning, bacterial transformation, opening the field of metagenomics *sensu-lato* [17, 18]. This conceptual revolution gives the opportunity to study not or not-yet-cultivated microorganisms by means of standard laboratory techniques, that may represent most of the environmental microbial diversity [19-21].

ARGD environmental prevalence has already been described many times with various methods; for example cultivation of bacteria from environmental samples such as soil, seawater, sediments, etc.) using selective media [22, 23] or direct identification of genetic markers of antibiotic-resistance genes by molecular biology methods, such as PCR or quantitative PCR [24-26]. But while being very sensitive on focused populations of bacteria and or genes, these methods cannot be used to perform global surveys of environmental antibiotic resistance abundance and diversity across multiple bacterial lineages, for a same sample. Proper risk-assessment of human activities and consumption of medical antibiotics by humans considering the whole picture of antibiotic-resistance, from environmental reservoirs to problematic cases of infections resisting all treatments in the clinic seems now crucial [1, 3, 27]. In this attempt, we are investigating the impact of a recent and precisely recorded source of medical antibiotic molecules in a very isolated area where the local population adopted sedentary lifestyle recently [28]. The hamlet of Trois-Sauts is a recent settlement built on the banks of the upper Oyapock River in French Guiana (Figure 1). A coordinated survey is ongoing since 2001 to study antibiotic resistance carriage in the local animals, humans and wildlife populations [29-31]. Together with this study, abundance of

human commensal bacteria up and downstream the village was measured in the Oyapock River. To complete this work, samples were collected during the 2010 campaign to assess such impacts in the soil environment and compare the local environment based on the recent human sedentary-lifestyle and recorded antibiotic consumption. Soil samples from the village, where humans impact landscape notably, are compared with forest soil samples with varying degrees of human impact along a line-transect. Metagenomic soil DNA extracted from each line-transect sample and either sequenced directly (random shotgun sequencing) or after a 16S rRNA gene PCR amplification of metagenomic DNA. In addition, the same DNA samples are used in quantitative PCR assays targeting selected antibiotic-resistance markers (*blaSHV*, *blaTEM*, *sul(I)*, *ermA*, *tetA(G)*, *tetA(H)*, *cfiA*) and mobile genetic element markers (such as *int1* for integrons and *trfA* for conjugative plasmids). Results showed that antibiotic resistance genes are ubiquitous and abundant in all French Guiana soil samples, as revealed by annotation of shotgun sequencing results or qPCR assays. But is the distribution of antibiotic resistance genes determinant significantly different in the village soil samples than in the forest soil samples? Are antibiotic-resistance genes more abundant in village samples, where human impact is much more important than in distant forest samples? Results obtained with shotgun sequencing and qPCR data were also compared with the differences in microbial taxonomic composition in village and forest soil samples observed with annotation of 16S rRNA genes amplicons from the same DNA. What factor between human urbanization and human antibiotic consumption is having the most significant impacts on the local environmental resistome and microbial communities composition? How can these results relate to the very current problem of acquired-antibiotic resistance in the clinic? [3, 32, 33].

## **Experimental procedures.**

### **Study site.**

Samples were collected in October 2010 during a field campaign in the Amazonian National Park, Camopi administrative area, near the Trois-Sauts settlements in French Guiana. Camopi is the third largest French municipality with an administrated area of 10,300 km<sup>2</sup> (Figure 1a). Access to the inner area of the Amazonian National Park, including Camopi district, is restricted to local indigenous inhabitants since 1970 and a special permit delivered by the local French authorities is required for anyone else to enter the park [34]. Trois-Sauts is one of Camopi settlement and is located upstream the Oyapock river that materialize the Brazil-France border (Figure 1b) [29]. At the time of sampling in October 2010 the Trois-Sauts population was of 597 inhabitants. Their lifestyle is still mainly traditional but is changing rapidly as expected from increased interactions with the more densely populated area downstream the Oyapock (Camopi, 1-2 days journey; Saint-Georges, 2-3 days journey) and facilitated access to the village by motor-boats [28]. The local health-post is still nowadays Trois-Sauts' population only source of antibiotics, delivered upon prescription by resident nurses, allowing a precise record of given drugs (Table 1) [31]. The restricted access to Trois-Sauts due to administrative regulations and its isolated situation deep upstream the Oyapock makes exogenous antibiotic pollution unlikely and is confirmed by local habits [28].

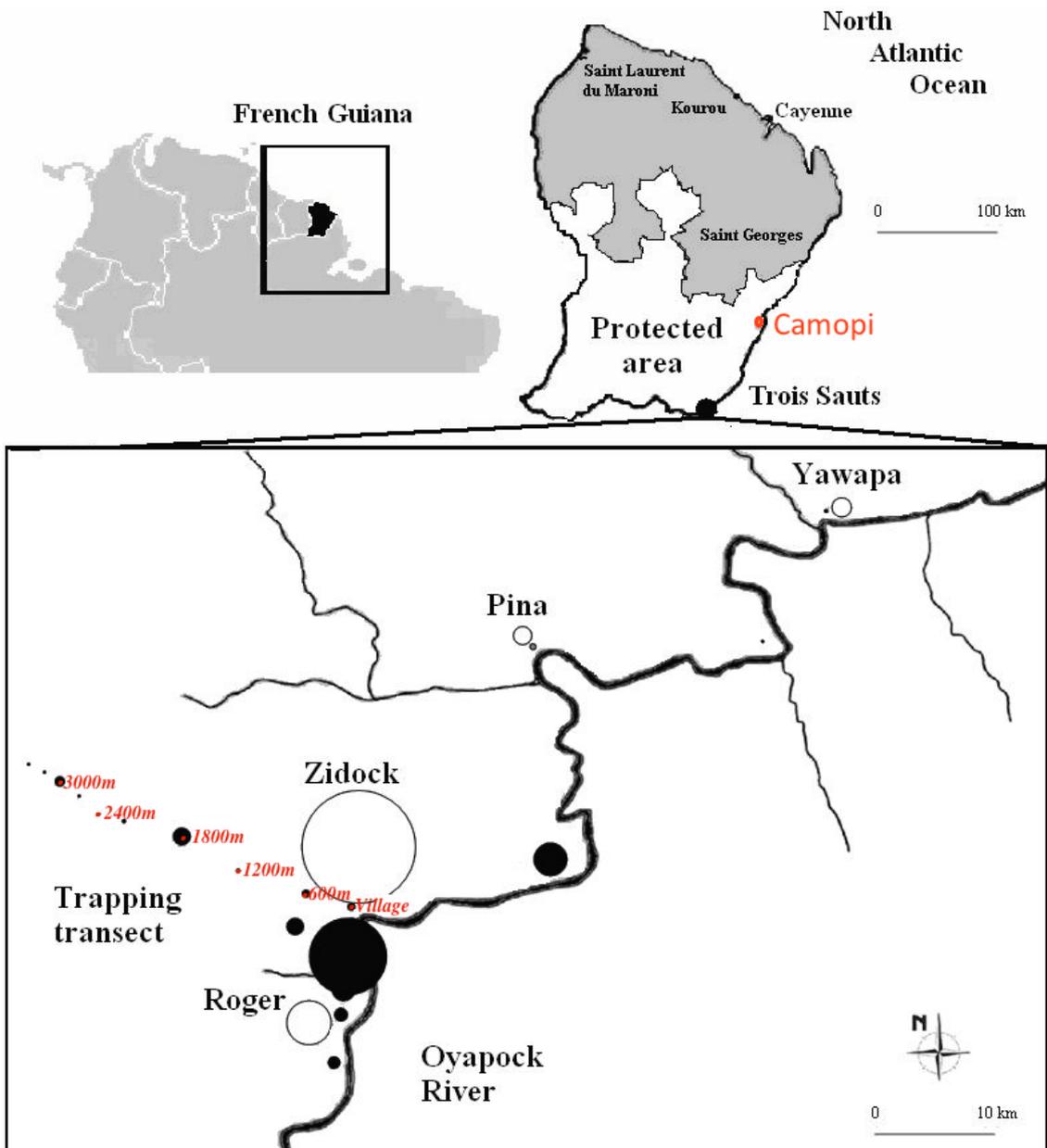
### **Soil sampling.**

Soil has been sampled at 0m; 600m; 1200m; 1800m; 2400m and 3000m distances along a line-transect going from Trois-Sauts village Health post (Village, 0m) to the primary forest at Northeast facing away from the river (Figure 1). This direction is chosen based on

the presence and maintenance of a hunting-trail used by the village hunter and since opening a new trail is a very intensive operation due to the dense vegetation surrounding the village. Samples of about 1kg are taken under the soil litter on the first 10 cm of soil using a clean shovel and immediately stored in plastic bags. For each distance, three soil samples are collected away from the trail, where the soil was not notably disturbed by its maintenance, at each tip of a 1m equilateral triangle where GPS coordinates are recorded in its center. This sampling plan result in three soil samples for each six distances with recorded GPS position (0m; 600m; 1200m; 1800m; 2400m and 3000m) for a total of 18 samples for which specific characteristics (coordinates, landscape type, distance) are recorded (Table 2). For comparison purpose we included in the analysis a soil sample from Rothamsted Research Park Grass long-term experiment control plot [35] from which environmental metagenomic DNA has been sequenced in a previous study [36].

### **Metagenomic DNA extraction**

Upon reception soil samples are sieved using a 2mm mesh to eliminate coarse stone and plant material. Metagenomic DNA is extracted from our samples using Power Max Soil DNA kit from MO BIO (Carlsbad, CA, USA) following manufacturer's instruction. Briefly, this kit consist in a first step of mechanical lysis using silica and glass beads followed by an adapted purification of soil DNA including humic-acids removal by precipitation, immobilization of DNA on a silica membrane followed by ethanol washes and elution in ultrapure water. DNA concentration following extraction is measured using a fluorescent



**Figure 1 – Maps of French Guiana and the Amazonian National Park.**

Maps of respectively: French Guiana administrative boundaries; the Amazonian National Park protected area and the Trois-Sauts settlements (Roger; Zidok; Pina and Yawapa) located on the East bank of the Oyapock river. Circle are proportional to the number of resident at each location. In red: soil sampling sites for each distance going from Zidok ("Village"; 0m) to 3000m in the primary forest; black dots: hunting trail followed for the line-transect soil sampling. Adapted from Lescat, et al. (2013). *Env Mic Reports* 5, 49–57.

molecular probe (QuBit, Life Technologies, Carlsbad, CA, USA) and diluted to a final concentration of 10ng.µL<sup>-1</sup> DNA for subsequent use in quantitative PCR assays, 16S rRNA PCR products and metagenomic DNA sequencing. The same DNA extraction method (PowerSoil MoBio, Carlsbad, USA) was used to extract Rothamsted soil samples at the time of the previous study (2008). The same procedure is used here since using different extraction protocols has been shown to introduce significant changes quantitative (abundance of reads) and qualitative (annotation of reads) in environmental bacterial community extracted-DNA making inter-samples comparisons hazardous [37].

Sample	Distance	Habitat	Latitude	Longitude
0m1	0m	Village	2.257016667	-52.87328333
0m2			2.257016667	-52.87328333
0m3			2.257016667	-52.87328333
600m1	600m	Very degraded forest	2.260333333	-52.87686667
600m2			2.260333333	-52.87686667
600m3			2.260333333	-52.87686667
1200m1	1200m	Secondary forest	2.263233333	-52.8834
1200m2			2.263233333	-52.8834
1200m3			2.263233333	-52.8834
1800m1	1800m	Secondary forest	2.265016667	-52.88873333
1800m2			2.265016667	-52.88873333
1800m3			2.265016667	-52.88873333
2400m1	2400m	Primary forest	2.267716667	-52.8935
2400m2			2.267716667	-52.8935
2400m3			2.267716667	-52.8935
3000m1	3000m	Primary forest	2.270666667	-52.89836667
3000m2			2.270666667	-52.89836667
3000m3			2.270666667	-52.89836667
Roth.H32	Roth.	Grassland	n.d.	

**Table 1 – Geographic characteristics of soil samples collected at Trois-Sauts, French Guiana and Rothamsted (UK) used in this study.**

Tables showing Geographic characteristics of soil samples collected at Trois-Sauts, French Guiana and Rothamsted (UK) used in this study. "Distance" represent the relative distance to the village of Trois-Sauts health-post (distance 0m) for each sampling sites along a hunting-trail line-transect (total length = 3.2km as determined by GPS coordinates). Latitude North and longitude West are given as WSG84 decimal coordinates and a single GPS measure is taken for each distance group. Description of the habitat is reported as a subjective indication of land cover at sampling sites (personal communication from F. Catzeflis & C. Angebault; Lescat et al. (2013).*Env Mic Reports* 5, 49–57.)

### **Environmental shotgun DNA and 16S rRNA amplicons sequencing.**

All 18 samples have been sequenced using Roche Titanium FLX+ 454 pyrosequencing technology directly on Environmental DNA (shotgun sequencing [38]) using a different multiplex identifier (MID or barcode) for each sequencing library resulting in 19 different datasets of sequences after filtering unique identifiers (or barcode [39]). A sum of  $3.09 \cdot 10^6$  sequence reads was obtained for shotgun sequencing. PCR targeting the 560pb V4-V6 region of the 16S rRNA using forward 16S-0515F and reverse 16S-1061R primers [40] and a unique MID resulting in PCR products with unique identifier for each sample. PCR products are then pooled in stoichiometric proportion and sequenced using Roche 454 pyrosequencing (Titanium FLX+) technology. A total of 156417 16 rRNA amplicons sequence reads were obtained. All sequencing reactions are performed by Beckman Genomics (Denver, MA, USA) after provision of at least 1 $\mu$ g of either purified metagenomic DNA for environmental sequencing or purified MID-tagged 16S rRNA PCR products for 16S rRNA amplicons sequencing. All details are available in Table S2.

### **Taxonomic assignation of 16S rRNA metagenomic sequences.**

Roche 454 pyrosequencing output (SFF files) obtained from 16S rRNA PCR amplicons are provided by our sequencing service (Beckman Genomics, Denver, MA, USA) and is treated using the QIIME pipeline (Caporaso et al. 2010) scripts for taxonomic and phylogenetic analysis as described in supplemental information (Supplemental Experimental Procedures). Briefly, sequences reads are filtered based on quality score (PHRED > 20) and read length (280pb > ok length > 560pb); when reverse primer is detected, sequences reads are cut at its 3'end and remaining 3' sequence is trimmed. Remaining sequence reads are clustered based on global alignment similarity. Threshold for *de novo* OTU clustering is set at

97% sequence similarity using CD-HIT (Niu et al. 2010; Fu et al. 2012). Taxonomic assignment of *de novo* OTU is performed with RDP Classifier version 2.6 (naïve-Bayesian classifier) trained on 16S rRNA training set 9 (Wang et al. 2007) (<http://rdp.cme.msu.edu/classifier>). Phylogenetic tree is computed using FastTree 2.1.3 (Price et al. 2010). Weighted and un-weighted UniFrac distances metrics are used to compare phylogenetic  $\beta$ -diversity between samples with and without respectively taking into account abundance (Lozupone et al. 2011).

Sample	Num. reads	Mean size	Obs. OTU ave.	Chao1 ave.	Shannon ave.
16S rRNA 0m1	3465	529	1155.5	2522.227	9.418
16S rRNA 0m2	10981	526	2988	5175.616	10.672
16S rRNA 0m3	7673	526	2266.3	4170.337	10.4
16S rRNA 600m1	9497	525	2319.8	3982.974	10.093
16S rRNA 600m2	13962	524	2184.5	3702.679	9.987
16S rRNA 600m3	11162	525	2454.8	4441.165	10.153
16S rRNA 1200m1	11973	524	2226.8	3683.98	10.002
16S rRNA 1200m2	15	530	n.d.	n.d.	n.d.
16S rRNA 1200m3	11570	526	2394.6	4189.277	10.198
16S rRNA 1800m1	8942	526	2291.3	3939.192	10.344
16S rRNA 1800m2	5217	524	1676.8	3498.404	9.901
16S rRNA 1800m3	13647	526	2690.4	4810.303	10.417
16S rRNA 2400m1	13248	526	2323.8	4106.202	10.068
16S rRNA 2400m2	6888	529	1795.7	3450.204	9.718
16S rRNA 2400m3	5203	527	1412.5	2415.833	9.565
16S rRNA 3000m1	5245	526	1957.3	3896.504	10.298
16S rRNA 3000m2	5740	525	1713.9	3180.905	9.997
16S rRNA 3000m3	12004	525	2563.5	4420.845	10.357

**Table 2 - French Guiana amazonian and Rothamsted H32 soil samples 16S rRNA gene V4-V6 region amplicons sequencing statistics and diversity.**

Table showing statistics for each French Guiana amazonian and Rothamsted H32 soil samples 16S rRNA gene V4-V6 region amplicons sequencing and diversity. Each soil samples metagenomic DNA is amplified by PCR using universal primers targeting Archaea and Bacteria 16S rRNA gene and products are sequenced with Roche 454 pyrosequencing. Samples diversity statistics are calculated using rarefaction analysis based on random sequence resampling. Sample 1200m.2 yield only 15 sequence reads and has been discarded for downstream analysis. Colors shading indicate relative differences per variable (column) for each value; from red (lowest value) to dark green (highest value).

### **Shotgun metagenome functional and SSU annotation.**

Sequence reads obtained from French Guiana soil sample metagenomic DNA environmental shotgun pyrosequencing runs are annotated using the MG-RAST platform (metagenomics.anl.gov) [41]. Sequence reads are uploaded as raw SFF files using default parameters for sequence data *denoising*. PCA plots are calculated at the function level using the SEED subsystem database [42]. Metagenomes are also searched for SSU and an LSU matches, at a threshold of e-value  $<10^{-5}$  with a 60% minimal similarity over an alignment length  $>15$ pb between shotgun soil metagenome sequence reads and RDP database project on the MG-RAST server. Taxonomic assignment serves both the purpose of comparing taxonomy profile based on shotgun or PCR sequencing and normalizing relative ARGD abundance and diversity in each sample by SSU/LSU counts (Table 3).

### **Antibiotic resistance gene annotation of environmental DNA sequences.**

After demultiplexing using MID-tag unique per sample as described above, artificial duplicates known to occur with Roche 454 pyrosequencing are removed using cd-hit-454 [43] with default parameters. Sequence similarities are searched between clean datasets and an amino-acid sequence database of known antibiotic resistance proteins (ARDB; [44]). Only reads with a 35% amino-acid sequence similarity match or above and a 60 BLASTx bit-score or above are considered positive matches. This threshold was used in another study using the same database [26] and also used in our previous analysis using public metagenomes datasets. This threshold offers a trade-off between sensitivity (less false-negatives ARGD reads annotations) and specificity (less false-positive ARGD read annotations). For a thorough description of the ARGD annotation process, please refer to Chapter 2.

Sample	Distance	MG-RAST Ac. Num.	Num. reads cd-hit	Mean size cd-hit (bp)	Archaea reads	Bacteria reads	Eukaryota reads	Viruses reads	Total 16S rRNA reads	Num. ARGD reads	%SSU	%ARGD	ARGD/SSU	Dist. grp. mean	SD
0m1		4520480.3	121805	332	129	1802	227	34	1931	702	1.59%	0.58%	0.36		
0m2	0m	4520481.3	228647	348	166	2035	285	31	2201	1410	0.96%	0.62%	0.64	0.44	0.18
0m3		4520482.3	98982	334	133	1703	193	22	1836	556	1.85%	0.56%	0.30		
600m1		4520495.3	141137	321	109	1594	190	11	1703	739	1.21%	0.52%	0.43		
600m2	600m	4520496.3	170526	319	121	1660	209	10	1781	842	1.04%	0.49%	0.47	0.52	0.12
600m3		4520497.3	216303	333	124	1697	229	14	1821	1186	0.84%	0.55%	0.65		
1200m1		4520483.3	121203	290	105	1573	163	16	1678	546	1.38%	0.45%	0.33		
1200m2	1200m	4520484.3	152839	287	126	1734	190	20	1860	698	1.22%	0.46%	0.38	0.39	0.08
1200m3		4520485.3	177845	295	114	1572	170	10	1686	810	0.95%	0.46%	0.48		
1800m1		4520486.3	88218	372	119	1809	218	21	1928	602	2.19%	0.68%	0.31		
1800m2	1800m	4520487.3	147999	380	120	1754	206	20	1874	1026	1.27%	0.69%	0.55	0.40	0.13
1800m3		4520488.3	101088	366	123	1802	236	23	1925	652	1.90%	0.64%	0.34		
2400m1		4520489.3	177368	377	108	1809	225	18	1917	1267	1.08%	0.71%	0.66		
2400m2	2400m	4520490.3	146376	397	134	1852	236	16	1986	1030	1.36%	0.70%	0.52	0.62	0.09
2400m3		4520491.3	186868	399	159	1893	250	34	2052	1419	1.10%	0.76%	0.69		
3000m1		4520492.3	146472	369	112	1687	198	16	1799	993	1.23%	0.68%	0.55		
3000m2	3000m	4520493.3	205050	340	128	1808	236	17	1936	1227	0.94%	0.60%	0.63	0.63	0.07
3000m3		4520494.3	269074	322	135	1886	245	33	2021	1398	0.75%	0.52%	0.69		
Roth.H32	Roth.	4453406.3	187214	360	151	2772	519	128	2923	1151	1.56%	0.61%	0.39	0.39	nd

**Table 3 - French Guiana amazonian and Rothamsted H32 soil samples environmental metagenomic DNA shotgun sequencing and annotation summary.**

Table displaying French Guiana amazonian and Rothamsted H32 soil samples environmental metagenomic DNA shotgun sequencing results statistics obtained by Roche 454 Titanium FLX+. The number of SSU reads are determined using the MG-RAST website; Total 16S rRNA reads = sum of Archaea- and Bacteria-annotated SSU reads. Num. ARGD reads = total number of ARGD-annotated reads found in each dataset using the ARDB database. %SSU = relative percentage of SSU-annotated reads per total number of reads. %ARGD = relative percentage of ARGD-annotated reads per total number of reads; ARGD/SSU = relative proportion of ARGD-annotated reads per SSU-annotated reads; in each dataset.

### **Quantitative PCR assays using antibiotic resistance marker genes.**

Metagenomic DNA obtained from soil samples are used in SYBR green quantitative PCR (qPCR) to quantify selected antibiotic resistance gene markers on the same DNA pools used for 16S rRNA and environmental shotgun sequencing (Table 2). SYBR green is a DNA intercalating agent that binds to double-stranded DNA during replication, *i.e.* during PCR elongation cycles, making possible to detect trace amount of specifically amplified DNA and calculate its concentration in the reaction tube before the reaction using a calibration-curve of fluorescence intensity as function of initial copy of positive template controls. For genes *bla<sub>SHV</sub>*, *ermA*, *tetA(G)* and *cfiA* used in these quantitative assays, positive controls corresponding to each gene targeted region are synthesized *de novo*, as described by David et al. (2008) [45] (Figure S1, Table S1). Briefly: long single-stranded oligonucleotides spanning the entire length of each target gene amplified region are synthesized with cohesive 3' ends and are hybridized and synthesized to full length in a PCR-like reaction using Taq polymerase. The resulting PCR products are purified on agarose-gel, after checking their correct size, using the Illustra Gel & PCR CleanUp kit (GE Healthcare). Resulting purified PCR products are quantified, serially diluted and used as template to determine quantitative PCR calibration curve for each assay (*i.e.*: target gene). For qPCR targeting genes *IncP-trfA*, *int1*, *sul1* and *tetA(H)*, plasmid pKJK5 is used a positive template (NCBI accession: NC\_008272.1 ; courtesy of Pr. Søren J. Sørensen, Copenhagen University [46]). Relative quantification between samples is normalized using qPCR results obtained with primers 16S rRNA gene universal primers for each reaction. Calibration curves for 16S rRNA gene quantification are performed using a serial-dilution of known copy-numbers of *Sphingobium francense* type-strain Sp+ 16S rRNA gene, a soil borne Proteobacteria [NCBI TaxId: 332057]. Reactions are performed in a final volume of 20µL using 10µL SensiFast

SYBR No-Rox (Bioline) reaction buffer containing Taq polymerase, 0.8 $\mu$ L of each primer at 100nM and 2 $\mu$ L of template DNA at 10ng. $\mu$ L<sup>-1</sup> and run with the following program in a qPCR-thermocycler RotorGen 3000 (Quiagen, Courtaboeuf, France): 10 mins at 95°C and 40 cycles of 30 secs at 95°C; 30 secs at specific T<sub>a</sub> (see Table S1); 30 secs at 72°C. A final melting analysis is performed recording fluorescence while temperature ramps from 50°C to 99°C at 1°C per second. A serial dilution of 10<sup>8</sup> to 10<sup>2</sup> copies of positive-control DNA template for each gene is used to assess reaction efficiency and serve as calibration curve for determining cycle threshold. Only reactions with a calibration curve goodness-of-fit R<sup>2</sup>-value above 0.99 and efficiency between 0.8 and 1.2 were considered valid for quantification purpose (Table 4). For each reaction, reaction tube CT is determined for three independent reactions using the same soil metagenomic DNA as template using Corbett Research RotorGene software version 2.1.0 (Build 9) built-in threshold calibration setting tool. CT values obtained for the same DNA sample are considered as experimental replicates and analyzed accordingly to extract technical standard deviation from quantitative experiments. Relative quantification between samples is possible after normalizing quantification data with an internal calibrator (16S rRNA gene). Results can then be expressed as fold change in relative target quantification (Figure 5) between experimental samples and control sample (Rothamsted soil sample extracted with the same procedure).

#### **Data analysis.**

Statistical computations are done using the R-software (<http://www.r-project.org/>) and additional libraries: vegan [47] (hierarchical analysis), ade-4 [48] (multivariate analysis), pvclust [49] (bootstrap analysis). Geographic system information displays and analyses are done using GenGIS 2.0 [50].

Target gene	Gene function	Resistance to	qPCR efficiency	Calibration curve R2	Accession Number	Ref.
<b>16S rRNA</b>	Ribosomal RNA	n.a.	0.91	0.9918	NR_042944.1	Takai et al. (2000)
<b>blaSHV</b>	SHV-11 type extended spectrum $\beta$ -lactamase	Penicillins, cephalosporins	0.86	0.9958	KF585137	Hammond et al. (2005)
<b>blaTEM</b>	TEM-1 type extended spectrum $\beta$ -lactamase	Penicillins, cephalosporins	0.95	0.9985	AY425988.1	Demanèche et al. (2008)
<b>cfiA</b>	cfiA metallo- $\beta$ -lactamase	Penicillins, cephalosporins, cephamycin, imipenem	1.08	0.9969	JF795577.1	Pumbwe et al. (2007)
<b>ermA</b>	Erythromycin rRNA adenine N-6-methyltransferase 1	Macrolides, lincosamides, type B streptogramins	1.01	0.9957	NC_002952.2	Merino-Diaz et al. (2007)
<b>sulI</b>	Sulfonamid-resistant dihydropteroate synthase	Sulfonamides	1.02	0.9999	AM261282	Heuer and Smalla (2007)
<b>tetA(G)</b>	Class G TetA tetracycline-specific efflux pump	Tetracyclines	0.95	0.9968	S52437	Fan et al. (2007)
<b>tetA(H)</b>	Class H TetA tetracycline-specific efflux pump	Tetracyclines	0.93	0.9997	U00792	Fan et al. (2007)
<b>Int1</b>	Class 1 integrons tyrosine recombinase	n.a.	0.92	0.9997	AM261282	Hoyle et al. (2006)
<b>TrfA</b>	Broad-host plasmid IncP trfA replication initiation	n.a.	1	0.9911	AM261282	Bahl et al. (2009)

**Table 4 – Characteristics of antibiotic resistance and mobile genetic elements marker genes used in this study and respective quantitative PCR reactions efficiency and calibration-curve goodness-of-fit.**

Accession Number indicate the respective sequence template used as positive control for each reaction on NCBI sequence database.

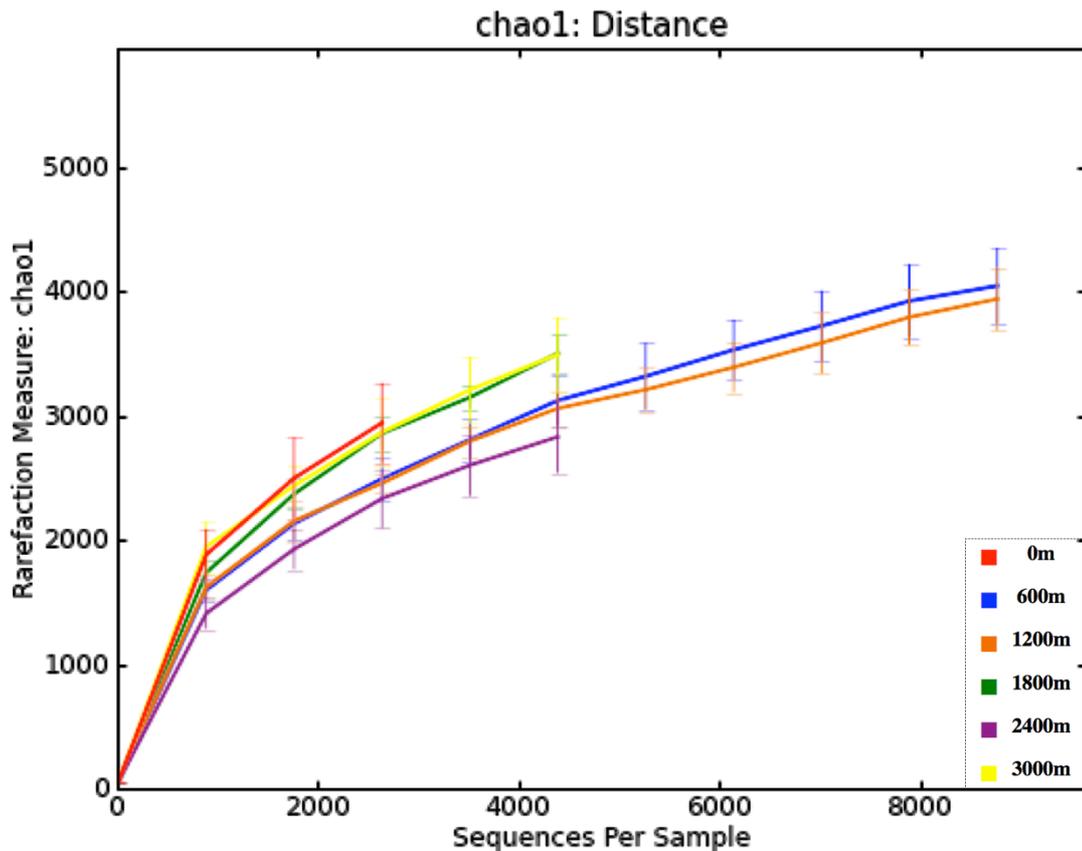
$$\text{ratio} = \frac{(E_{\text{target}})^{\Delta C_p \text{ target (control - sample)}}}{(E_{\text{Ref}})^{\Delta C_p \text{ Ref (control - sample)}}$$

**Equation 1 - Ratio formula.**

Formula used to calculate a gene quantification ratio (target ratio) between two samples using a reference gene (16S rRNA gene) as sample calibrator, in reference to a control sample (Rothamsted , Park Grass soil). Pfaffl, M. W. Et al. (2002). Nucleic Acids Res 30, e36.

## Results

Microbial community profile differences along the line-transect forest gradient are assessed analyzing 16S rRNA amplicons annotation profile for each soil sample. Before taxonomic affiliation, sequence reads (n=2897800; 0.995Gbp) are clustered at 97% similarity as described in experimental procedures. A total of 12647 different OTUs (i.e. clusters of sequences at 97% similarity using cd-hit) are found in French Guiana soil samples. Representative sequence for each cluster is compared to a reference database of full-length 16S rRNA sequences annotated with taxonomic information (RDP GreenGenes ver. 12.10). Sample 1200m.2 yielded only 15 sequences reads after demultiplexing and was therefore discarded from subsequent analyses. Sample OTU-species richness and evenness (alpha-diversity [51]) are estimated using respectively Chao1 estimator and Shannon's index of diversity. Sequence read datasets are rarefied with random read resampling in an effort to standardize sampling depth before alpha-diversity estimation for each sample (Table S3). Rarefaction curve of Chao1 for each distance group is displayed with 95% confidence interval error bars (Figure 2). Chao1 estimator; Shannon's index and total number of observed OTUs alpha-diversity metrics are calculated independently for each sample within each distance group and results are given in Table 5. Datasets with fewer sequences such as 0m.1 (n=3465 seqs.) or 2400m.3 (n=5203 seqs) are typically giving lower observed diversity (respectively: 1155.2 and 1412.5 different observed OTUs; Chao1=2522.3 and 2415.8; Shannon=9.42 and 9.57). Despite these differences due to unequal sequencing depth; datasets are relatively homogenous in terms of intra-diversity and can be further compared without introducing significant biases.

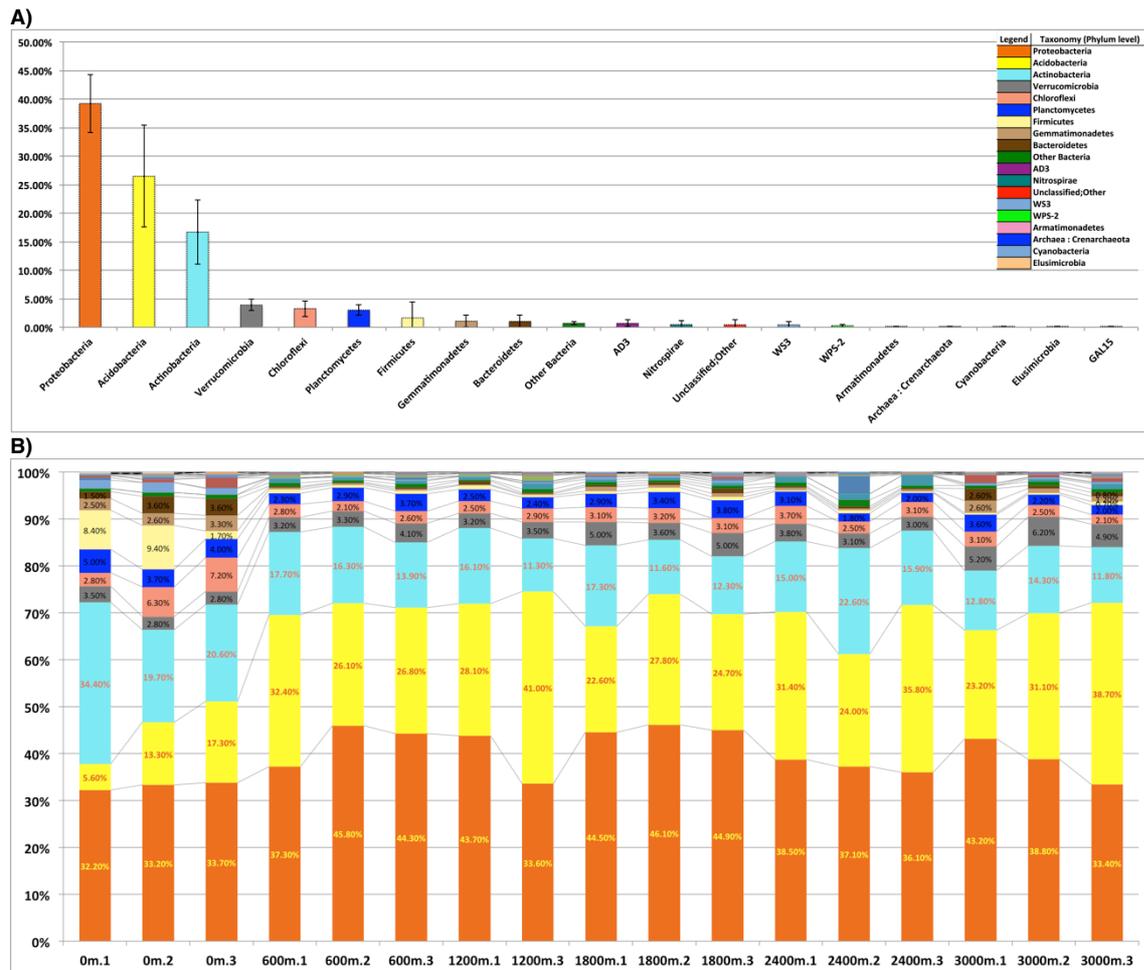


**Figure 2 - Chao1 index rarefaction curve for each distance groups of 16S rRNA datasets.**

Chao1 index is calculated for each sample based on independent resampling of the *de novo* OTU abundance table in each sample with similar OTU defined at 97% nucleotide similarity of 16S rRNA gene V4-V6 region sequence. Rarefaction curves are plotted using independent Chao1 index for each sample, grouped by respective distance: 0m; 600m; 1200m; 1800m; 2400m and 3000m. Error-bars are indicating 95% confidence interval for Chao1 within the same distance group.

### Community taxonomy diversity along the soil line-transect.

To assess significant per sample differences in the soil transect taxonomic composition, relative phylum abundance is analyzed using results from RDP Classifier (Figure 2B). The dominant taxa distribution between samples is relatively homogeneous at the phylum level. At a global abundance scale (Figure 3A), most phyla are ordered in most samples as follow: Proteobacteria (tot. mean. 39.2%  $\pm$ 5.08 SD), Acidobacteria (26.50%  $\pm$ 8.88), Actinobacteria (16.70%  $\pm$ 5.62) (Figure 3A). The three next most abundant phyla in most samples are Verrucomicrobia, Chloroflexi and Planctomycetes accounting respectively for 3.89% $\pm$ 1; 3.27% $\pm$ 1.38; 3.02% $\pm$ 0.88 (Figure 3A).



**Figure 3 – Relative phylum abundance in 16S rRNA amplicons sequence reads datasets for all samples (A) and per sample (B).**

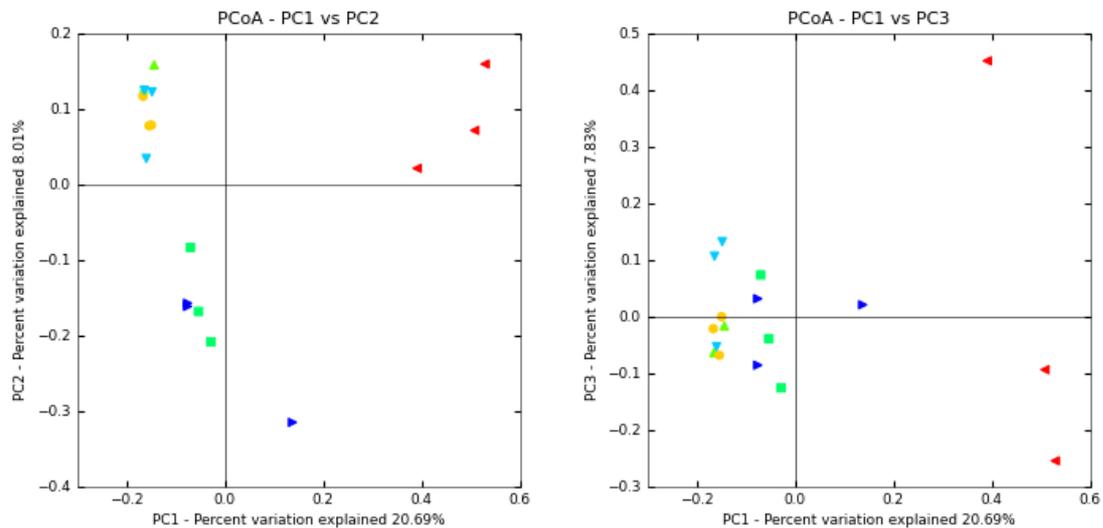
Taxonomic profile are displayed at the phylum level of soil samples 16s rRNA gene PCR amplification products (amplicons) sequenced (Roche 454) and annotated using RDP Classifier 2.6. Relative percentage of each taxon is labeled for most dominant phylums in panel B: Proteobacteria, Acidobacteria, Actinobacteria, Verrucomicrobia, Chloroflexi, Planctomycetes, Firmicutes (0m.1; 0m.2 and 0m.3) and Gemmatimonadetes (0m.1; 0m.2; 0m.3 and 3000m.1).

Figure 3B displays relative phylum abundance in each soil sample dataset. In village samples (0m.1 and 0m.2), Firmicutes are notably more abundant than in other samples with  $8.9\% \pm 0.71$  (sample 0m.3:  $1.70\%$ ) of total OTU compared to  $0.6\% \pm 0.21$  in all other samples. Gemmatimonadetes, Bacteroidetes are also more abundant in village (respectively:  $2.9\% \pm 1.21$  and  $2.8\% \pm 0.44$ ) than forest samples, with the exception of sample 3000m.1 showing similar proportion (forest:  $0.55\% \pm 0.3$  and  $0.48\% \pm 0.24$ ; 3000m.1:  $2.6\%$ ). Nitrospirae are also more abundant in village samples ( $1.7\% \pm 0.4$  versus  $0.2\% \pm 0.16$  in forest samples) but all other phyla are detected below 1% relative abundance in all samples (Table S3). Also,

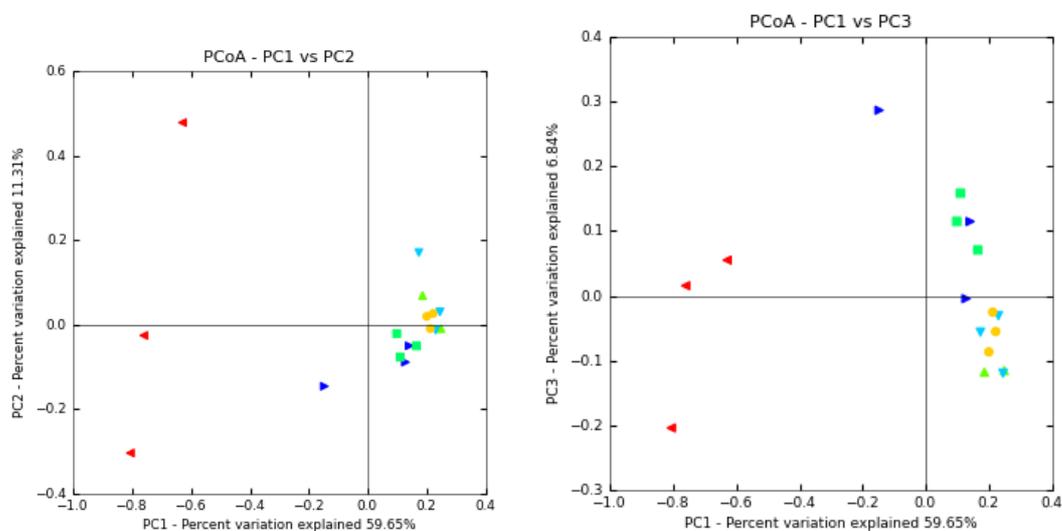
in sample 2400m.2, the relative abundance of unclassified sequences but not belonging to bacteria ("Unclassified; other") accounts for 2.8% whereas in other samples unclassified sequences account only between 0% and 0.5% of all reads in dataset. Two phyla are not detected in any sample in this analysis: Euryarchaeota (methanogens, large range of extremophiles) and BRC1 candidate subdivision defined on the basis of 16S rRNA divergence of sequence retrieved from rice paddy microcosm soil metagenome. It is acknowledged that due to specific divergence in their 16S rRNA gene sequence, BRC1 is not amplified correctly with the primers used in this study [40]. All taxa relative abundance tables and respective bar plots, for each sample at the: class, order, family and genus taxonomic level are available as annexes, hosted online for convenience (<http://bit.ly/1ilxj9z> ; 15 pages). Differences in soil sample bacterial community composition ( $\beta$ -diversity [51]) can be estimated using UniFrac phylogenetic-distance [52] between 16S rRNA sequences in each dataset. Results are plotted using a principal coordinates analysis (PCoA) in Figure 4. Weighted and un-weighted UniFrac distances express taxonomic composition dissimilarity in samples respectively taking into account differences in OTU relative abundance or not (presence / absence only); ANOSIM p-value < 0.001 for both measures. In both plots, samples belonging to 0m distance group (Village) are clustered out of the forest samples along PC1 axis in un-weighted and weighted distances PCoA plots with: 20.69% and 59.65% variations explained respectively. In qualitative terms (i.e.: neglecting relative abundance: un-weighted UniFrac distance), PC2 axis (8.01% variations explained) is dividing forest samples in two groups of distance: group 1: 600m; 1200m and 2400m samples and group2: 1800m and 3000m samples (Figure 4A). This distinction between forest samples is due to taxa with a low relative abundance since this difference is not reflected by the analysis taking into consideration the quantitative aspect of differences in sample taxonomic

composition (i.e.: weighted UniFrac distance, using taxa relative abundance) where forest samples are clustered together (Figure 4B).

### A) Unweighted



### B) Weighted



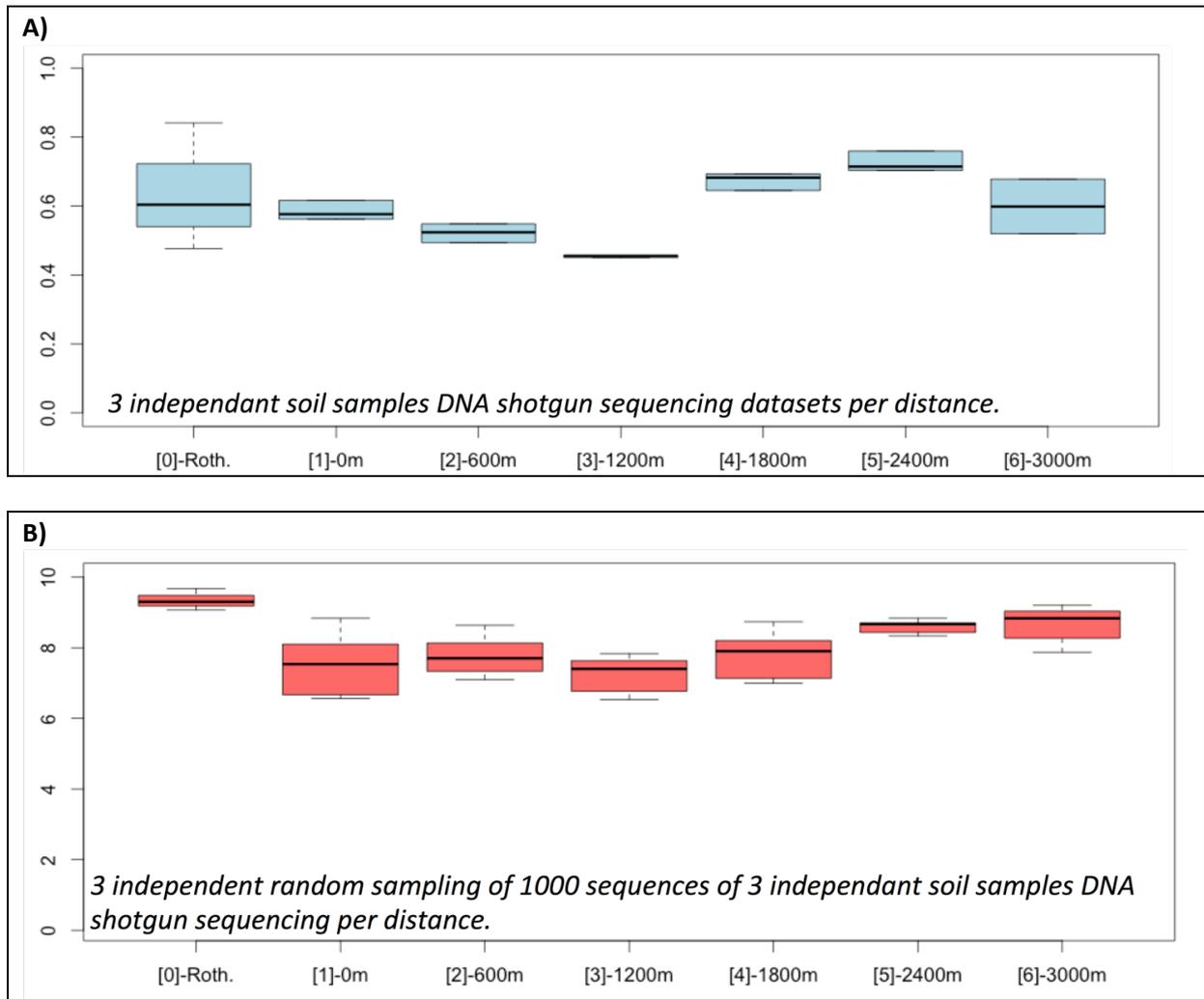
**Figure 4 – Plots showing results of unweighted (A) and weighted (B) UniFrac distances principal coordinate analysis.**

16S rRNA gene PCR amplification products (amplicons) are sequenced by Roche 454 pyrosequencing and aligned using PyNAST and each 18 soil samples 16S rRNA gene amplicons datasets are compared using UniFrac distance. Unweighted UniFrac distance (A) measure OTU composition dissimilarity between sample (presence / absence) ; whereas weighted UniFrac distance (B) also measures dissimilarity between OTU relative abundance in each sample. Samples are colored as follow: red: 0m; yellow: 600m; light green: 1200m; light blue: 2400m; dark blue: 3000m. Village samples (red) are clearly grouping away from forest samples in all analysis along the axis component explaining most of variance between samples. Unweighted and weighted UniFrac ANOSIM:  $R=0.4521$  and  $R=0.4885$  respectively;  $p\text{-value}=0.0001$  for both, determined on 9999 permutations.

### **Antibiotic resistance abundance and diversity in soil shotgun sequencing.**

ARGD-annotated sequence reads are detected in all datasets obtained from French Guiana soil samples metagenomic DNA shotgun sequencing (Figure 5). In order to characterize each sample in term of ARGD composition, abundance and diversity of ARGD is calculated as follow. Antibiotic-resistance gene abundance is determined as the fraction of ARGD-annotated reads on the total number of reads per samples (Figure 5A). Soil samples are not showing significant differences in term of relative abundance of ARGD-annotated reads that accounts for  $0.644\% \pm 0.314\%$  of total sequences. This value is in the range of the Rothamsted Park Grass soil ARGD abundance. Relative ARGD diversity (Figure 5B) is estimated as the number of different ARGD sequences (ARGD<sub>id</sub>) found per sample divided by the total number of sequence in ARDB (n=2999) and similarly does not show a significant gradient effect when grouping samples based on relative distance to the village. Using similar clustering method as described above for measuring taxonomic distribution dissimilarity considering each ARGD sequence in the Antibiotic Resistance Database (ARDB, [44]) as a species. Antibiotic resistance genes are classified in the database as a function of their resistance type (same enzyme / protein variants), resistance class (same enzyme family) and the antibiotic molecule they confer resistance (Table S3). Clustering (Ward algorithm) of samples based on Bray-Curtis distance dissimilarity computed using ARGD-annotated reads is showing a distinct pattern of ARGD types distribution in the different soil samples (Figure 8). Most ARGD found in French Guiana soil samples are belonging to multi-drug efflux pumps proteins, as found in other soil metagenomes (cf. Chapter 2). ARGD-annotated reads affiliated with Vancomycin resistance operons are the second most abundant category of resistance genes found in French Guiana soil samples and accounts for  $0.08\% \pm 0.0033$ ; followed by bacitracin-specific efflux ( $0.04\% \pm 0.0014$ ), beta-lactams

( $0.03\% \pm 0.0014$ ) and tetracyclines ( $0.03 \pm 0.0008$ ) in French Guiana soil samples. The distribution of these different ARGD is however changing along the soil line-transect but no gradient of ARGD-annotated read abundance as a function of distance village is observed (Figure 7).

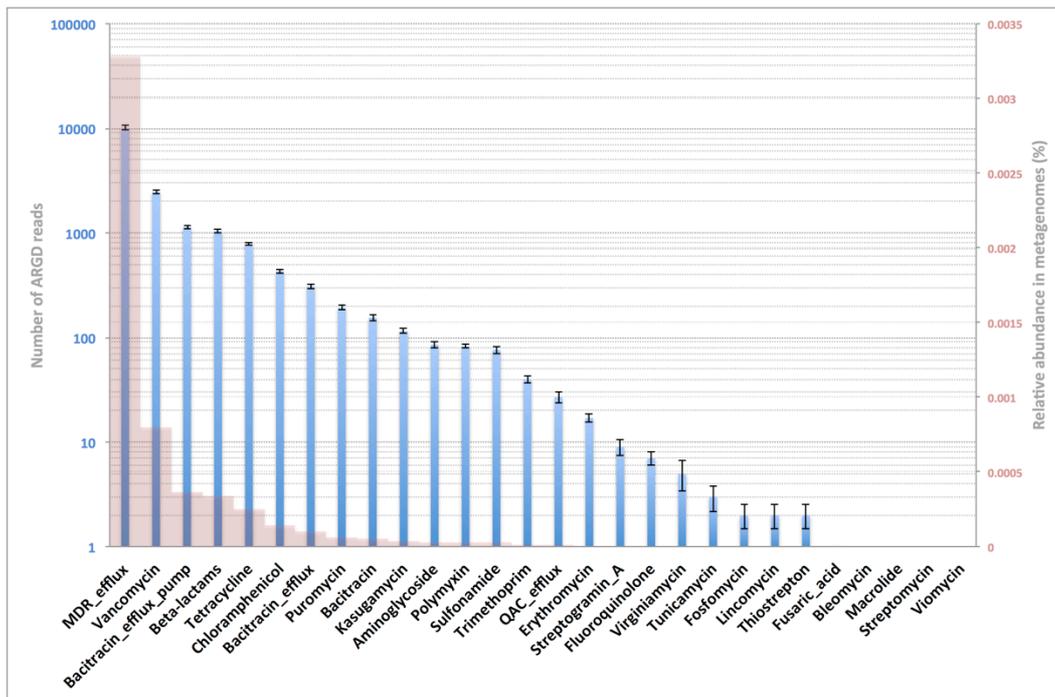


**Figure 5 - Relative abundance (A) and diversity (B) of antibiotic-resistance gene determinants in French Guiana and Rothamsted soils.**

Boxplots showing the relative abundance of reads annotated as ARGD (A), expressed as number of ARGD hits divided by total number of reads in the dataset and ARGD diversity (B), expressed as the percentage of 2999 unique sequences in ARDB database found in each dataset using three random sampling of 1000 sequences of three independent datasets per distance.

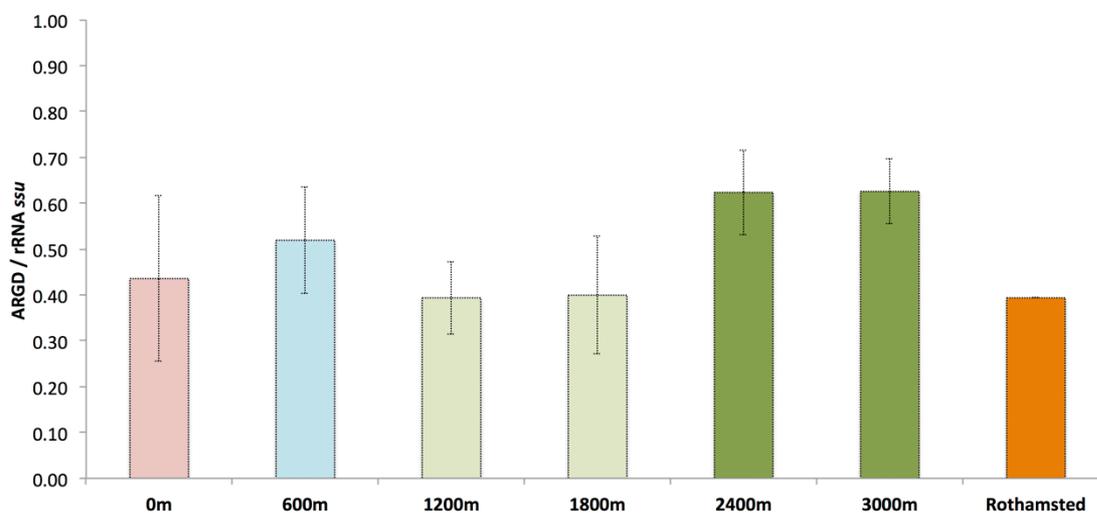
### Relative ARGD quantification in soil samples by qPCR.

Chosen antibiotic-resistance markers genes are quantified in soil samples using qPCR quantification. These genes (Table 4) are representative markers of widespread antibiotic-resistance genes in the clinical context and already detected in natural environments context (references in Table 4). French Guiana soil samples can be compared for relative ARGD markers gene copies detected after a normalization of each reaction with results obtained with the same assay using Rothamsted soil DNA as template according to Equation 1, as specified in experimental procedures. Results are therefore expressed as a quantification ratio between each sample and Rothamsted Park Grass soil DNA (Table 5, Figure 8). Results show that no significant distance effect is observed (null hypothesis of equal means between distances  $p\text{-value}>0.1$ ) for any ARGD markers, though, all genes but *tetA(H)* are detected by qPCR in French Guiana and in Rothamsted soil samples. The detected copy-numbers of class-B beta-lactamase gene *cfiA*, conferring resistance to carbapenems; cefotixin and cephamycins, is 1.5 to two fold more abundant in French Guiana soil samples than in Rothamsted soil ( $1.83\pm 0.34$  fold increase relative to Rothamsted *cfiA* abundance). Inversely, class 1 integrons tyrosine-recombinase gene *int1* gene is less abundant in French Guiana than in Rothamsted ( $0.64\pm 0.07$  fold). Beside these two markers, all other genes are not showing a significantly different abundance between Rothamsted, Park Grass soil and samples collected along the French Guiana soil line-transect (Figure 8).



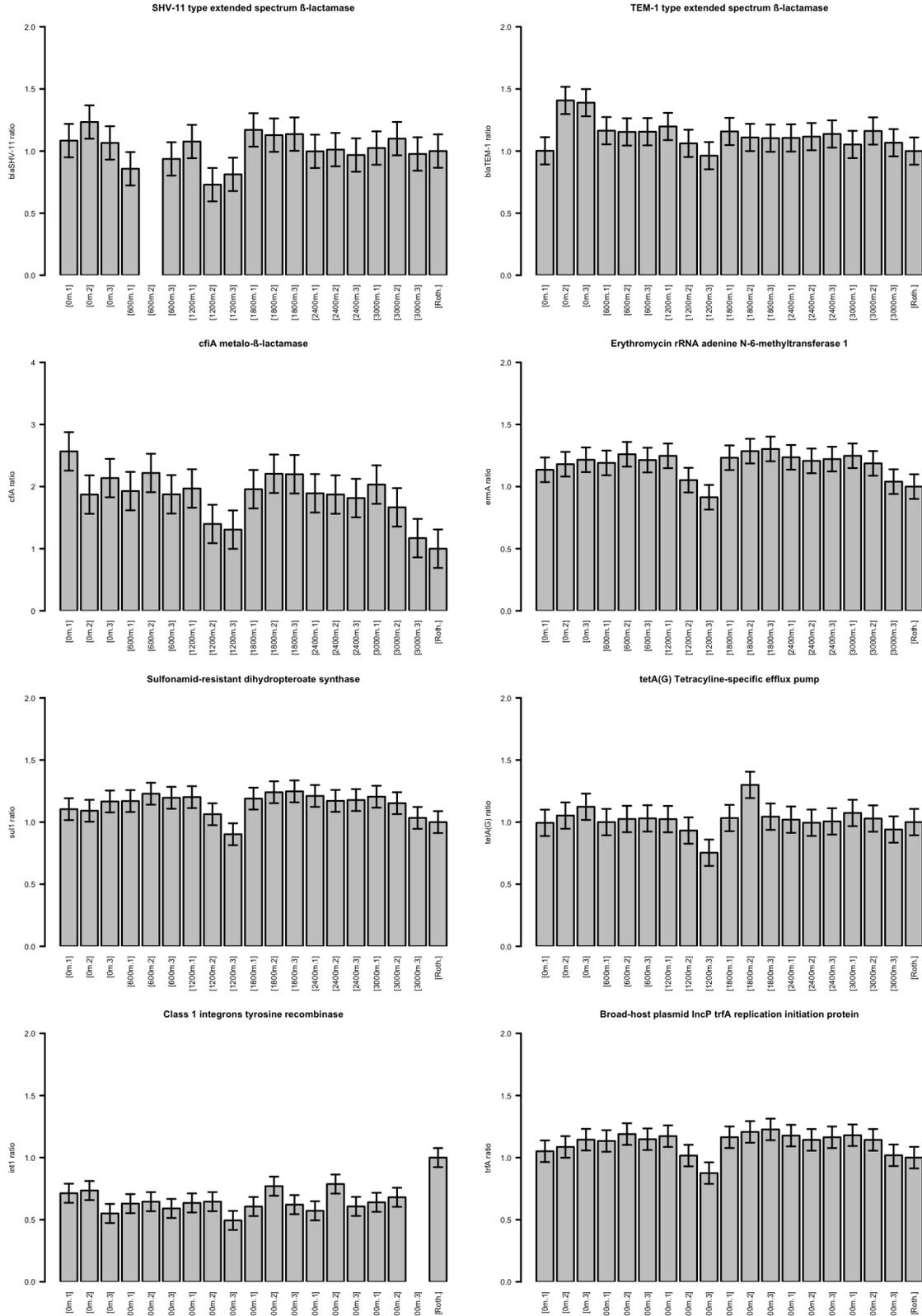
**Figure 6 - Global abundance of antibiotic-resistance genes found in French Guiana soil samples.**

The total number of ARGD-annotated reads found by BLASTx analysis are grouped per antibiotic molecule resistance-conferring properties and plotted by abundance rank order. In order to account for variations between individual sample distribution and uneven sequencing depth per sample, abundance counts (blue bars) are corrected by the total size of the dataset for all 18 soil metagenomic DNA shotgun sequencing datasets (n=3,085,014 reads) to express relative ARGD abundance per read in dataset (red bars).



**Figure 7 - Relative ARGD abundance detected in environmental shotgun sequencing normalized by found 16S rRNA ssu copies in the same datasets.**

Barplots showing the mean relative abundance of ARGD-annotated reads in each sample relatively to the number of SSU-annotated reads in shotgun soil metagenomic DNA sequencing results. Colour fillin indicates habitat type red: village; blue: Very degraded forest; light green: secondary forest; dark green: primary forest; orange: grassland (Rothamsted Park Grass).



**Figure 8 - Barplot showing qPCR relative quantification for antibiotic-resistance and mobile genetic elements marker genes for each individual metagenomic DNA sample.**

Barplots showing fold increase in qPCR relative quantification for each antibiotic-resistance and mobile genetic element marker genes assays relative to the same reaction using Rothamsted soil DNA as template and 16S rRNA quantification results as internal sample calibrator.

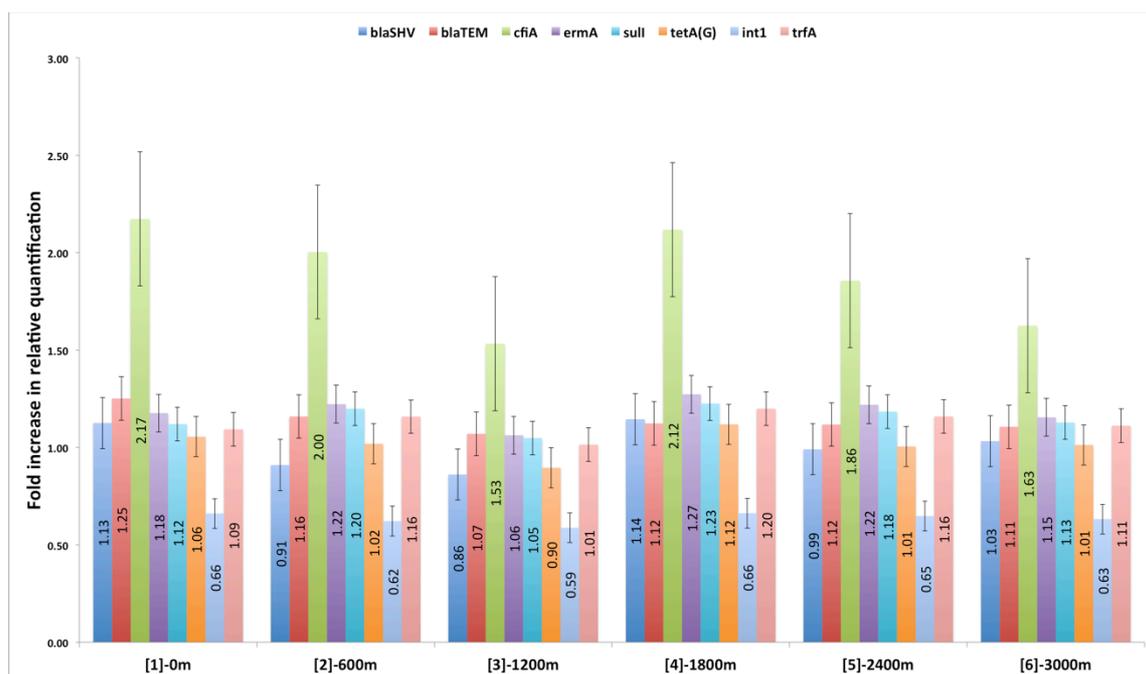
## Discussion.

Using metagenomic approaches, we investigated antibiotic resistance prevalence and diversity near a secluded French Guiana settlement: Trois-Sauts; where history of antibiotics delivered to the local population is known. In this context, we can question the impact of recent introduction of antibiotics molecule delivery to the local population on antibiotic resistance gene determinant abundance and diversity present in the local soil environment. Soil samples are collected in the vicinity of the Trois-Sauts village, municipality of Camopi, 97356, French Guyana, France during October 2010. Population records of the Camopi administrative area began in 1967 with only 276 individuals and 96 buildings identified and reach 1605 individuals and 418 buildings in 2009 (source: INSEE [53]). The Trois-Sauts hamlets population itself is reaching 597 individuals at the time of sampling in 2010, with 320 living in the direct vicinity of the local health-post and the soil line-transect (designated as "Village", Zidok hamlet of Trois-Sauts, Camopi municipality (Figure 1). The chosen sampling strategy is based on the fact that the local population relies essentially on the water hydrographic network for movements and all settlements or soil-impacting human activities (agriculture, buildings) are restricted to the immediate vicinity of the Oyapock River's banks (Figure 1). As described in experimental procedures, the soil line-transect follows a hunting-trail, used by the village hunters. Samples were collected a few meters away from the trail, where the soil was observed as not disturbed by its maintenance and usage.

	16S_rRNA	blaSHV	blaTEM	cfiA	ermA	sull	tetA(G)	int1	trfA
d.0m1	0.951244016	1.083993581	1.001941824	2.5666575	1.135153684	1.103952752	0.994092488	0.713182622	1.05125497
d.0m2	0.920936069	1.233753777	1.407631817	1.871731664	1.18015871	1.091168683	1.052582563	0.73501728	1.085851704
d.0m3	0.873834508	1.065532603	1.389234567	2.137000308	1.215751353	1.166036929	1.123251616	0.550183074	1.144381449
d.600m1	0.881837419	0.857462771	1.164052152	1.927333448	1.190655736	1.169342523	1.000079004	0.62943884	1.133995881
d.600m2	0.840690543	NA	1.154039743	2.218972724	1.26016614	1.228600776	1.024915447	0.645189362	1.189498334
d.600m3	0.870954846	0.936649853	1.155649237	1.874862355	1.213556519	1.195734369	1.029860505	0.590571707	1.148165149
d.1200m1	0.852667977	1.076719312	1.19791947	1.96937781	1.247502353	1.201230235	1.023910168	0.635066441	1.17278944
d.1200m2	0.983785706	0.729450934	1.06196128	1.398072026	1.051981813	1.06328711	0.932078274	0.645451862	1.016481531
d.1200m3	1.142224935	0.812299016	0.96279527	1.306829744	0.914271429	0.902177428	0.753189526	0.494395231	0.875484302
d.1800m1	0.859125563	1.170429021	1.157623167	1.95808815	1.232307906	1.189214542	1.032855006	0.606238427	1.163974211
d.1800m2	0.828620822	1.128201829	1.109458012	2.206693965	1.285324715	1.240259022	1.299338079	0.770160525	1.20682461
d.1800m3	0.814929091	1.136251289	1.103866493	2.198753476	1.302678455	1.247353799	1.043664776	0.621425184	1.227100629
d.2400m1	0.849457409	0.997757558	1.105811723	1.891976962	1.235672785	1.210475395	1.019902698	0.571822153	1.177222059
d.2400m2	0.874934025	1.011505945	1.116100488	1.872089783	1.206875883	1.170737607	0.994786583	0.787110263	1.142943321
d.2400m3	0.859395688	0.968043649	1.137640949	1.815266655	1.221261612	1.177594043	1.005698064	0.607059382	1.163608352
d.3000m1	0.847323749	1.024186778	1.052986419	2.032254252	1.24773255	1.204585039	1.073528418	0.640138549	1.180186441
d.3000m2	0.874934025	1.10009697	1.1616596	1.665842651	1.186829401	1.152184683	1.028864269	0.680810735	1.142943321
d.3000m3	0.981623184	0.976595591	1.067033864	1.170086457	1.039817741	1.033895522	0.940542169	NA	1.018720845
d.Roth	1	1	1	1	1	1	1	1	1

**Table 5 - Relative quantification results from qPCR antibiotic-resistance and mobile element marker genes normalized by 16S rRNA gene copy number in DNA sample.**

All quantitative reactions are normalized by 16S rRNA gene copy number detected in the same DNA sample as follow : gene ratio =  $\log[\text{target copies} / \text{rRNA copies}]$ . Quantification standard deviations is assessed using three independant replicates for each quantification reaction of the same DNA sample. Values are reported in Table S3 and displayed in Figure 6 as error-bars.



**Figure 9 - Boxplot showing qPCR relative quantification for antibiotic-resistance and mobile genetic element marker genes as a function of distance.**

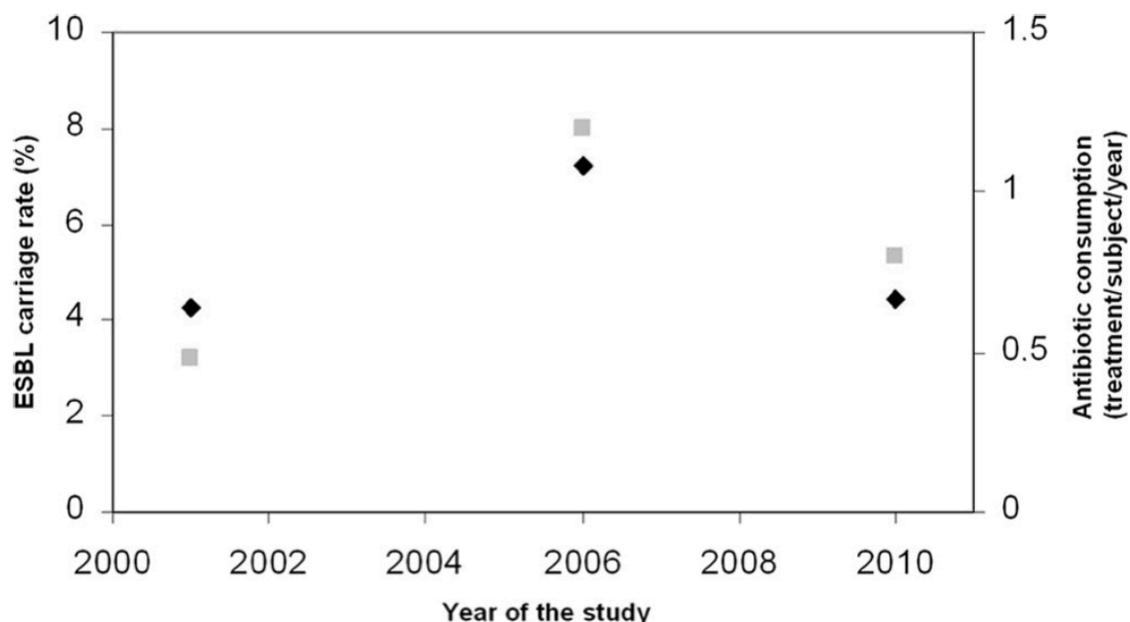
Barplots showing fold increase in qPCR relative quantification for each antibiotic-resistance and mobile genetic elements marker genes assays relative to the same reaction using Rothamsted soil DNA as template and 16S rRNA quantification results as internal sample calibrator as a function of distance.

Most prescribed drug in 2010 was penicillin with 0.43 treatments per subject per year (t/sub./yr.) and was significantly most prescribed than any other antibiotics, followed by metronidazole and macrolide with both 0.09 t/sub./yr. Total antibiotic consumption rose from 0.64 t/sub./yr in 2001 to 1.09 t/sub./yr in 2006 accounts and diminished in 2010 to 0.67 t/sub./yr. This evolution of total antibiotic consumption was shown to be correlated to an increase in extended-spectrum beta-lactamase producing *Enterobacteriaceae* fecal carriage by the local population tested cohort (Figure 10) [30]. In this study, authors identified on selective media the occurrence of total cultivable *Enterobacteriaceae* in rectal swab samples collected on a cohort of Trois-Sauts villagers, previously sampled in 2006 and 2001 [29-31, 54, 55]. Human settlement has strong impacts on the surrounding environment such as localized deforestation for agriculture, building houses [28, 56]. These activities are located on the banks of the Oyapack and revealed for example by an increase in *Escherichia coli* density in the river water downstream the village [54] and distinct microbial communities in village soil samples revealed in our study (Figure 4). Also, a strong correlation ( $R^2=0.87$ ) was reported between increase in antibiotic consumption and total fecal carriage of extended-spectrum beta-lactamases producing *Enterobacteriaceae* (Figure 10, [30]). Microbial antibiotic-resistance abundance was evaluated with similar culture-based methods in a broad range of surrounding environments: local human intestinal microbiome; wild mammals [29], river streams [54] and soil (this study) in a coordinated project funded by the French Agency for Food, Environmental and Occupational Health & Safety (ERAES2, ANSES [54]).

Antibiotic	2001	2006	2010
Penicillin	0.19	0.58	0.43
Cephalosporin	0	0.04	0.03
Macrolide	0.09	0.13	0.09
Metronidazole	0.29	0.22	0.09
Ofloxacin	0	0.01	0.01
Other	0.08	0.1	0.02
<b>Total</b>	<b>0.64</b>	<b>1.09</b>	<b>0.67</b>

**Table 6 – Comparison of antibiotic exposure among volunteers and villagers in 2001, 2006, and 2010 expressed in treatment per individual per year.**

Antibiotic consumed at Trois-Sauts are all delivered by a local health-post recoding every prescription. Most prescribed drugs per habitant are penicillin-derivatives. Penicillin antibiotics consumption increased by more than two fold between 2001 and 2006 and is diminishing in 2010. Metronidazole consumption is also reduced by two fold between 2006 and 2010 and was the most prescribed drug in 2001. Data from Woerther et al. (2013). Antimicrob Agents Chemother 57, 5060–5066.



**Figure 10 – Plot showing the relationship between resistant Escherichia coli carriage and total antibiotic consumption at Trois-Sauts.**

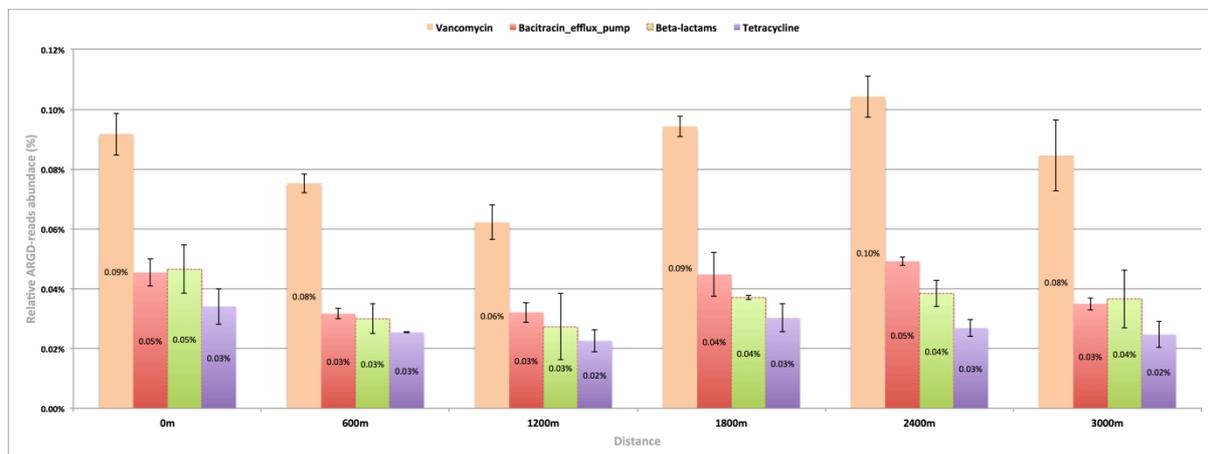
ESBL-E carriage rate in Wayampi volunteers (gray squares) and overall antibiotic exposure of the whole community (black diamonds) in 2001, 2006, and 2010. Linear regression ( $y = 8.675x + 1.396$ ;  $R^2 = 0.87$ ) and Pearson's correlation ( $P = 0.24$ ) were used to evaluate the evolution. Data from Woerther et al. (2013). Antimicrob Agents Chemother 57, 5060–5066.

In order to test if a similar impact in terms of antibiotic abundance and diversity was observable along a line-transect gradient going from the village to 3km deep in the forest (Figure 2), soil metagenomic DNA were sampled at different distances from the village. The soil is indeed a considerable reservoir of antibiotic producers and antibiotic resistant microorganisms (see Chapter 1). The observed increase in antibiotic-resistant bacteria found

in Trois-Sauts population does not solve the question of these antibiotic-resistance gene origins. Authors also observed that *Escherichia coli* strains collected from humans, human-associated animals (i.e.: dogs, chicken) and wild animals (i.e.: rodents, bats) belong to different genotypes and are distinct populations with different traits, such as presence of virulence factors or antibiotic-resistance genes on different types of plasmid replicons [29, 31]. These two observations of structured populations of related strains harboring multiple different genetic elements question the acquisition and dissemination of clinically problematic traits in human-associated microbial populations by means of horizontal gene transfers. How and from where human-associated bacterial antibiotic resistance genes are acquired? [2, 57] What is the role of the local environmental resistome reservoir of genetic diversity? How is it impacted by human activities or antibiotic consumption [3, 30, 58, 59]?

Soil samples were collected at a regular distance (600m), for 3km starting from the village and going into the primary forest to assess if a similar impact of a localized consumption of antibiotics can be observed in the soil resistome abundance and diversity and the potential relationship between human-associated microbiome and the local environment. Considering the difficulty to cultivate most microorganisms from environmental samples such as soil with conventional laboratory techniques and the need of holistic approaches in environmental microbial ecology [19, 21], a metagenomic approach was developed in order to 1) circumvent cultivation-related bias and 2) explore diversity and abundance of ARGs without *a priori* and limited sampling biases. In this respect, direct sequencing of soil metagenomic DNA extracted directly from soil samples was performed (shotgun sequencing) and sequencing reads were compared with a database of known antibiotic resistance related sequences (ARDB), as described in experimental procedures. In parallel, the same DNA solutions were used to amplify the V4-V6 region of

the 16S rRNA gene with "universal" primers targeting conserved regions for a more in-depth analysis of sample taxonomic composition. Random shotgun sequencing offers only a limited number of reads ( $1.19\% \pm 0.38$ ) that can be confidently associated with the 16S rRNA gene taxonomic markers (Table 3, Figure S2). Due to pyrosequencing technology, reads obtained are of  $344 \pm 34$  nucleotides length and span random region of the SSU gene. Sequence alignment with 16S rRNA database will therefore only occur on the read region spanning the 16S rRNA gene, lowering confidence in taxonomic affiliation as function of phylogenetic ranks depth: Kingdom > Phylum > Class > Order > Family > Genus > Species. Such a strategy is only applicable if species diversity is limited in the DNA sample submitted to random sequencing. In this case, sequence read redundancy in each dataset will be sufficient to allow sequence alignment, genome reconstruction and confident taxonomic affiliation. In an environment such as a forest soil where species richness



**Figure 11 - Relative abundance per distance of ARGD-annotated reads for four different antibiotics.**

The total number of ARGD-annotated reads found by BLASTx analysis are grouped per antibiotic molecule resistance-conferring properties for each distance. Standard deviation of the mean is calculated for each group of antibiotic molecule with the three soil samples per distance (18 French Guiana soil samples).



**Figure 12 - Projection of soil samples GPS coordinates on satellite imaging of Trois-Sauts, French Guiana with Bray-Curtis dissimilarity matrices hierarchical clustering of ARGD (A) and (B) Phylum-level sequence reads relative abundance found in each soil sample metagenomic DNA.**

GPS coordinates recorded at each sampling distance are projected on photography from LANDSAT-7 satellite imagery public data (landsat.usgs.gov). Despite unavoidable cloud cover above the Amazonian rainforest, the Oyapock river is clearly visible and sampling site location can be compared with the on-field hand-drawn map in Figure 1. Bray-Curtis dissimilarity matrices between samples are calculated for each abundance count datasets (i.e: ARGD resistance type annotated read counts and Phylum-level annotated read counts) on 1000 resampling of minimum sequence in samples (respectively:  $n=546$  and  $n=3438$ ; see Table ) and ordered as trees using UPGMA clustering. With both datasets, Village samples (1; red box) and forest soil samples (2; green box) are grouped separately. Tree significance test:  $p\text{-value}=0.088$  and  $0.178$  on 1000 replicates.

is estimated to vary between  $10^6$  to  $10^{10}$  microorganism species per gram intrinsic species diversity is preventing such approaches for complex ecosystem studies [15, 60]. A dual approach was used in this study: sample taxonomic composition was assessed by targeted 16S rRNA gene sequencing (amplicons sequencing) while ARGD abundance and diversity was calculated using direct-sequencing of soil metagenomic DNA. Differences observed between directed or shotgun metagenomic sequencing has been tested in other soil environments [61]. Both approaches provided similar results interpretation according to authors but with different trade-offs: amplicons sequencing approach was more sensitive and allowed taxonomic assignation on the same region for all sequences but skew evenness, due to the non-linearity of PCR amplification. Shotgun sequencing provided a more

unbiased representation of species evenness in DNA samples, not taking into account extraction-related biases; but taxonomic identification was less robust and only a fraction of reads could be used as taxonomic markers [62]. In another approach, Eisen *et al.* (2010) [63] simulated microbial communities *in vitro* to measure differences between known abundance of each genome in a mock community and results obtained by pyrosequencing directly after metagenomic DNA was extracted with different protocols. They showed that no sequence-based method represented true relative abundance of an *in vitro* simulated microbial community. Each DNA extraction protocol produced a specific community-composition profile in term of taxa relative abundance; however, these results were reproducible (same differences) when using the same protocol. Bearing in mind these limitations, a single protocol was used in this study for all metagenomic DNA extractions of soil samples, all treated independently. These replicates are essential to assess if observed differences between samples of different distances are significant using appropriate ecological theory statistics [51, 64].

While carefully considering all the above-cited limits, our results show clearly that ARGD are highly abundant and diverse in French Guiana soil samples (Figure 5) and falls into the abundance range of Rothamsted Park Grass soil. An average of  $0.59\% \pm 0.1$  ARGD-annotated reads per dataset (Figure 5A) was found in Guyana soil samples. ARGD diversity (Figure 5B) was calculated as the number of unique ARGD-annotated reads found in each dataset divided by the number ( $n=2999$ ) of unique antibiotic-resistance related amino-acid sequences in the Antibiotic Resistance Database (ARDB). No significant differences in ARGD diversity was observable (Figure 5B; Table 3) and was also in the range of Rothamsted Park Grass ( $9.97\% \pm 1.77$ ). Soil line-transect was actually not showing a gradient of increasing ARGD contamination from the village to the forest when comparing ARGD abundance

(Figure 5A) or diversity (Figure 5B) as a function of distance to the village. Nonetheless, ARGD-annotated reads were detected in every soil sample metagenomic DNA by shotgun sequencing, even when considering the relatively low sensitivity of such approaches [63]. A full Roche 454 Titanium pyrosequencing plate, as used in this study, typically yields  $10^6$  sequence reads with a length mode around 400 to 600bp per sequencing run (0.4 to 1 Gbp respectively) and orders of magnitude more throughput than with Sanger-based methods [65]. Though, when considering the overall soil diversity such big numbers are actually representing only a small proportion. When averaging di-nucleotide mass to 650 Daltons, a full Roche 454 Titanium plate represents roughly only 5.4-13 gram mass equivalent of actually sequenced data:  $[(10^6 \text{ reads} \times 500\text{pb} \times 650\text{g.mol}^{-1})/\text{NA}] = 5.4 \cdot 10^{-13} \text{ g}$ ; but most platforms sequencing processes require micrograms of DNA when considering sequencing library preparation, running controls etc. Therefore, any confidently annotated read sequenced at random in the overall pool of different soil metagenomic DNA samples indicates that the corresponding gene was actually very abundant in the metagenomes, including the ARGD-annotated reads in French Guiana soil samples in this study. When normalizing samples with the number of *ssu*-annotated reads in metagenomes, it seemed that the proportion of ARGD per *ssu* was correlated to landscape type (village; secondary, primary or degraded forest) rather than respective distance from the village; with a higher proportion of ARGD found in primary forest samples.

Our results obtained by metagenomic sequencing were confirmed by quantitative PCR using the same DNA samples as previously used for metagenomic sequencing. All qPCR assays targeting selected antibiotic resistance marker genes were positive (Table 4&5), but no significant gradient of increased ARGD copy number between distance samples was observed. Each quantification reaction was compared to the same assay performed using

Rothamsted soil metagenomic DNA as template to average variations in each French Guiana soil sample metagenomic DNA quantification for better inter-samples and inter-qPCR marker quantification comparisons (Figure 8). Similarly, like for metagenomic sequencing, no significant gradient effect was observed, two out of three samples from the village (0m.2 and 0m.3) showing a significantly higher ratio of *bla*<sub>TEM-1</sub> per *ssu* copy than any other sample but the same trend was not observed in sample 0m.1. Quantitative PCR results were presented as relative quantification normalized by the same reaction performed in the Rothamsted soil, using 16S rRNA gene quantification as inter-samples internal calibrator. Such qPCR assays were very sensitive to variations likely to be introduced in environmental molecular biology experiments such as: DNA extraction biases, PCR inhibitors or uneven primers specificity with different templates. Such biases were unavoidable when objective was to compare samples from various locations, where all things were not equal given the complexity of the soil matrix, in respect to qPCR assays, orders of magnitude more sensitive than random environmental shotgun sequencing [66, 67]. Local variations between soil samples from the same distance group were likely to be reflected by qPCR assays but not observed in our shotgun sequencing of the same soil samples (Figure 7).

While not observing an obvious gradient effect, replicating soil sampling is nonetheless allowing us to confidently assert that the soil bacterial community near the village (distance 0m) was significantly different from the forest soil samples. This is true both in terms of microbial community taxonomic composition (Figure 3, 4 & 12) and ARGD composition, as shown by our annotated-read count abundance (Figure 7 & 11) and Bray-Curtis dissimilarity tree clustering (Figure 12) based on two different sequencing results (i.e. amplicons & shotgun sequencing, respectively). Interestingly, forest soil samples from distances 2400m and 3000m (Primary forest) are showing an increased abundance of ARGD-

annotated reads per SSU in comparison to all other samples (null hypothesis: p-value=0.001; non-parametric bilateral T-test); inversely, samples from distance 1200m and 1800m (Secondary forest) are showing a relatively less abundant share of ARGD annotated reads (p-value=0.01). At a finer scale, however, shotgun metagenomic sequencing of French Guiana soil DNA collected near the village showed increased relative abundance of ARGD annotated reads related to beta-lactams resistance (Figure 11). When considering only two out of three of these village samples, this difference was significant (p-value=0.001) and the same trend was shown by qPCR where a relatively higher abundance of *bla*<sub>TEM-1</sub> gene relative to ssu copies in the same two samples (0m.2 and 0m.3; Figure 8) was found, comforting metagenomic read annotation results.

When considering the complex relationships between soil microorganisms and their ecosystem, involving a large range of different of biotic and abiotic factors and defining as many micro-habitats and different ecological niches, it would have been extremely striking to not separate samples on the basis of very obvious environmental factors, such as land cover. Forest samples with a higher tree density than degraded and villages' samples are showing increased *Acidobacteria* relative abundance, which is a widely documented observation in soil microbial ecology [68]. In a previous investigation by using 16S rRNA gene clone libraries of Amazonian forest soil metagenomic microbial DNA, authors also observed a significant shift in bacterial community following transformation of a forest in pasture [69]. They found that the relative abundance of Bacilli and *Clostridium* spp. (Phylum: Firmicutes); *Cytophaga*, *Flexibacteria* and *Bacteroides* spp. (Phylum: Bacteroidetes) increased in soil following deforestation, our own results confirming their data (Figure 3B). These two phyla are prominent in the human-gut microbiota as revealed in recent deep-sequencing results of humans' microbiome [70, 71]; but are also commonly found in

environmental samples such as soils [36, 72] or oceans [73-76]. In our results, it was not possible to correlate an increase of these specific taxa relative abundance and antibiotic-resistance gene abundance and diversity, despite the same clustering of samples was achieved with both of resistance and taxonomy annotation.

In conclusion, our results are reaffirming the major role of soil as a reservoir of antibiotic-resistance genes determinants, as revealed by shotgun metagenomic. Our analysis did not show that a localized human antibiotic consumption was having a significant impact on the abundance or diversity of antibiotic resistance genes in DNA extracted from soil collected at varying distance from the source of prescription and population activities (Figure 7). The impact of deforestation, defined by differences in landscape type between samples, was however significant in terms of different taxonomic composition between samples from varying locations. We also showed that ARGD distribution was not random but depended on local factors, such as intrinsic taxonomic composition (Figure 12) or landscape type (Table 1), rather than raw distance from the local source of antibiotic molecules. It is, however, uneasy in the light of our results to infer more on the complex microbial ecology involved in ARGD dissemination between telluric bacteria and pathogens, despite the significant impact on the soil taxonomic composition in the vicinity of the village following anthropization. This impact on taxonomic distribution in the village samples and the augmentation of beta-lactams resistance observed both by shotgun metagenomic sequencing and qPCR in soil samples near the village is of concern when observing emergence and dissemination of antibiotic-resistant commensal bacteria among villagers following an increase in antibiotic selective-pressure; notably when most of such resistance genes were identified on plasmid replicons [30, 31]. Observations on the human-impact on the mobile DNA pool were reported in many studies, as reviewed elsewhere [59], and

illustrated the need of a careful consideration of the soil environment in dissemination of antibiotic resistance [3, 27]. Studies focusing on *in situ* mechanisms of horizontal gene transfer in the environment [77, 78] coupled with a systematic approach involving carefully planned experimental-design, involving replicated time-series and spatial sampling and the always evolving sequencing technologies (*i.e.* single-molecule sequencing [79], higher throughput, lower DNA requirements, etc...) will hopefully shed new light on the complex systems biology leading from the reservoir of environmental diversity to problematic pathogen emergence and increasing antibiotic-resistance in the clinic.

## References

1. Nesme, J., Cécillon, S., Delmont, T. O., Monier, J.-M., Vogel, T. M., and Simonet, P. (2014). Large-Scale Metagenomic-Based Study of Antibiotic Resistance in the Environment. *Curr Biol*.
2. Forsberg, K. J., Reyes, A., Wang, B., Selleck, E. M., Sommer, M. O. A., and Dantas, G. (2012). The Shared Antibiotic Resistome of Soil Bacteria and Human Pathogens. *Science* 337, 1107–1111.
3. Martinez, J. L. (2008). Antibiotics and Antibiotic Resistance Genes in Natural Environments. *Science* 321, 365–367.
4. D'Costa, V. M., Griffiths, E., and Wright, G. D. (2007). Expanding the soil antibiotic resistome: exploring environmental diversity. *Curr Opin Microbiol* 10, 481–489.
5. Heuer, H., Schmitt, H., and Smalla, K. (2011). Antibiotic resistance gene spread due to manure application on agricultural fields. *Curr Opin Microbiol* 14, 236–243.
6. Bérdy, J. (2005). Bioactive microbial metabolites. *J. Antibiot.* 58, 1–26.
7. Smith, T. (2000). Antibiotics from soil bacteria. *Nat Struct Biol* 7, 189–190.
8. Walsh, F., and Duffy, B. (2013). The Culturable Soil Antibiotic Resistome: A Community of Multi-Drug Resistant Bacteria. *PLoS ONE* 8, e65567.
9. Allen, H. K., Moe, L. A., Rodbumrer, J., Gaarder, A., and Handelsman, J. (2009). Functional metagenomics reveals diverse beta-lactamases in a remote Alaskan soil. *ISME J* 3, 243–251.

10. Torres-Cortés, G., Millán, V., Ramírez-Saad, H. C., Nisa-Martínez, R., Toro, N., and Martínez-Abarca, F. (2011). Characterization of novel antibiotic resistance genes identified by functional metagenomics on soil samples. *Environ Microbiol* *13*, 1101–1114.
11. Davies, J. (2011). How to discover new antibiotics: harvesting the parvome. *Curr Opin Chem Biol* *15*, 5–10.
12. Pidot, S. J., Coyne, S., Kloss, F., and Hertweck, C. (2014). Antibiotics from neglected bacterial sources. *International Journal of Medical Microbiology* *304*, 14–22.
13. Stackebrandt, E., Frederiksen, W., Garrity, G. M., Grimont, P. A. D., Kämpfer, P., Maiden, M. C. J., Nesme, X., Rosselló-Mora, R., Swings, J., Trüper, H. G., et al. (2002). Report of the ad hoc committee for the re-evaluation of the species definition in bacteriology. *Int J Syst Evol Microbiol* *52*, 1043–1047.
14. Zhang, J., Kapli, P., Pavlidis, P., and Stamatakis, A. (2013). A general species delimitation method with applications to phylogenetic placements. *Bioinformatics*.
15. Torsvik, V. (2011). Prokaryotic Diversity--Magnitude, Dynamics, and Controlling Factors. *Science* *296*, 1064–1066.
16. Whitman, W. B., Coleman, D. C., and Wiebe, W. J. (1998). Prokaryotes: the unseen majority. *Proc Natl Acad Sci USA* *95*, 6578–6583.
17. Monier, J.-M., Demanèche, S., Delmont, T. O., Mathieu, A., Vogel, T. M., and Simonet, P. (2011). Metagenomic exploration of antibiotic resistance in soil. *Curr Opin Microbiol* *14*, 229–235.
18. Schloss, P. D., and Handelsman, J. (2003). Biotechnological prospects from metagenomics. *Curr Opin Biotechnol* *14*, 303–310.
19. Schloss, P. D., and Handelsman, J. (2005). Metagenomics for studying unculturable microorganisms: cutting the Gordian knot. *Genome Biol* *6*, 229.
20. Riesenfeld, C. S., Goodman, R. M., and Handelsman, J. (2004). Uncultured soil bacteria are a reservoir of new antibiotic resistance genes. *Environ Microbiol* *6*, 981–989.
21. Rappé, M. S., and Giovannoni, S. J. (2003). The uncultured microbial majority. *Annu Rev Microbiol* *57*, 369–394.
22. Thompson, J. M., Gündoğdu, A., Stratton, H. M., and Katouli, M. (2013). Antibiotic resistant *Staphylococcus aureus* in hospital wastewaters and sewage treatment plants with special reference to methicillin-resistant *Staphylococcus aureus* (MRSA). *J Appl Microbiol* *114*, 44–54.
23. Demanèche, S., Sanguin, H., Poté, J., Navarro, E., Bernillon, D., Mavingui, P., Wildi, W., Vogel, T. M., and Simonet, P. (2008). Antibiotic-resistant soil bacteria in transgenic plant fields. *Proc Natl Acad Sci USA* *105*, 3957–3962.

24. Zhu, Y.-G., Johnson, T. A., Su, J.-Q., Qiao, M., Guo, G.-X., Stedtfeld, R. D., Hashsham, S. A., and Tiedje, J. M. (2013). Diverse and abundant antibiotic resistance genes in Chinese swine farms. *Proc Natl Acad Sci USA*.
25. Heuer, H., Kopmann, C., Binh, C. T. T., Top, E. M., and Smalla, K. (2009). Spreading antibiotic resistance through spread manure: characteristics of a novel plasmid type with low %G+C content. *Environ Microbiol* *11*, 937–949.
26. Looft, T., Johnson, T. A., Allen, H. K., Bayles, D. O., Alt, D. P., Stedtfeld, R. D., Sul, W. J., Stedtfeld, T. M., Chai, B., Cole, J. R., et al. (2012). In-feed antibiotic effects on the swine intestinal microbiome. *Proc Natl Acad Sci USA* *109*, 1691–1696.
27. Perry, J. A., and Wright, G. D. (2013). The antibiotic resistance “mobilome”: searching for the link between environment and clinic. *Front Microbiol* *4*, 138.
28. Damien, D., Tritsch, I., and Grenand, P. (2012). Construction and territorial restructuring of the Wayãpi and Teko Indians from the municipality of Camopi, French Guiana. *confins*.
29. Lescat, M., Clermont, O., Woerther, P.-L., Glodt, J., Dion, S., Skurnik, D., Djossou, F., Dupont, C., Perroz, G., Picard, B., et al. (2013). Commensal *Escherichia coli* strains in Guiana reveal a high genetic diversity with host-dependant population structure. *Env Mic Reports* *5*, 49–57. Available at: <http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=23757130&retmode=ref&cmd=prlinks>.
30. Woerther, P.-L., Angebault, C., Jacquier, H., Clermont, O., Mniai, El, A., Moreau, B., Djossou, F., Peroz, G., Catzeflis, F., Denamur, E., et al. (2013). Characterization of fecal extended-spectrum- $\beta$ -lactamase-producing *Escherichia coli* in a remote community during a long time period. *Antimicrob Agents Chemother* *57*, 5060–5066.
31. Woerther, P.-L., Angebault, C., Lescat, M., Ruppé, E., Skurnik, D., Mniai, A. E., Clermont, O., Jacquier, H., Costa, A. D., Renard, M., et al. (2010). Emergence and dissemination of extended-spectrum beta-lactamase-producing *Escherichia coli* in the community: lessons from the study of a remote and controlled population. *J. Infect. Dis.* *202*, 515–523.
32. Benveniste, R., and Davies, J. (1973). Aminoglycoside antibiotic-inactivating enzymes in actinomycetes similar to those present in clinical isolates of antibiotic-resistant bacteria. *Proc Natl Acad Sci USA* *70*, 2276–2280.
33. Boucher, H. W., Talbot, G. H., Bradley, J. S., Edwards, J. E., Gilbert, D., Rice, L. B., Scheld, M., Spellberg, B., and Bartlett, J. (2009). Bad Bugs, No Drugs: No ESKAPE! An Update from the Infectious Diseases Society of America. *Clin Inf Dis*.
34. de l'écologie et du développement durable, M. (2007). Décret n° 2007-266 du 27 février 2007 créant le parc national dénommé « Parc amazonien de Guyane ». *Journal Officiel de la République Française*, 1–7.

35. Silvertown, J. (1980). The dynamics of a grassland ecosystem: botanical equilibrium in the Park Grass Experiment. *Journal of Applied Ecology*.
36. Delmont, T. O., Prestat, E., Keegan, K. P., Faubladiere, M., Robe, P., Clark, I. M., Pelletier, E., Hirsch, P. R., Meyer, F., Gilbert, J. A., et al. (2012). Structure, fluctuation and magnitude of a natural grassland soil metagenome. *ISME J* 6, 1677–1687.
37. Delmont, T. O., Robe, P., Cécillon, S., Clark, I. M., Constancias, F., Simonet, P., Hirsch, P. R., and Vogel, T. M. (2011). Accessing the soil metagenome for studies of microbial diversity. *Appl Environ Microbiol* 77, 1315–1324.
38. Eisen, J. A. (2007). Environmental Shotgun Sequencing: Its Potential and Challenges for Studying the Hidden World of Microbes. *PLoS Biol* 5, e82.
39. Parameswaran, P., Jalili, R., Tao, L., Shokralla, S., Gharizadeh, B., Ronaghi, M., and Fire, A. Z. (2007). A pyrosequencing-tailored nucleotide barcode design unveils opportunities for large-scale sample multiplexing. *Nucleic Acids Res* 35, e130–e130.
40. Klindworth, A., Pruesse, E., Schweer, T., Peplies, J., Quast, C., Horn, M., and Glöckner, F. O. (2013). Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res* 41, e1.
41. Meyer, F., Paarmann, D., D'Souza, M., Olson, R., Glass, E. M., Kubal, M., Paczian, T., Rodriguez, A., Stevens, R., Wilke, A., et al. (2008). The metagenomics RAST server - a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics* 9, 386.
42. Overbeek, R. A., Begley, T., Butler, R. M., Choudhuri, J. V., Chuang, H.-Y., Cohoon, M., de Crécy-Lagard, V., Diaz, N., Disz, T., Edwards, R. A., et al. (2005). The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. *Nucleic Acids Res* 33, 5691–5702.
43. Niu, B., Fu, L., Sun, S., and Li, W. (2010). Artificial and natural duplicates in pyrosequencing reads of metagenomic data. *BMC Bioinformatics* 11, 187.
44. Liu, B., and Pop, M. (2009). ARDB--Antibiotic Resistance Genes Database. *Nucleic Acids Res* 37, D443–7.
45. David, M. M., Sapkota, A. R., Simonet, P., and Vogel, T. M. (2008). A novel and rapid method for synthesizing positive controls and standards for quantitative PCR. *J Microbiol Methods* 73, 73–77.
46. Bahl, M. I., Hansen, L. H., Goesmann, A., and Sørensen, S. J. (2007). The multiple antibiotic resistance IncP-1 plasmid pKJK5 isolated from a soil environment is phylogenetically divergent from members of the previously established alpha, beta and delta sub-groups. *Plasmid* 58, 31–43.
47. Dixon, P. (2003). VEGAN, a package of R functions for community ecology. *J Veg Sci* 14, 927–930.

48. Thioulouse, J., Chessel, D., Dole dec, S., and Olivier, J. M. (1997). ADE-4: a multivariate analysis and graphical display software. *Statistics and computing* 7, 75–83.
49. Suzuki, R., and Shimodaira, H. (2006). Pvcust: an R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics* 22, 1540–1542.
50. Parks, D. H., Mankowski, T., Zangoeei, S., Porter, M. S., Armanini, D. G., Baird, D. J., Langille, M. G. I., and Beiko, R. G. (2013). GenGIS 2: geospatial analysis of traditional and genetic biodiversity, with new gradient algorithms and an extensible plugin framework. *PLoS ONE* 8, e69885.
51. Whittaker, R. H. (1972). Evolution and measurement of species diversity. *Taxon*.
52. Lozupone, C., Lladser, M. E., Knights, D., Stombaugh, J., and Knight, R. (2011). UniFrac: an effective distance metric for microbial community comparison. *ISME J* 5, 169–172.
53. Insee (2012). Camopi (97356 - Commune) - Évolution et structure de la population. INSEE, 1–16.
54. Angebault, C., Ruimy, R., Woerther, P.-L., Djossou, F., Denamur, E., Clermont, O., Vandenesch, F., Catzeflis, F., Rousset, F., Guillemot, D., et al. (2010). Ecologie de la résistance aux antibiotiques d'*Escherichia coli* et *Staphylococcus aureus* dans les flores commensales de l'homme et des animaux en milieu naturel : période 2 (ERAES2). *Les Rencontres Scientifiques de l'AFFSET*, 1–1.
55. Angebault, C., Djossou, F., Abélanet, S., Permal, E., Ben Soltana, M., Diancourt, L., Bouchier, C., Woerther, P.-L., Catzeflis, F., Andremont, A., et al. (2013). *Candida albicans* is not always the preferential yeast colonizing humans: a study in Wayampi Amerindians. *J. Infect. Dis.* 208, 1705–1716.
56. Béliveau, A., Davidson, R., and Lucotte, M. (2008). Dynamiques des propriétés des sols de fermes familiales de la région du Tapajos (Amazonie brésilienne) après la première année de culture sur brûlis. *Université du Québec à Montréal*.
57. Wellington, E. M. H., Boxall, A. B., Cross, P., Feil, E. J., Gaze, W. H., Hawkey, P. M., Johnson-Rollings, A. S., Jones, D. L., Lee, N. M., Otten, W., et al. (2013). The role of the natural environment in the emergence of antibiotic resistance in gram-negative bacteria. *Lancet Infect Dis* 13, 155–165.
58. Nardelli, M., Scalzo, P. M., Ramírez, M. S., Quiroga, M. P., Cassini, M. H., and Centrón, D. (2012). Class 1 integrons in environments with different degrees of urbanization. *PLoS ONE* 7, e39223.
59. Gaze, W. H., Krone, S. M., Larsson, D. G. J., Li, X.-Z., Robinson, J. A., Simonet, P., Smalla, K., Timinouni, M., Topp, E., Wellington, E. M., et al. (2013). Influence of humans on evolution and mobilization of environmental antibiotic resistome. *Emerging Infect Dis* 19.

60. Torsvik, V., Daae, F. L., Sandaa, R. A., and Øvreås, L. (1998). Novel techniques for analysing microbial diversity in natural and perturbed environments. *J Biotechnol.*
61. Steven, B., Gallegos-Graves, L. V., Starckenburg, S. R., Chain, P. S., and Kuske, C. R. (2012). Targeted and shotgun metagenomic approaches provide different descriptions of dryland soil microbial communities in a manipulated field study. *Env Mic Reports* 4, 248–256.
62. Sunagawa, S., Mende, D. R., Zeller, G., Izquierdo-Carrasco, F., Berger, S. A., Kultima, J. R., Coelho, L. P., Arumugam, M., Tap, J., Nielsen, H. B., et al. (2013). Metagenomic species profiling using universal phylogenetic marker genes. *Nat Methods* 10, 1196–1199.
63. Morgan, J. L., Darling, A. E., and Eisen, J. A. (2010). PLOS ONE: Metagenomic Sequencing of an In Vitro-Simulated Microbial Community. *PLoS ONE*.
64. Zinger, L., Gobet, A., and Pommier, T. (2012). Two decades of describing the unseen majority of aquatic microbial diversity. *Molecular Ecology* 21, 1878–1896.
65. Mardis, E. R. (2008). Next-Generation DNA Sequencing Methods. *Annu. Rev. Genom. Human Genet.* 9, 387–402.
66. Schriewer, A., Wehlmann, A., and Wuertz, S. (2011). Improving qPCR efficiency in environmental samples by selective removal of humic acids with DAX-8. *J Microbiol Methods* 85, 16–21.
67. Smith, C. J., and Osborn, A. M. (2009). Advantages and limitations of quantitative PCR (Q-PCR)-based approaches in microbial ecology. *FEMS Microbiol Ecol* 67, 6–20.
68. Rosch, C., Mergel, A., and Bothe, H. (2002). Biodiversity of Denitrifying and Dinitrogen-Fixing Bacteria in an Acid Forest Soil. *Appl Environ Microbiol* 68, 3818–3829.
69. Borneman, J., and Triplett, E. W. (1997). Molecular microbial diversity in soils from eastern Amazonia: evidence for unusual microorganisms and microbial population shifts associated with deforestation. *Appl Environ Microbiol* 63, 2647–2653.
70. Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K. S., Manichanh, C., Nielsen, T., Pons, N., Levenez, F., Yamada, T., et al. (2010). A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* 464, 59–65.
71. Turnbaugh, P. J., Hamady, M., Yatsunenko, T., Cantarel, B. L., Duncan, A., Ley, R. E., Sogin, M. L., Jones, W. J., Roe, B. A., Affourtit, J. P., et al. (2009). A core gut microbiome in obese and lean twins. *Nature* 457, 480–484.
72. Jacquiod, S., Franqueville, L., Cécillon, S., Vogel, T. M., and Simonet, P. (2013). Soil bacterial community shifts after chitin enrichment: an integrative metagenomic approach. *PLoS ONE* 8, e79699.
73. Gilbert, J. A., Field, D., Huang, Y., Edwards, R., Li, W., Gilna, P., and Joint, I. (2008).

- Detection of large numbers of novel sequences in the metatranscriptomes of complex marine microbial communities. *PLoS ONE* 3, e3042.
74. Agogué, H., Lamy, D., Neal, P. R., Sogin, M. L., and Herndl, G. J. (2011). Water mass-specificity of bacterial communities in the North Atlantic revealed by massively parallel sequencing. *Molecular Ecology* 20, 258–274.
  75. Sogin, M. L., Morrison, H. G., Huber, J. A., Mark Welch, D., Huse, S. M., Neal, P. R., Arrieta, J. M., and Herndl, G. J. (2006). Microbial diversity in the deep sea and the underexplored “rare biosphere.” *Proc Natl Acad Sci USA* 103, 12115–12120.
  76. Galand, P. E., Casamayor, E. O., Kirchman, D. L., and Lovejoy, C. (2009). Ecology of the rare microbial biosphere of the Arctic Ocean. *Proc Natl Acad Sci USA* 106, 22427–22432.
  77. Thomas, C. M., and Nielsen, K. M. (2005). Mechanisms of, and barriers to, horizontal gene transfer between bacteria. *Nat Rev Microbiol* 3, 711–721.
  78. Nielsen, K. M., and Townsend, J. P. (2004). Monitoring and modeling horizontal gene transfer. *Nat Biotechnol* 22, 1110–1114.
  79. Schneider, G. F., and Dekker, C. (2012). DNA sequencing with nanopores : Nature Biotechnology : Nature Publishing Group. *Nat Biotechnol*.

# Chapitre 5

## **Conclusions sur les résultats apportés par les approches environnementales évaluant l'abondance, la diversité et le potentiel de dissémination des ARGD dans le sol et futures perspectives pour l'étude du résistome du sol.**

### **Résumé.**

Le problème critique de santé publique représenté par l'acquisition de phénotypes de résistance par des bactéries pathogènes rendant ainsi les traitements antibiotiques inefficaces doit être évalué dans une perspective écologique globale. En effet, si de nombreux microorganismes producteurs, mais aussi de nombreuses bactéries saprophytes résistantes à des antibiotiques sont abondantes dans le sol, ils constituent alors un réservoir de mécanismes de résistance, transférables aux bactéries commensales et/ou aux pathogènes, et sélectionnés en réponse aux traitements antibiotiques en milieu hospitalier. Ce contexte théorique est communément admis mais les approches expérimentales peinent à fournir des résultats tester rigoureusement ces hypothèses. Le but de ces travaux était d'abord de développer et utiliser une méthode permettant d'établir avec le minimum de biais, en l'état actuel des connaissances et des récents progrès technologiques, la prévalence et la diversité des gènes de résistance à des antibiotiques dans l'environnement. Cette méthode bioinformatique à l'avantage de permettre le ré-examen d'une grande quantité de données de séquences disponibles publiquement dans les bases de données internationales. Les résultats obtenus confirment l'importance du sol en tant que réservoir de gènes de résistance à des antibiotiques et ont conduit à l'élaboration d'une stratégie double pour l'examen de l'abondance et la dynamique des gènes de résistance à des

antibiotiques dans le sol. D'une part, avec un étude *in vitro* utilisant des microcosmes d'un sol de référence amendés avec des concentrations connues soit de molécules antibiotiques pures, soit d'effluents provenant d'animaux d'élevages traités ou non avec des antibiotiques a d'abord été réalisée. Cette expérience nous a permis de mettre en évidence un impact contrasté entre la pression de sélection des molécules antibiotiques d'une part, relativement faible, et celle des communautés bactériennes provenant des effluents de ferme, très important, sur la structure des communautés bactériennes d'un même sol, au cours des 6 mois d'incubation en microcosmes. Cet effet n'est cependant pas détecté avec les molécules antibiotiques pures et nos résultats questionnent sérieusement la forme de la pression de sélection qui est exercée dans le sol, qui n'est vraisemblablement pas liée directement aux molécules antibiotiques, aux concentrations testées. Afin d'explorer cette situation *in situ*, le résistome du sol d'un site particulièrement intéressant dans le contexte de l'études d'impact de la consommation humaine de molécules antibiotiques thérapeutiques a été étudié. Le village de Trois-Sauts, en forêt Amazonienne (Guyane Française) est très isolé et la consommation antibiotique de la population est très récente et précisément répertoriée. Des échantillons de sols ont été collectés dans et à proximité du village, à intervalle régulier, le long d'un transect. Les mêmes méthodes d'analyse (qPCR, séquençages métagénomique) ont été utilisées et permettent de mettre en évidence un effet manifeste de l'urbanisation sur la communauté bactérienne du sol et une distribution caractéristique des gènes de résistance à des antibiotiques dans le sol du village, en comparaison avec celui de la forêt. L'impact sur les communautés microbiennes du sol, soit du fait de la fertilisation du sol par des matières fécales, soit de la colonisation humaine d'un habitat, est très marqué dans nos deux approches expérimentales. En revanche, l'effet de sélection des molécules antibiotiques proprement dites n'est lui pas significatif. Ces résultats suggèrent que la sélection naturelle agit très différemment, soit sur le phénotype de résistance, soit sur son hôte, en fonction de l'environnement: sol ou bactéries cliniques et commensales, soumises à des traitements antibiotiques et souligne le rôle important des éléments génétiques mobiles dans ces dynamiques d'adaptation. Les études futures sur l'écologie microbienne de la résistance antibiotique devront tenter de relier plus précisément qu'actuellement les données obtenues sur la taxonomie et le type de réplicons et les données codant des traits du type résistance antibiotique, ce qui est aujourd'hui très difficile car la taille des lectures de séquences ne permet pas d'avoir cette double

information sur une seule lecture. Malgré les débits élevés de séquençage actuels qui permettent de tester de plus en plus d'hypothèses écologiques à l'aide des données massives de séquençage et une planification expérimentale rigoureuse, ceux-ci sont encore très loin de pouvoir capturer l'ensemble de la diversité génétique d'un sol. Il est cependant possible de réaliser ce lien entre taxonomie, ou type de réplicon, et phénotype de résistance antibiotique. Ceci à condition de pouvoir réaliser un assemblage des séquences, ce qui reste actuellement un défi dans des environnements où la diversité des génomes est très élevée. Il est aussi possible d'imaginer des stratégies qui dès l'extraction d'ADN vont cibler directement et uniquement une fraction de la diversité, biaisant de fait les résultats si tant est que l'on veuille une approche globale systématique, mais permettant ainsi de réduire la diversité observée (*evenness*).

Des détails sont apportés dans la version anglaise du présent chapitre et explicite les perspectives expérimentales développer au laboratoire pour répondre à ces questions et cibler le mobilome.



# Chapter 5

## **Conclusions on experimental approaches results evaluating ARGD abundance, diversity and dissemination of antibiotic-resistance and future perspectives for soil resistome studies.**

The critical threat of antibiotic-resistance phenotype acquisition and dissemination by pathogenic bacteria for human healthcare need to be considered at a global environmental level, taking into account the complex ecology of microorganism's interactions. Many antibiotic producers, but also antibiotic-resistant saprophytic bacteria are found in the environment, and more especially in soil, since most antibiotics molecules are isolated from soil microorganisms. Soil is therefore considered now as a major reservoir of resistance genes found in pathogens, readily transferable to pathogens and/or commensal bacteria and selected in response to antibiotic treatments in our gut, or those of treated animals. This theoretical context has led to an extensive review of the literature available on the subject of environmental origins of antibiotic-resistance genes and is now widely accepted but environmental microbiologists are still struggling to obtain experimental data to test rigorously these hypotheses. One of the main objectives of this thesis was to provide a method to assess antibiotic resistance gene abundance and diversity in a broad range of environments with minimum biases and in respect to currently available technologies. This method was developed based on sequence-similarity screening of known antibiotic-resistance genes of publicly available metagenomic sequencing data. Results obtained analyzing environmental metagenomes confirmed that soil is a prominent reservoir of antibiotic-resistance genes and led to development of a double experimental

approach to study further their abundance and dynamic in soil. An *in vitro* study was setup with soil microcosms amended either with pure-antibiotic molecules or farm-effluents to study the impact of such pollutions on a reference soil never exposed previously to antibiotics (Rothamsted, Park Grass). This experiment evidenced that the sole antibiotic-molecule selection pressure does not elicit antibiotic-resistance genes abundance in soil at concentrations tested in our microcosm conditions while farm effluents are considerably modifying bacterial soil communities as viewed with RISA-PCR fingerprinting results. This questions how antibiotic-resistance genes are selected in soil since antibiotic-pressure exerted by concentrations usually found in agricultural soils are not significantly modifying antibiotic-resistance genes abundance in soil. In order to address this question *in situ*, we focused our interest on a very interesting study site in the frame of environmental impact assessment of human antibiotic consumption. Trois-Sauts is an isolated village in French Guiana Amazonian forest, built recently and where antibiotics consumption by the local population is precisely registered. Soil samples were collected inside and in the vicinity of the village, along a regularly inter-spaced line-transect and analyzed using methods described previously (qPCR, metagenomic sequencing). Results show that soil samples from the village are significantly different in term of taxonomic and ARGD distributions when compared to forest soil samples. Impacts of either soil-fertilization by farm-effluents or recent urbanization are manifest in our results but not correlated with increased antibiotic molecules concentration that. Indeed, whenever antibiotis concentration can be measured with rapidly in soil, they rapidly decay overtime.

This suggest that natural selection acts differently on the same antibiotic-resistance phenotype when considering either a different host (i.e. environmental or pathogens/commensal bacteria) or a different environment (i.e: a soil, or a patient treated with antibiotics) and underlies mobile genetic elements role of in these rapid adaptations and gene dynamics. Future experimental studies on the environmental aspect of antibiotic-resistance, but also clinical resistance, should be developped with the objective of linking data on antibiotic-resistance phenotype and replicons encoding these traits and their dynamic. We are arguing that while recent advances in sequencing technology throughput can now enable microbial ecologists to test more and more general ecology hypotheses, with the help of large amounts of sequence data and proper exerimental planning, but are still orders of magnitude to fully grasp soil complexity. It is however possible to achieve this

goal with actual technologies, when for example, replicons assembly can be confidently asserted and give the opportunity to annotate on a same contig a taxonomy or replicon type marker and an antibiotic resistance gene (ARGD) marker.

### **Experimental perspectives: targeting the mobilome.**

A conclusion of our work is that antibiotic resistance gene dissemination implies a complex ecology and we are still very limited, despite the results we obtained, in these conclusions, due to the very complexity of soil microbial ecology. A firm link must be established between taxonomy, replicon type or antibiotic resistance gene markers, such as their simultaneous presence on an assembled sequence for example. This is only with such results that we'll be able to address more specific questions, such as "*Which resistance genes are harbored by which hosts, on which replicons?*" It is possible, when considering a full plasmid sequence for example to understand if it has a narrow or broad-hosts spectrum, based on its incompatibility group, on which hosts it is susceptible to be transferred and expressed but also they're behavior: Are they auto-conjugative? Do they need a helper strain for transfer? Only replicated with the chromosome? Answers to these questions in the soil ecosystem are critical for the pharmaceuticals industry for example, to orient drug development on less prone to be transferred genes, if such exists. Microbial ecologists, by comparing which genes are harbored by which replicons can decipher their association network and further hypothesize on antibiotic resistance genes origins, role and selection in soil. However, assemblies of large short-reads sequence datasets obtained by direct sequencing of complex environment samples were mostly huge failures, in respect to sequencing cost per total assembled sequence length and the very limited gain in knowledge such results provide, considerably limiting the interest of such massive approaches. Indeed, successful metagenomes assembly require a limited diversity, for various reasons, the main one being that algorithms used for computing alignments assume all molecule in DNA samples are from the same genome. This is true also for metagenomic-specific assemblers. When diversity is low in samples, or rather that a same genome or few genomes are highly abundant, authors were able to assemble near full-length chromosome with confidence from metagenomic sequences. Providing these results, they were able to establish a metabolic model, coherent with surrounding environment (an acid mine drainage biofilm) and gave them further insights on the bacteria harboring this genome.

This study was ultimately published by in Nature by Jillian Banfield and colleagues in 2004, with this statement: "*Conversely, if the genomic heterogeneity within a species is dominated by large rearrangements, deletions, or insertions, it may be impossible to define composite genomes for species populations from natural communities.*" This work was published in 2004 and performed with Sanger cloning-sequencing of 3.2Kb plasmid. Both due to the fact such methods are really time-consuming and implies a lot of steps between the environment and sequences data, it will not be reiterated as often as science actually demands. *A contrario*, direct DNA sequencing with next-generation sequencing produces a lot of sequence data with minimal efforts. However, diversity is never considered prior to DNA extraction. We argue that specific DNA molecule isolation can considerably reduce DNA molecules diversity in DNA solutions. Such metagenomic studies are already published, focusing for example only on viruses' particles (virome), single-cell sequencing after cell sorting, etc. In every method, aim is to provide for sequencing, a targeted pool of DNA molecules.

A method is currently under active-development within our research team to focus and target on specific pools of mobile genetic elements: such as viral DNA, concentrated by filtration of viral particles from environmental samples but also plasmids. Plasmid DNA is covalently closed and circular (CCC); while chromosomal DNA is not or loose it rapidly upon extraction. These differences in spatial conformation confer an increased density to CCC DNA when bind to dye such as ethidium bromide. This method is not new and plasmid isolation with cesium chloride isopycnic density gradient is an old and standard laboratory procedure. We are however developing it specifically for soil, using a first step where microbial cells are extracted from soil first. Supercoiled DNA can loose its conformation rapidly, and most likely do not maintain it when in a dormant state, limiting considerably the recoverable plasmid fractions. Even considering all these limitations, plasmid DNA can be specifically extracted from these cells, using cesium chloride ultracentrifugation density gradients with precisely set cell density. Such a protocol is very time-consuming and therefore costly for now: a total of 5kg of soil are used to obtain sufficient amounts of DNA. However, we were able to obtain plasmid metagenomic DNA from soil and animal feces, in quantities now compatible with some sequencing platforms, building next-generation sequencing libraries with nanogram DNA quantities instead of micrograms. This method is still under development and will soon provide results in the form of plasmid-specific DNA

sequence read datasets obtained without culture specific biases, expression in a heterologous host, recruitment and eventually at a higher throughput than single-cell sequencing.

# Annexes.

## Chapter 2 - Annexes.

Table S1 - Metagenome datasets details.

Figure S1 - Figure S1 - Fifty most prevalent resistance type found in all metagenomes.

Table S2 - Fifty most abundant classes of antibiotic resistance gene found in environmental metagenomic datasets.

Figure S3 –ARGD-annotated reads SEED subsystems annotation.

Table S3 - Co-occurrence of antibiotic resistance classes in Roche 454 environmental metagenomic datasets.

## Chapter 3 - Annexes.

Figure S1 - PCR amplification of tetracyclines-specific antibiotic-resistance efflux-pump tetA(G) in tetracyclines-amended microcosms and PT2 pigs.

Figure S2 - Oxytetracycline decay measured in Rothasmted soil.

## Chapter 4 - Annexes.

Figure S1 – Steps involved in de novo positive primer synthesis.

Figure S2 - Relationship between number of sequence reads and number of detected 16S rRNA SSU reads by BLAST.

Figure S3 - Boxplots representing differences in relative abundance as a function of habitat.

Figure S4 - PCoA plots of unweighted UniFrac distance dissimilarity matrix between forest soil samples.

Table S1 – Oligonucleotide sequences and references of primers used in ARGD quantitative PCR assays.

Table S2 - Soil samples alpha-diversity analysis based on 16S rRNA gene amplicons sequencing results.

Table S3 - Relative abundance of each taxonomic phylum.

Table S3 - List of all antibiotics and their corresponding antibiotic-resistance classes detected in our analysis.

Table S4 - Absolute quantification results from qPCR antibiotic-resistance and mobile elements markers genes normalized by 16S rRNA gene copy number in DNA sample with respective 95% CI.



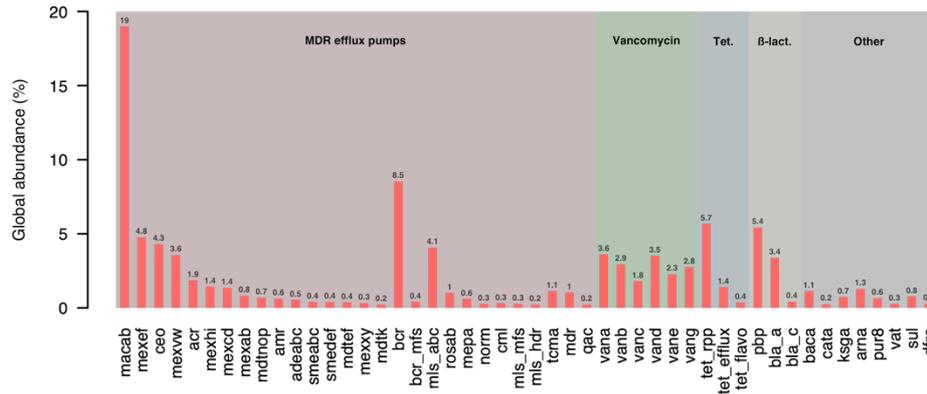
## **Chapter 2. - Annexes.**

### ***Large-scale metagenomic-based study of antibiotic resistance in the environment.***

Table S1

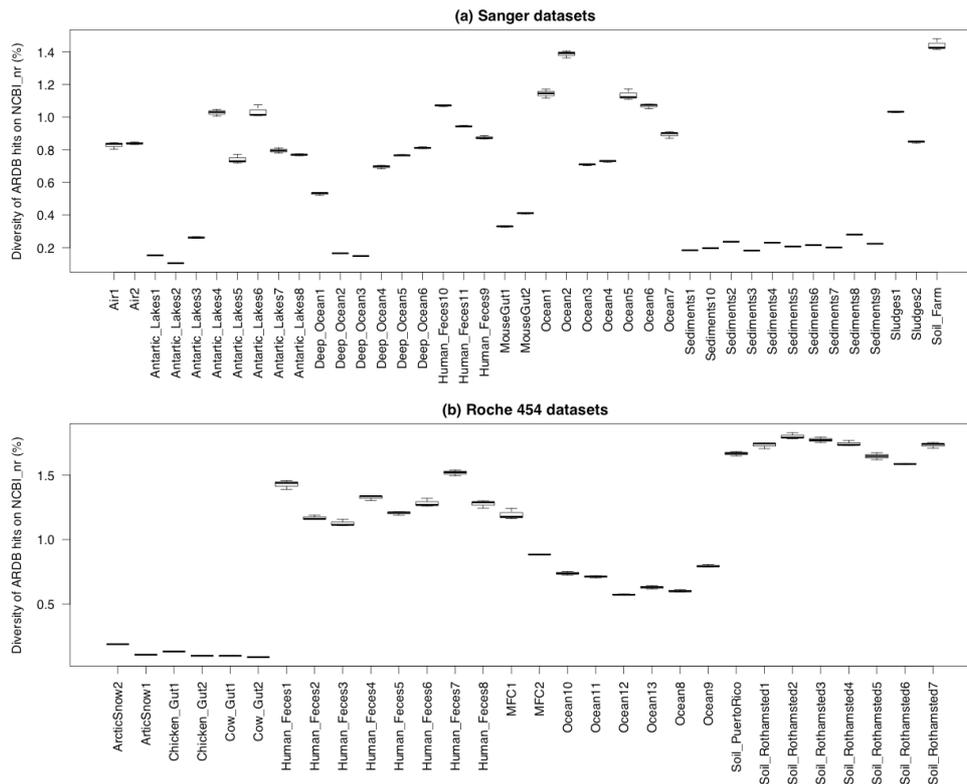
	Short Name	Metagenome ID	Project	Seq. technology	Nbr. reads	Average read length (bp)	Avg. Nb. Seq	Avg. Length	Abundance (%)	Diversity (%)
1	Antartic_Lakes1	Antartic_Lakes-4443679.3	Antartic Lakes	Sanger	10042	971.45	117247.13	996.8	2.29	4.63
2	Antartic_Lakes2	Antartic_Lakes-4443680.3	Antartic Lakes	Sanger	9672	994.85			1.46	3.37
3	Antartic_Lakes3	Antartic_Lakes-4443681.3	Antartic Lakes	Sanger	54446	1008.89			3.88	8.87
4	Antartic_Lakes4	Antartic_Lakes-4443682.3	Antartic Lakes	Sanger	283663	1001.43			1.54	18.11
5	Antartic_Lakes5	Antartic_Lakes-4443683.3	Antartic Lakes	Sanger	100085	1012.24			1.38	11.97
6	Antartic_Lakes6	Antartic_Lakes-4443684.3	Antartic Lakes	Sanger	281490	998.46			1.44	17.24
7	Antartic_Lakes7	Antartic_Lakes-4443686.3	Antartic Lakes	Sanger	103058	985.59			2.78	14.34
8	Antartic_Lakes8	Antartic_Lakes-4443687.3	Antartic Lakes	Sanger	95521	1001.5			2.83	13.57
9	ArticSnow1	Artic_Snow_High_Hg-4443127.3	Arctic Snow	Roche 454	16222	330.62	27175.5	354.325	0.43	1.80
10	ArticSnow2	Artic_Snow_Low_Hg-4443128.3	Arctic Snow	Roche 454	38129	378.03			0.33	3.10
11	Chicken_Gut1	Chicken_Gut-4440283.3	Chicken Cecum	Roche 454	294682	104.04	266311	103.94	0.06	1.70
12	Chicken_Gut2	Chicken_Gut-4440284.3	Chicken Cecum	Roche 454	237940	103.84			0.05	1.43
13	Cow_Gut1	Cow_Gut-4441679.3	Cow Rumen	Roche 454	264849	100.6	221781	101.09	0.06	1.03
14	Cow_Gut2	Cow_Gut-4441680.3	Cow Rumen	Roche 454	178713	101.58			0.06	0.97
15	Deep_Ocean4	Deep_Ocean-A_pompejana_Epibiont-4441663.3	Alvinella Pompejana Epibiont Metagenome	Sanger	293065	982.7			1.03	11.40
16	Deep_Ocean6	Deep_Ocean-Dissolved_Oxygen_min.layer-4442503.3	Microbial Community Genomics at the HOT/ALOHA	Sanger	11223	990.81			1.97	4.13
17	Deep_Ocean5	Deep_Ocean-Upper_euphotic-4442500.4	Microbial Community Genomics at the HOT/ALOHA	Sanger	11478	1029.06			1.93	4.50
18	Deep_Ocean1	Deep_Ocean-Whale_Fail_Bone-4441619.3	Whale Fall Metagenome	Sanger	40549	1017.83	210073	615.42	3.03	11.77
19	Deep_Ocean3	Deep_Ocean-Whale_Fail_Mat-4441656.4	Whale Fall Metagenome	Sanger	38281	990.68			3.23	12.57
20	Deep_Ocean2	Deep_Ocean-Whale_Fail_Rib-4441620.3	Whale Fall Metagenome	Sanger	38496	1016.07			4.20	13.97
21	Human_Feces1	Human-feces-Twin_Study-4440452.7	Twin Study	Roche 454	217386	237.87	432283.5	239.22	1.25	13.90
22	Human_Feces2	Human-feces-Twin_Study-4440460.5	Twin Study	Roche 454	414754	229.06			0.33	10.54
23	Human_Feces3	Human-feces-Twin_Study-4440595.4	Twin Study	Roche 454	510972	201.02			0.42	11.87
24	Human_Feces4	Human-feces-Twin_Study-4440611.3	Twin Study	Roche 454	495040	198.07			0.48	14.47
25	Human_Feces5	Human-feces-Twin_Study-4440613.3	Twin Study	Roche 454	302780	335.01			0.63	11.10
26	Human_Feces6	Human-feces-Twin_Study-4440614.3	Twin Study	Roche 454	519072	177.22			0.34	9.34
27	Human_Feces7	Human-feces-Twin_Study-4440616.3	Twin Study	Roche 454	502399	345.12			0.93	16.17
28	Human_Feces8	Human-feces-Twin_Study-4440825.3	Twin Study	Roche 454	495865	190.39			0.44	11.47
29	Human_Feces9	Human-feces-Kurokawa-4440939.3	28 Healthy Human Gut	Sanger	28900	1315.25	31851	1300.53	5.35	12.90
30	Human_Feces10	Human-feces-Kurokawa-4440942.3	29 Healthy Human Gut	Sanger	36455	1259.25			5.60	15.34
31	Human_Feces11	Human-feces-Kurokawa-4440943.3	29 Healthy Human Gut	Sanger	30198	1327.11			5.56	13.30
32	MFC1	MFC-C02-4447261.3	Microbial Fuel Cell	Roche 454	201041	NA			0.96	14.14
33	MFC2	MFC-W02-4447259.3	Microbial Fuel Cell	Roche 454	251718	NA			0.54	11.70
34	MouseGut1	Mouse_Gut-Lean-4440463.3	Obese and Lean Mouse Gut microbiota	Sanger	10845	781.8	11351	773.25	4.35	5.67
35	MouseGut2	Mouse_Gut-Obese-4440464.3	Obese and Lean Mouse Gut microbiota	Sanger	11857	764.71			4.60	5.17
36	Ocean1	Ocean-4441573.3	Global Ocean Survey	Sanger	317180	1012.13	203963	666.96	3.50	25.44
37	Ocean2	Ocean-4441574.3	Global Ocean Survey	Sanger	368835	1007.74			3.20	28.04
38	Ocean3	Ocean-4441576.3	Global Ocean Survey	Sanger	142352	1006.78			1.07	11.60
39	Ocean4	Ocean-4441577.3	Global Ocean Survey	Sanger	90901	1000.55			1.76	12.07
40	Ocean5	Ocean-4441591.3	Global Ocean Survey	Sanger	131798	1088.44			3.52	20.17
41	Ocean6	Ocean-4441660.3	Global Ocean Survey	Sanger	127122	1081.48			3.17	18.01
42	Ocean7	Ocean-4443688.3	Global Ocean Survey	Sanger	71068	1066.62			3.54	14.00
43	Ocean8	Ocean-4443693.3	North Atlantic Spring Bloom	Roche 454	276234	231.84			0.15	4.83
44	Ocean9	Ocean-4443713.3	North Atlantic Spring Bloom	Roche 454	217549	238.87			0.27	5.90
45	Ocean10	Ocean-4443714.3	North Atlantic Spring Bloom	Roche 454	186172	237.36			0.29	6.20
46	Ocean11	Ocean-4443716.3	North Atlantic Spring Bloom	Roche 454	216796	237.01			0.23	5.94
47	Ocean12	Ocean-4443725.3	North Atlantic Spring Bloom	Roche 454	244125	228.76			0.14	5.10
48	Ocean13	Ocean-4443729.3	North Atlantic Spring Bloom	Roche 454	261387	232.92			0.15	5.37
49	Sediments1	Halophile_Sediments-4440963.3	Guerrero Negro Hypersaline Sediments	Sanger	11562	646.02	12906.3	653.315	2.39	4.23

Table S1 - Details about metagenomic datasets used in this study.



**Figure S1 - Fifty most prevalent resistance type found in all metagenomes.**

Global percentage of hits in all metagenomes attributed to a specific resistance type as classified in ARDB. Most prevalent resistance type found in our analysis are MDR efflux pumps (red box), vancomycin resistance (green box), tetracycline resistance (blue box),  $\beta$ -lactams resistance (pale yellow box) and resistance to other antibiotic molecules (decreasing abundance order): Polymixin ; Bacitracin ; Sulfonamide ; Kasugamycin ; Puromycin ; Trimethoprim ; Streptogramin ; Chloramphenicol (grey box).



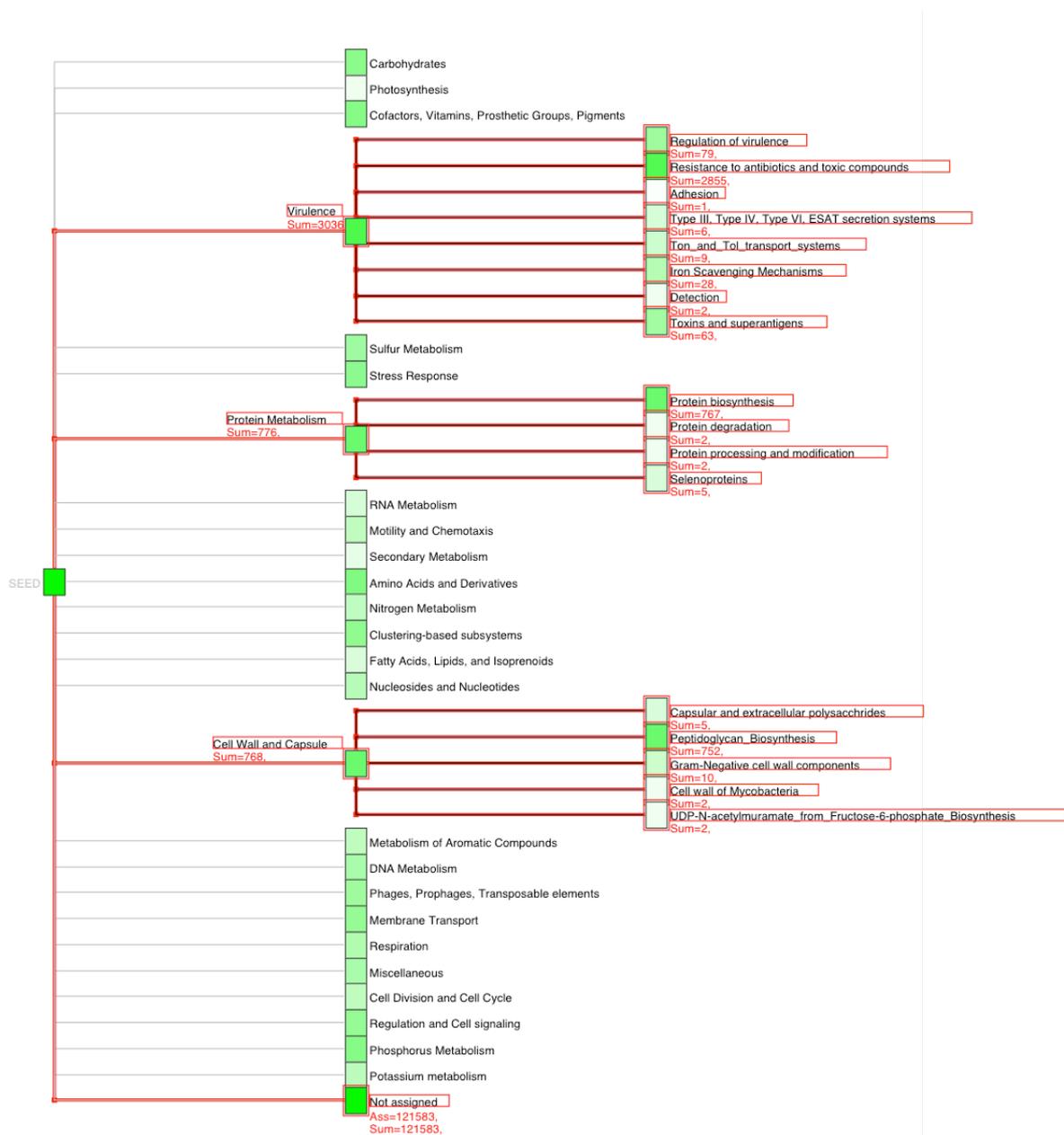
**Figure S2 - Relative diversity found on NCBI\_nr of ARDB annotated reads for each metagenome.**

Metagenomic reads giving a positive hit against ARDB were retrieved (n=128972) and compared with the same BLASTx parameters against NCBI\_nr. A total of 127865 reads (99.14%) from either Sanger datasets (a) or Roche 454 datasets (b) were annotated above the same stringency cut-off used for the ARDB BLASTx search. Diversity percentage has been calculated for each metagenome (i=71) based on three independent resampling of 1000 reads from each dataset annotated previously with ARDB. The number of unique hits was divided by the total number of different hits found in the whole ARDB positive reads subset of all metagenomic dataset (N=47839).

ARDB Resistance Class name	ARDB Mechanism of resistance	Antibiotic specificity	Global abundance in metagenomes (%)
macab	RND transporter Macrolide	Unspecific (MDR efflux)	18.99
mexef	RND transporter	Unspecific (MDR efflux)	4.76
ceo	RND transporter	Unspecific (MDR efflux)	4.29
mexvw	RND transporter	Unspecific (MDR efflux)	3.56
acr	RND transporter	Unspecific (MDR efflux)	1.86
mexhi	RND transporter	Unspecific (MDR efflux)	1.44
mexcd	RND transporter	Unspecific (MDR efflux)	1.35
mexab	RND transporter	Unspecific (MDR efflux)	0.81
mdtnop	RND transporter	Unspecific (MDR efflux)	0.69
amr	RND transporter	Unspecific (MDR efflux)	0.62
adeabc	RND transporter	Unspecific (MDR efflux)	0.54
smeabc	RND transporter	Unspecific (MDR efflux)	0.4
smedef	RND transporter	Unspecific (MDR efflux)	0.4
mdtef	RND transporter	Unspecific (MDR efflux)	0.37
mexxy	RND transporter	Unspecific (MDR efflux)	0.32
mdtk	RND transporter	Unspecific (MDR efflux)	0.22
bcr	ABC transporter system Bacitracin	Unspecific (MDR efflux)	8.53
bcr_mfs	ABC transporter system Bacitracin	Unspecific (MDR efflux)	0.41
mls_abc	ABC transporter Macrolide	Unspecific (MDR efflux)	4.06
rosab	Potassium antiporter system	Unspecific (MDR efflux)	1.01
mepa	MATE transporter	Unspecific (MDR efflux)	0.61
norm	MATE transporter	Unspecific (MDR efflux)	0.27
cml	MFS transporter Chloramphenicol	Unspecific (MDR efflux)	0.31
mls_mfs	MFS transporter Macrolide	Unspecific (MDR efflux)	0.27
mls_hdr	MFS transporter Macrolide	Unspecific (MDR efflux)	0.22
tcma	MFS transporter	Unspecific (MDR efflux)	1.13
mdr	MFS transporter	Unspecific (MDR efflux)	1.04
qac	MFS transporter	Unspecific (MDR efflux)	0.22
vana	VanA type operon	Vancomycin	3.62
vanb	VanB type operon	Vancomycin	2.93
vanc	VanC type operon	Vancomycin	1.8
vand	VanD type operon	Vancomycin	3.52
vane	VanE type operon	Vancomycin	2.27
vang	VanG type operon	Vancomycin	2.76
tet_rpp	Tetracycline ribosomal protection protein	Tetracycline	5.67
tet_efflux	Tetracycline-specific efflux pump	Tetracycline	1.4
tet_flavo	Flavoproteins resistance to Tetracycline	Tetracycline	0.36
pbp	Penicillin binding protein	Beta-lactams	5.42
bla_a	Class A beta-lactamase	Beta-lactams	3.38
bla_c	Class B beta-lactamase	Beta-lactams	0.42
baca	Bacitracin resistance	Bacitracin	1.14
cata	Chloramphenicol acetyltransferase	Chloramphenicol	0.25
ksga	Kasugamycin resistance	Kasugamycin	0.73
arna	Polymixin resistance	Polymixin	1.28
pur8	Puromycin resistance	Puromycin	0.65
vat	Virginiamycin resistance	Streptogramin	0.28
sul	Sulfonamide resistance	Sulfonamide	0.77
dfra	Trimethoprim resistance	Trimethoprim	0.29

**Table S2 - Fifty most abundant classes of antibiotic resistance gene found in environmental metagenomic datasets.**

Most prevalent resistance type found in our analysis are MDR efflux pumps, vancomycin resistance, tetracycline resistance,  $\beta$ -lactams resistance and resistance to other antibiotic molecules: Polymixin ; Bacitracin ; Sulfonamide ; Kasugamycin ; Puromycin ; Trimethoprim ; Streptogramin ; Chloramphenicol. Resistance mechanisms involved and antibiotic molecule targets are specified for each group of the 50 resistance classes found. Resistance classes are groups of genes implicated in the same resistance phenotype as defined in ARDB (ardb.cbcb.umd.edu).



**Figure S3 –ARGD-annotated reads SEED subsystems annotation.**

All reads from our initial query of 71 metagenomic datasets with a match in ARDB are compared to NCBI\_nr protein database using the same BLASTx parameters and stringency cut-off values and annotated with SEED. Darker or lighter green-shaded boxes indicate respectively higher or lower relative reads abundance for each node. All reads have a match in NCBI\_NR but most are not assigned to any functional category (n=121583; 94%) in SEED. Among assigned reads, Virulence (n=3036; 2.35%) is the most abundant category and contains antibiotic-resistance related functions. Resistance proteins are also found in other SEED nodes such as Protein metabolism (i.e.: Tetracycline ribosomal protection resistances), Cell wall and Capsule (i.e.: Penicillin-binding proteins) or Phages and transposable element (i.e.:  $\beta$ -lactamases). Read-counts for or leaves containing at least one ARDB-annotated reads are available in Table S5.

	Soil	Ocean	Human_Feces				
dfrb	0	0	0	mdtk	1	1	1
fusb	0	0	0	mdtm	1	1	1
mph	0	0	0	mdtnop	1	1	1
rif	0	0	0	meca	1	1	1
tet_other	0	0	0	mexab	1	1	1
vph	0	0	0	mexcd	1	1	1
aac	1	0	1	mexef	1	1	1
ant	1	0	1	mexhi	1	1	1
aph	1	0	1	mexvw	1	1	1
bla_a	1	0	1	mexxy	1	1	1
bla_b	1	0	1	mls_abc	1	1	1
bla_c	1	0	1	nora	1	1	1
bla_d	1	0	1	norm	1	1	1
blt	1	0	1	pbp	1	1	1
cata	1	0	1	pur8	1	1	1
lnu	1	0	1	qac	1	1	1
mdfa	1	0	1	rosab	1	1	1
mdtg	1	0	1	smeabc	1	1	1
mdth	1	0	1	smedef	1	1	1
mdtl	1	0	1	sul	1	1	1
mecr1	1	0	1	tcma	1	1	1
mepa	1	0	1	tet_efflux	1	1	1
mls_mfs	1	0	1	tet_rpp	1	1	1
sta	1	0	1	tet_xprt	1	1	1
str	1	0	1	vana	1	1	1
tet_flavo	1	0	1	vanb	1	1	1
tet_mod	1	0	1	vanc	1	1	1
tsnr	1	0	1	vand	1	1	1
vanz	1	0	1	vane	1	1	1
ykk	1	0	1	vang	1	1	1
lmp	0	0	1	vat	1	1	1
acr	1	1	1	ble	1	0	0
adeabc	1	1	1	bmr	1	0	0
amr	1	1	1	emea	1	0	0
arna	1	1	1	ere	1	0	0
baca	1	1	1	fush	1	0	0
bcr	1	1	1	mls_hdr	1	0	0
bcr_mfs	1	1	1	pmra	1	0	0
catb	1	1	1	qacab	1	0	0
ceo	1	1	1	qnr	1	0	0
cml	1	1	1	tmr	1	0	0
dfra	1	1	1	fos	1	1	0
emrd	1	1	1	mfpa	1	1	0
emre	1	1	1	pac	1	1	0
erm	1	1	1				
ksga	1	1	1				
lsa	1	1	1				
macab	1	1	1				
mdr	1	1	1				
mdtef	1	1	1				

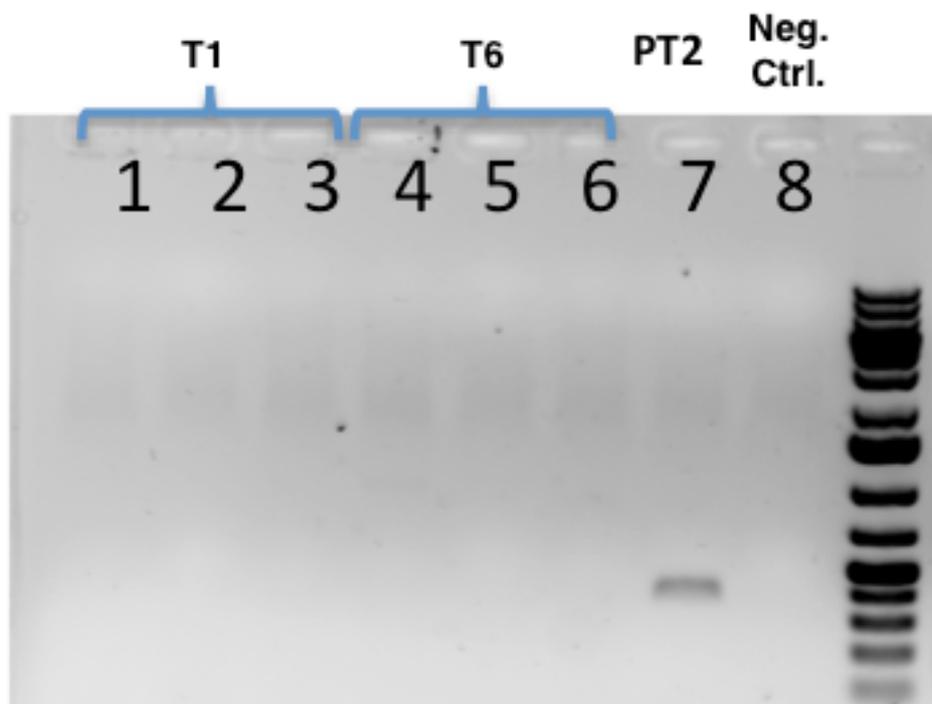
Note: "1" indicate the designed antibiotic resistance type has been found in any of the datasets (Roche 454 only) from the corresponding environments, inversly "0" indicate its absence.

**Table S3 - Co-occurrence of antibiotic resistance classes in Roche 454 environmental metagenomic datasets.**

Co-occurrence table of antibiotic resistance classes found in Roche 454 metagenomic datasets and used to draw the Venn diagram (Figure 3).

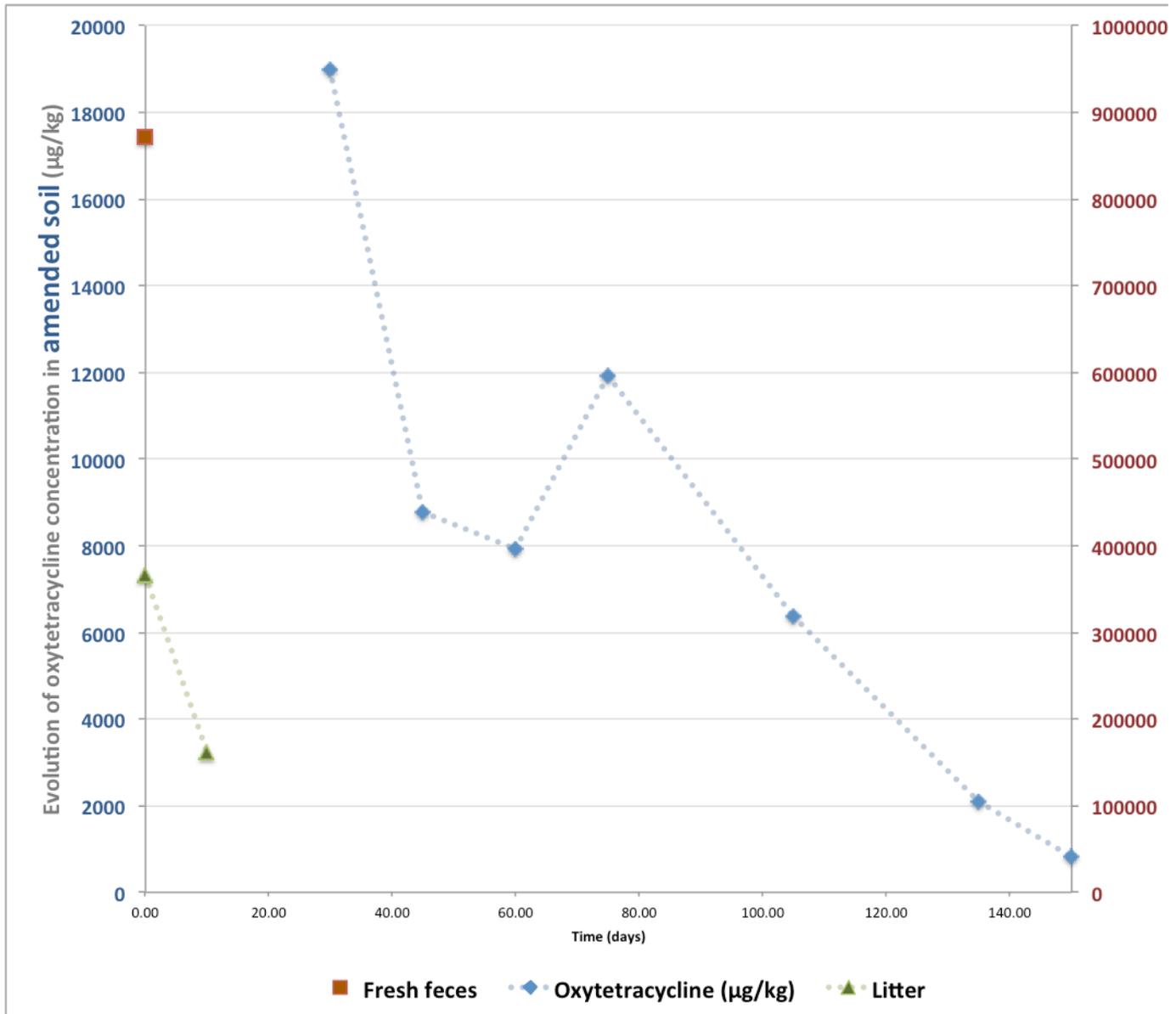
## Chapter 3 - Annexes.

### Antibiotic-amended Rothamsted, Park Grass soil microcosms experiment.



**Figure S1 - PCR amplification of tetracyclines-specific antibiotic-resistance efflux-pump tetA(G) in tetracyclines-amended microcosms and PT2 pigs.**

Agarose electrophoresis gel photography loaded with the products of tetA(G) antibiotic-resistance gene PCR amplification using target-specific primers pair. PCR reaction are performed in a 25 $\mu$ L reaction using 25ng of metagenomic DNA extracted from soil microcosms, respectively for each gel lanes: 1: T0\_PT2; 2: T0\_TetLC; 3: T0\_TetHC; 4: T3\_PT2; 5: T3\_TetLC; 6: T3\_TetHC and DNA extracted directly from PT2 pigs (treated with tetracyclines). Only pigs PT2 DNA is positive for tetA(G) at a PCR-based level sensitivity. Lane 8: no-template DNA PCR control. DNA size ruler: 1Kb+ Fermentas. Gene

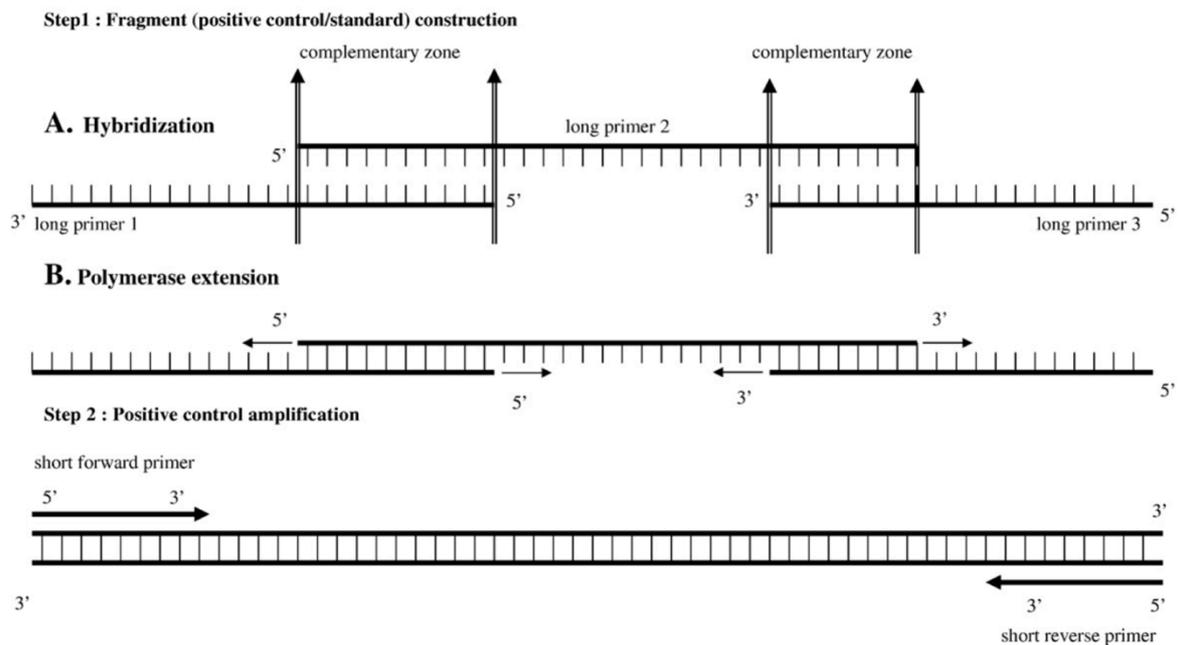


**Figure S2 - Decrease of oxytetracycline concentration in manure over time measured by LC/MS-MS.**

Plot showing the concentration of oxytetracycline measured in pig manure (blue) during its maturation time (150 days); fresh feces (red) used for manure preparation and animal litters (green). Concentrations of oxytetracycline are expressed as  $\mu\text{g}/\text{kg}$  of dry matter. These results were not obtained with the same samples used in amendment but similar ones from the same animals. These results illustrate antibiotic-molecule rapid decay in manure and also explain why tetracyclines are not observed in microcosms experiment were both dilution with soil molecules degradation and sorption to soil organic matter or clay particles will considerably lower the total available tetracyclines measured by LC/MS-MS chemical analyses.

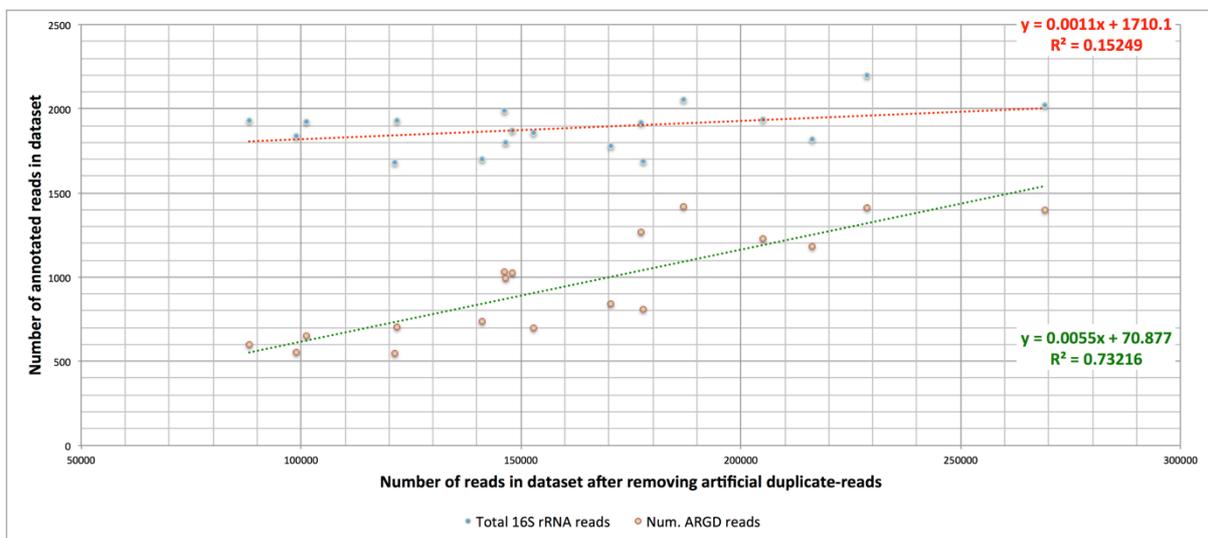
## Chapter 4 - Annexes.

# ***Amazonian resistome: evaluating antibiotic resistance abundance and diversity across a French Guiana forest soil gradient through metagenomic approaches.***



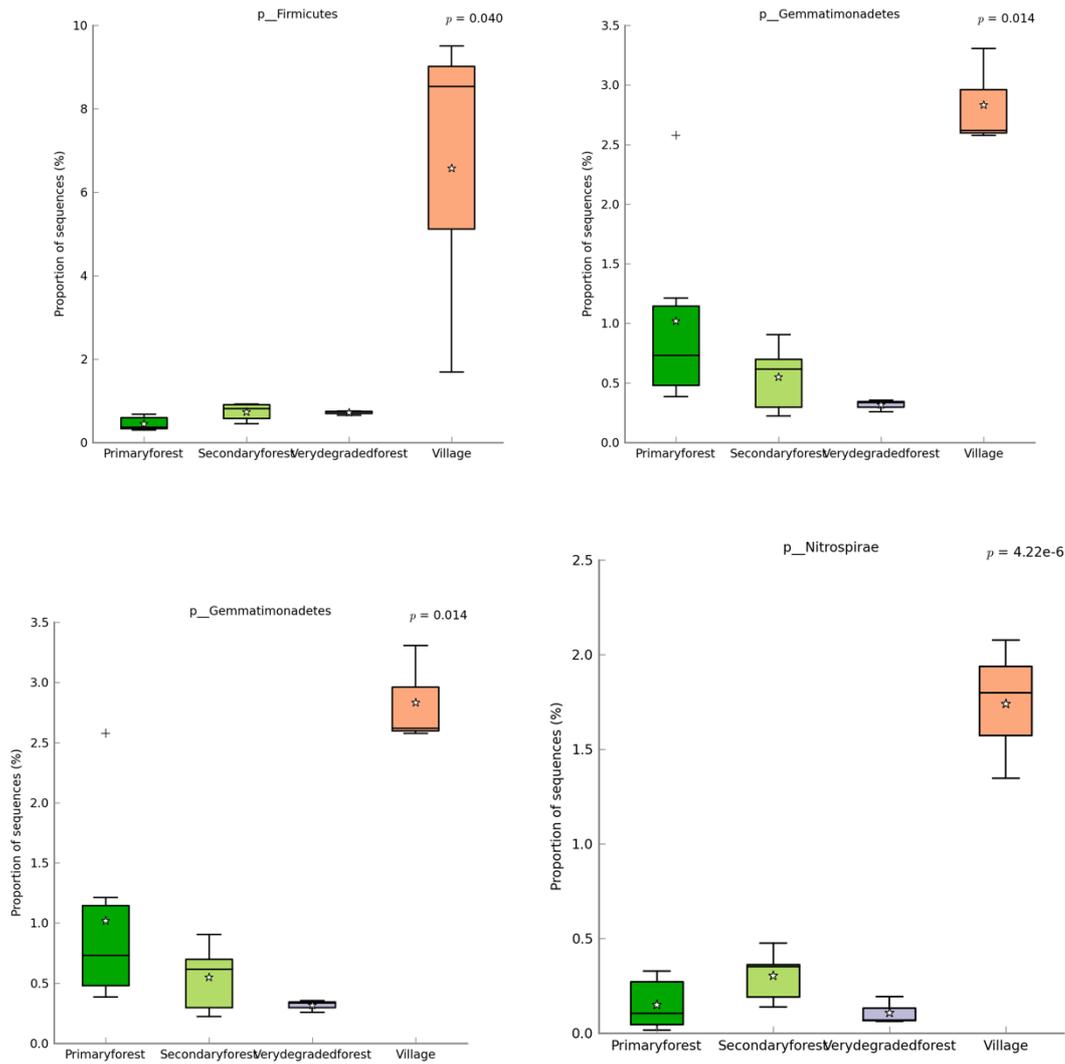
**Figure S1 – Steps involved in *de-novo* positive primer synthesis.**

Experimental steps in the synthesis of positive controls/standards. Step 1A: The three long oligonucleotides hybridize on each other's complementary sequence, which was bioinformatically designed. Step 1B: After hybridization, the polymerization without supplementary primers begins. Step 2: The last step of the positive control synthesis consists of classic PCR cycles with the supplementary short primers. {David:2008hv}. David, M. M et al. (2008). *A novel and rapid method for synthesizing positive controls and standards for quantitative PCR*. J Microbiol Methods 73, 73–77.

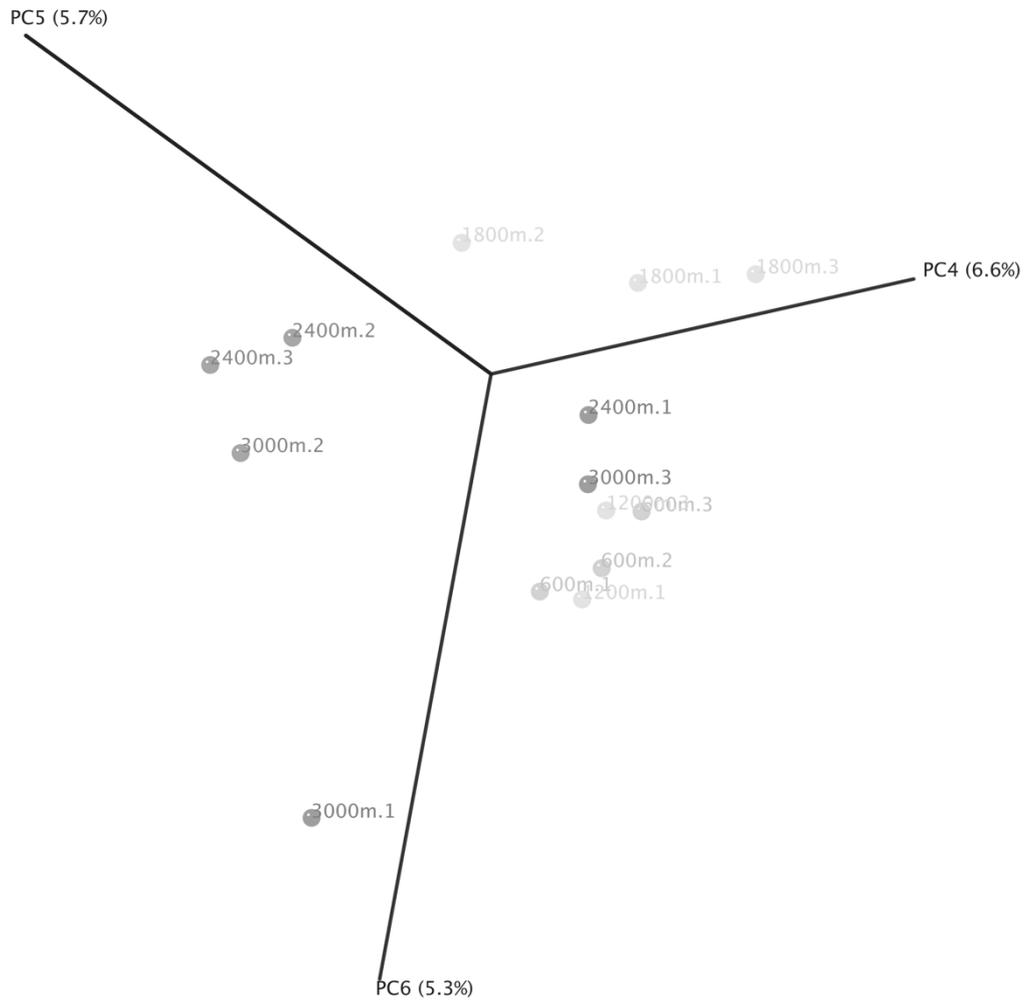


**Figure S2 - Relationship between number of sequence reads and number of annotated 16S rRNA SSU and ARGD reads and number of reads in dataset.**

16S rRNA SSU annotated-reads number in each dataset (blue dots) is determined by best-hit classification of shotgun-sequenced soil DNA metagenomic reads using the RDP SSU database on the MG-RAST website (see experimental procedures). ARGD-annotated reads number per dataset (red circles) is determined using a best-hit classification of shotgun-sequenced soil DNA metagenomic reads using an antibiotic resistance specific sequence database (ARDB; <http://ardb.cbcb.umd.edu>).



**Figure S3 - Boxplots representing differences in relative abundance as a function of habitat.** Boxplots are plotted using phylum level taxonomic assignment of 16S rRNA amplicons sequencing results. Given p-value are calculated using Turkey-Kramer post-hoc test at 95% confidence. Given p-values are adjusted for multiple-tests using Bonferroni correction.



**Figure S4 - PCoA plots of unweighted UniFrac distance dissimilarity matrix between forest soil samples.**

When only considering samples from the forest and visualizing the data using principal component axis 4, 5 and 6; forest samples are also forming groups based on distance and landscape type or use. PCoA plot is viewed in simulated-3D using KiNG software on data provided by QIIME analysis pipeline of 16S rRNA amplicons sequencing from metagenomic soil samples.

## Supplementary Information.

### Table S1 (1/2)

Name	Oligonucleotide sequence (5'-3')	T <sub>a</sub> * (°C)	Reference
Uni340F	CCTACGGGRBGCASCAG	55	{Takai:2000vn}
Uni806R	GGACTACNNGGGTATCTAAT	55	
blaSHV_for	TGCTTGGCCCGAATAACAA	60	{Hammond:2005ju}
blaSHV_rev	GCGTATCCCGCAGATAAATCA	60	
blaSHV_LF1	TGCTTGGCCCGAATAACAAAGCAGAGCGCATCGTGGTGATTTATCTGCGGGATACGC	65	This study
blaSHV_LF2	GCGTATCCCGCAGATAAATCACCACGATGCGCTCTGCTTTGTTATTCGGGCCAAGCA	65	
cfiA_for	GTCGCAGTTATGGCACAGAA	55	{Pumbwe:2007gq}
cfiA_rev	TCATTGATCGGTGTGTCCAG	55	
cfiA_LF1	GTCGCAGTTATGGCACAGAAAAGCGTAAAAATATCCGATGATATCAGTATCACCCAACCTCTCGG ACAAAGTGACACTTAT	65	This study
cfiA_LFmid	GAAGGTACCATACCCCATCCTTCGATTCGCGGAGGGATACATAAGTGTACACTTTGTCCG	65	
cfiA_LF2	GGATGGGGTATGGTACCTTCCAACGGAATGATTGTTATCAACAACCATCAGGCAGCGTTGCTGG ACACACCGATCAATGA	65	
ermA_for	TATCTTATCGTTGAGAAGGGATT	60	{MerinoDiaz:2007uu}
ermA_rev	CTACACTTGGCTTAGGATGAAA	60	
sull_for	CGCACCGGAAACATCGCTGCAC	65	{Pei:2007to}
sull_rev	TGAAGTCCGCCGCAAGGCTCG	65	
sull_LF1	CGCACCGGAAACATCGCTGCACGTGCTGTCGAACCTTCAAAGCTGAAGTCGGCGTTGGGGCTT CCGCTA	65	This study
sull_LFmid	AAGAAGGATTTCCGCGACACCGAGACCAATAGCGGAAGCCCCAACGCCGA	65	
sull_LF2	GTGTCGGGAAATCCTTCTTGGGCGCCACCGTTGGCCTTCTGTAAAGGATCTGGGTCCAGCGA GCCTTGGCGGAACTTCA	65	
tetG_for	GCTCGGTGGTATCTCTGCTC	62	{Fan:2007ee}
tetG_rev	AGCAACAGAATCGGGAACAC	62	
tetG_LF1	TGCTCGGTGGTATCTCTGCTCATGCCCGTTTATCGCCGCCCTTCTCAACGGGTTGCGGTTCC TGCTTGCTGCATTTCTCAAGGAGACTCATCA	70	This study

Table S1 (2/2)			
tetG_LF2	CTAGCCCGCAATGCATCATCCAGCCGTAACAGAACGAATGGTTTGATGCGAACCGGC TTCCGGTCCCGCCATGGCTGTGATGAGTCTCCTTGAGGAA	70	
tetG_LF3	TGATGCATTGCGGGGCTAGGTGCGCTTTTCGAGTTTTCTTCATTATCAACTGATCGG CCAAGTGCCTGCAGCCCTATGGGTCATATATGGCGAGGAC	70	
tetG_LF4	ACAAACGCTTGAAGATCGCATGTGTTGCCCAAACGCCGAGCGACAAACCAACGG TCGCGGTGTTCCACTGAAAACGGTCTCGCCATATATGACCC	70	
tetG_LF5	GCGATCTTCCAAGCGTTTGTACCAGCCCGCTTCAAGCCGGCTTGAGAGCGGCGCAC GCTGCTGTTGGCATGGCTGCGGATGCGACTGGCTTCGTTT	70	
tetG_LF6	AGCAACAGAATCGGGAACACCATCCATCCCTGCGTGGCAAAGCCAGAAGAACAAGC CAGTCGCATCCGAGCCATGCCAAACAGCAGCGTGCAGCCGCT	70	
tetG_LF7	GTGTTCCCGATTCTGTTGCT	70	
tetH_for	GTGATGTGACTCCCGCTAAAAAT	62	{Fan:2007ee}
tetH_rev	CCAGAACCGCCAAAGACATACC	62	
int1.F	gggtcaaggatctggatttcg	62	{Hoyle:2006ba}
int1.R	acatgctgtaaatcatgctcg	62	
trfA_qpcr-incP_for	ttcacsttctacgagmtkgccaggac	66	{Bahl:2009fg}
trfA_qpcr-incP_rev	gwcagcttgcttacttctcca	66	

**Table S1 – Oligonucleotide sequences and references of primers used in ARGD quantitative PCR assays.**

\*T<sub>a</sub> : annealing temperature used during qPCR reactions.

Distance	Seqs/Sample	Phylo. Dist. Ave.	Phylo. Dist. Err.	chao1 Ave.	chao1 Err.	observed OTUs Ave.	observed OTUs Err.	shannon Ave.	shannon Err.
0m	10.00	6.63	0.10	53.93	1.51	9.97	0.05	3.32	0.01
0m	886.00	175.45	11.07	1880.93	206.25	611.27	43.25	8.96	0.18
0m	1762.00	269.51	20.51	2495.51	326.85	1016.93	90.01	9.53	0.24
0m	2638.00	339.79	25.95	2940.44	324.80	1331.00	127.99	9.80	0.27
600m	10.00	6.30	0.12	52.87	1.51	9.93	0.05	3.31	0.01
600m	886.00	119.48	4.38	1596.69	90.29	571.03	7.18	8.83	0.02
600m	1762.00	176.01	7.50	2131.12	138.25	929.80	23.39	9.35	0.04
600m	2638.00	221.08	11.38	2489.68	175.47	1204.33	35.10	9.60	0.04
600m	3514.00	254.53	12.37	2804.66	169.58	1422.17	48.41	9.73	0.05
600m	4390.00	285.65	16.25	3120.86	205.66	1616.03	55.90	9.83	0.05
600m	5266.00	313.56	19.11	3314.80	268.23	1789.00	75.46	9.91	0.06
600m	6142.00	336.93	20.10	3528.05	245.18	1930.80	80.34	9.96	0.06
600m	7018.00	360.98	23.04	3720.03	286.99	2074.33	96.43	10.01	0.07
600m	7894.00	382.49	24.89	3920.60	294.66	2204.57	105.02	10.05	0.07
600m	8770.00	402.28	26.81	4042.27	304.39	2319.70	110.35	10.08	0.07
1200m	10.00	5.95	0.01	51.80	3.20	9.90	0.10	3.30	0.02
1200m	886.00	121.94	5.43	1619.24	94.62	580.95	11.35	8.87	0.05
1200m	1762.00	176.58	8.77	2152.70	163.22	930.05	28.15	9.36	0.07
1200m	2638.00	218.78	12.54	2454.98	78.04	1199.75	39.35	9.61	0.08
1200m	3514.00	255.47	14.79	2793.38	123.35	1423.00	54.20	9.75	0.10
1200m	4390.00	286.48	14.57	3056.86	139.01	1614.75	46.15	9.86	0.08
1200m	5266.00	315.73	19.40	3207.53	174.93	1780.60	64.80	9.93	0.09
1200m	6142.00	339.22	19.86	3387.31	206.55	1931.90	64.50	9.99	0.09
1200m	7018.00	362.89	23.86	3583.23	246.96	2068.95	72.75	10.03	0.09
1200m	7894.00	386.71	24.91	3790.71	222.86	2201.25	79.65	10.07	0.10
1200m	8770.00	406.06	26.44	3936.63	252.65	2310.70	83.90	10.10	0.10
1800m	10.00	6.37	0.16	51.80	0.00	9.90	nan	3.30	nan
1800m	886.00	135.82	4.80	1735.40	94.42	607.87	18.06	8.97	0.08
1800m	1762.00	201.95	6.54	2368.46	119.82	996.20	32.24	9.53	0.09
1800m	2638.00	254.05	8.36	2855.68	136.94	1297.30	41.87	9.79	0.10
1800m	3514.00	295.15	10.21	3145.32	98.16	1538.73	51.07	9.94	0.11
1800m	4390.00	331.01	12.15	3498.63	153.50	1756.37	60.32	10.05	0.11
2400m	10.00	5.95	0.36	52.87	1.51	9.93	0.05	3.31	0.01
2400m	886.00	114.92	6.30	1408.32	127.88	543.20	20.69	8.71	0.09
2400m	1762.00	165.77	11.23	1925.19	165.95	871.47	42.26	9.20	0.10
2400m	2638.00	205.79	15.24	2331.59	238.30	1127.13	57.46	9.44	0.11
2400m	3514.00	237.49	18.54	2599.86	249.20	1329.37	70.20	9.57	0.11
2400m	4390.00	263.90	19.78	2828.23	293.26	1503.13	80.74	9.66	0.11
3000m	10.00	6.42	0.23	51.80	2.61	9.90	0.08	3.30	0.02
3000m	886.00	145.72	14.05	1944.02	201.52	620.63	24.28	9.01	0.08
3000m	1762.00	215.39	20.32	2435.00	168.91	1023.03	43.90	9.58	0.10
3000m	2638.00	266.82	25.90	2864.45	273.86	1330.97	69.58	9.84	0.11
3000m	3514.00	310.88	30.18	3206.49	265.38	1591.20	88.27	10.01	0.13
3000m	4390.00	349.51	33.27	3491.17	299.78	1815.57	103.33	10.12	0.13

**Table S2 - Soil samples alpha-diversity analysis based on 16S rRNA gene amplicons sequencing results.**

Alpha-diversity metrics are calculated using random sequence resampling of increasing number of sequences (min: 10; max: 8770 or maximum number of sequences in each sample).

Legend	Taxonomy (Phylum level)	OTUs count	Total	Mean	SD	0m.1	0m.2	0m.3	600m.1	600m.2	600m.3	1200m.1	1200m.3	1800m.1	1800m.2	1800m.3	2400m.1	2400m.3	2400m.3	3000m.1	3000m.2	3000m.3
	Protobacteria	7	39.20%	39.20%	5.08%	32.20%	33.20%	33.70%	37.30%	45.80%	44.30%	43.70%	33.60%	44.50%	46.10%	44.90%	38.50%	37.10%	36.10%	43.20%	38.80%	33.40%
	Acidobacteria	4	26.50%	26.46%	8.88%	5.60%	13.30%	17.30%	32.40%	26.10%	26.80%	28.10%	41.00%	22.60%	27.80%	24.70%	31.40%	24.00%	24.00%	23.20%	31.10%	38.70%
	Actinobacteria	3	16.70%	16.68%	5.62%	34.40%	19.70%	20.60%	17.70%	16.30%	13.90%	16.10%	11.30%	17.30%	11.60%	12.30%	15.00%	22.60%	15.90%	12.80%	14.30%	11.80%
	Verrucomicrobia	1	3.90%	3.89%	1.00%	3.50%	2.80%	2.80%	3.20%	3.30%	4.10%	3.20%	3.50%	5.00%	3.60%	5.00%	3.80%	3.10%	3.00%	5.20%	6.20%	4.90%
	Chloroflexi	1	3.30%	3.27%	1.38%	2.80%	6.30%	7.20%	2.80%	2.10%	2.60%	2.50%	2.90%	3.10%	3.20%	3.10%	3.70%	2.50%	3.10%	3.10%	2.50%	2.10%
	Planctomycetes	1	3.00%	3.02%	0.88%	5.00%	3.70%	4.00%	2.30%	2.90%	3.70%	2.50%	2.40%	2.90%	3.40%	3.80%	3.10%	1.80%	2.00%	3.60%	2.20%	2.00%
	Firmicutes	0	1.60%	1.64%	2.76%	8.40%	9.40%	1.70%	0.80%	0.60%	0.70%	0.90%	0.40%	0.60%	0.90%	0.80%	0.30%	0.60%	0.30%	0.40%	0.40%	0.70%
	Gemmatimonadetes	0	1.10%	1.06%	1.01%	2.50%	2.60%	3.30%	0.40%	0.30%	0.40%	0.20%	0.30%	0.90%	0.60%	0.70%	0.50%	0.40%	0.40%	2.60%	0.90%	1.20%
	Bacteroidetes	0	1.00%	1.04%	1.13%	1.50%	3.60%	3.60%	0.30%	0.20%	0.40%	0.70%	0.30%	0.40%	0.50%	1.00%	0.30%	0.40%	0.30%	2.60%	0.70%	0.80%
	Other Bacteria	0	0.70%	0.70%	0.22%	0.60%	0.80%	0.80%	0.70%	0.50%	0.70%	0.30%	0.70%	0.60%	0.50%	0.70%	0.80%	1.40%	1.30%	0.60%	0.70%	0.80%
	AD3	0	0.70%	0.69%	0.55%	0.00%	0.00%	0.00%	0.90%	0.60%	0.70%	0.60%	1.10%	0.20%	0.20%	0.10%	0.20%	0.30%	0.20%	0.00%	0.60%	1.10%
	Nitrospirae	0	0.50%	0.47%	0.64%	1.80%	2.10%	1.30%	0.10%	0.10%	0.20%	0.10%	0.20%	0.40%	0.40%	0.50%	0.00%	0.00%	0.00%	0.30%	0.20%	0.30%
	Unclassified;Other	0	0.40%	0.44%	0.88%	0.50%	0.00%	0.20%	0.10%	0.10%	0.10%	0.10%	0.40%	0.20%	0.10%	0.40%	0.20%	3.80%	0.30%	0.20%	0.10%	0.20%
	WS3	0	0.40%	0.42%	0.60%	0.50%	0.70%	2.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.30%	0.30%	0.40%	0.00%	0.00%	0.00%	1.70%	0.60%	0.70%
	WPS-2	0	0.30%	0.28%	0.29%	0.00%	0.00%	0.00%	0.60%	0.60%	0.50%	0.50%	1.10%	0.20%	0.20%	0.10%	0.20%	0.30%	0.20%	0.00%	0.10%	0.20%
	Armatimonadetes	0	0.20%	0.17%	0.08%	0.10%	0.30%	0.30%	0.20%	0.10%	0.10%	0.10%	0.20%	0.20%	0.20%	0.20%	0.20%	0.20%	0.20%	0.10%	0.20%	0.20%
	Archaea : Crenarchaeota	0	0.10%	0.05%	0.11%	0.20%	0.40%	0.20%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	Cyanobacteria	0	0.10%	0.08%	0.06%	0.10%	0.20%	0.00%	0.10%	0.10%	0.10%	0.00%	0.10%	0.10%	0.10%	0.10%	0.10%	0.00%	0.10%	0.10%	0.10%	0.10%
	Elusimicrobia	0	0.10%	0.06%	0.15%	0.10%	0.00%	0.20%	0.30%	0.10%	0.10%	0.10%	0.10%	0.10%	0.10%	0.10%	0.10%	0.00%	0.10%	0.10%	0.10%	0.10%
	GAL15	0	0.10%	0.06%	0.15%	0.10%	0.00%	0.20%	0.30%	0.10%	0.10%	0.10%	0.10%	0.10%	0.10%	0.10%	0.10%	0.00%	0.10%	0.10%	0.10%	0.10%
	Archaea : Euryarchaeota	0	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	BRC1	0	0.00%	0.04%	0.05%	0.10%	0.00%	0.10%	0.00%	0.00%	0.00%	0.00%	0.10%	0.00%	0.00%	0.10%	0.00%	0.10%	0.00%	0.10%	0.00%	0.00%
	Chloridia	0	0.00%	0.04%	0.08%	0.00%	0.10%	0.30%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.10%	0.00%	0.00%
	Chlorobi	0	0.00%	0.01%	0.03%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.10%
	FCPU426	0	0.00%	0.01%	0.02%	0.00%	0.10%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	Fibrobacteres	0	0.00%	0.01%	0.02%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	GN02	0	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	GN04	0	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	LD1	0	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	NC10	0	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	NKB19	0	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	OD1	0	0.00%	0.01%	0.02%	0.00%	0.00%	0.10%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	OP3	0	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	OP9	0	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	SBR1093	0	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	SC4	0	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	Spirochaetes	0	0.00%	0.02%	0.04%	0.10%	0.10%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.10%	0.00%	0.00%	0.00%	0.00%	0.00%	0.10%
	TM6	0	0.00%	0.02%	0.04%	0.10%	0.10%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.10%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	TM7	0	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	Tenerites	0	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	Thermi	0	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	WS1	0	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	WS4	0	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	WYO	0	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	ZB3	0	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%

Table S3 - Relative abundance of taxonomic phylum in each sample based on 16S rRNA gena amplicons sequencing.

<ul style="list-style-type: none"> <li>• <b>Aminoglycoside</b> <ul style="list-style-type: none"> <li>aac</li> <li>ant</li> <li>aph</li> <li>emea</li> <li>sta</li> </ul> </li> <li>• <b>Bacitracin</b> <ul style="list-style-type: none"> <li> Baca</li> <li>Bacitracin_efflux</li> <li>bcra</li> <li>Bacitracin_efflux_pump</li> <li>bcr</li> </ul> </li> <li>• <b>Beta-lactams</b> <ul style="list-style-type: none"> <li>bla_a</li> <li>bla_b</li> <li>bla_c</li> <li>bla_d</li> <li>meca</li> <li>mecr1</li> <li>pbp</li> </ul> </li> <li>• <b>Bleomycin</b> <ul style="list-style-type: none"> <li>ble</li> </ul> </li> <li>• <b>Chloramphenicol</b> <ul style="list-style-type: none"> <li>cata</li> <li>catb</li> <li>ceo</li> <li>cml</li> </ul> </li> <li>• <b>Erythromycin</b> <ul style="list-style-type: none"> <li>ere</li> <li>erm</li> </ul> </li> <li>• <b>Fluoroquinolone</b> <ul style="list-style-type: none"> <li>emea</li> <li>mfpa</li> <li>qnr</li> </ul> </li> <li>• <b>Fosfomycin</b> <ul style="list-style-type: none"> <li>fos</li> </ul> </li> <li>• <b>Fusaric_acid</b> <ul style="list-style-type: none"> <li>fusb</li> <li>fush</li> </ul> </li> <li>• <b>Kasugamycin</b> <ul style="list-style-type: none"> <li>ksga</li> </ul> </li> <li>• <b>Lincomycin</b> <ul style="list-style-type: none"> <li>lnu</li> </ul> </li> <li>• <b>Macrolide</b> <ul style="list-style-type: none"> <li>mph</li> </ul> </li> <li>• <b>MDR_efflux</b> <ul style="list-style-type: none"> <li>acr</li> <li>acrab</li> <li>adeabc</li> <li>amr</li> <li>arn</li> <li> Baca</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>bcr</li> <li>blt</li> <li>bmr</li> <li>ceo</li> <li>emrd</li> <li>emre</li> <li>lsa</li> <li>lmra</li> <li>lsa</li> <li>macab</li> <li>mar</li> <li>mdr</li> <li>mdt</li> <li>mdtef</li> <li>mdtg</li> <li>mdth</li> <li>mdtk</li> <li>mdtl</li> <li>mdtm</li> <li>mdtnop</li> <li>mepa</li> <li>mex</li> <li>mexab</li> <li>mexcd</li> <li>mexef</li> <li>mexhi</li> <li>mexih</li> <li>mexvw</li> <li>mexxy</li> <li>mey</li> <li>mls_abc</li> <li>mls_hdr</li> <li>mls_mfs</li> <li>msra</li> <li>norm</li> <li>pmra</li> <li>rosab</li> <li>smeabc</li> <li>smedef</li> <li>srmb</li> <li>tcma</li> <li>vgaa</li> <li>ykk</li> </ul> <ul style="list-style-type: none"> <li>• <b>Polymyxin</b> <ul style="list-style-type: none"> <li>arn</li> <li>arna</li> </ul> </li> <li>• <b>Puromycin</b> <ul style="list-style-type: none"> <li>pac</li> <li>pur8</li> </ul> </li> <li>• <b>QAC_efflux</b> <ul style="list-style-type: none"> <li>qac</li> <li>qacab</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>• <b>Streptogramin_A</b> <ul style="list-style-type: none"> <li>vat</li> </ul> </li> <li>• <b>Streptomycin</b> <ul style="list-style-type: none"> <li>str</li> </ul> </li> <li>• <b>Sulfonamide</b> <ul style="list-style-type: none"> <li>sul</li> </ul> </li> <li>• <b>Tetracycline</b> <ul style="list-style-type: none"> <li>tet_efflux</li> <li>tet_flavo</li> <li>tet_mod</li> <li>tet_other</li> <li>tet_rpp</li> <li>tet_xprt</li> </ul> </li> <li>• <b>Thiostrepton</b> <ul style="list-style-type: none"> <li>tsnr</li> </ul> </li> <li>• <b>Trimethoprim</b> <ul style="list-style-type: none"> <li>dfra</li> <li>dfrb</li> <li>Tunicamycin</li> <li>tmr</li> </ul> </li> <li>• <b>Vancomycin</b> <ul style="list-style-type: none"> <li>vana</li> <li>vanb</li> <li>vanc</li> <li>vand</li> <li>vane</li> <li>vang</li> <li>vanz</li> </ul> </li> <li>• <b>Viomycin</b> <ul style="list-style-type: none"> <li>vph</li> </ul> </li> <li>• <b>Virginiamycin</b> <ul style="list-style-type: none"> <li>vat</li> <li>Vatb</li> </ul> </li> </ul>
---	---	---

**Table S2 – List of all antibiotics and their corresponding antibiotic-resistance classes detected in our analysis.**

Antibiotic Resistance Genes Determinants (ARGD) are classified in the Antibiotic Resistance Database according to three hierarchical level: ARDBid (non-redundant sequences), ARGD type (gene), ARGD class (gene family, enzyme type) and ARGD mechanism (molecular mechanism of resistance when known). For sake of clarity, only ARGD classes detected in our analyses are reported in this list under the name of the antibiotic molecule or class of molecule they confer resistance to.

(1/2)

	Distance	Habitat	16S rRNA	S.D.16S	blaSHV/16S	S.D.shv	blaTEM/16S	S.D.tem	cfIA/16S	S.D.cfia
EES_0m1	a.0m	Village	0.0E+00	2.5E-01	-2.4E+00	5.3E-01	-4.0E+00	4.7E-01	-2.1E+00	2.4E+00
EES_0m2	a.0m	Village	0.0E+00	-2.4E-01	-2.6E+00	5.4E-01	-5.9E+00	7.2E-01	-3.6E+00	3.0E+00
EES_0m3	a.0m	Village	0.0E+00	-2.5E-01	-2.4E+00	5.1E-01	-5.7E+00	6.7E-01	-3.4E+00	2.9E+00
EES_600m1	b. 600m	Degraded forest	0.0E+00	-2.5E-01	-2.0E+00	4.9E-01	-4.7E+00	5.3E-01	-3.8E+00	1.4E+00
EES_600m2	b. 600m	Degraded forest	0.0E+00	-2.5E-01	-2.3E+00	7.3E-01	-4.5E+00	5.1E-01	-3.6E+00	1.4E+00
EES_600m3	b. 600m	Degraded forest	0.0E+00	-3.9E-01	-1.0E+00	5.0E-01	-3.4E+00	4.9E-01	-2.8E+00	1.1E+00
EES_1200m1	c. 1200m	Secondary forest	0.0E+00	-2.5E-01	-2.5E+00	5.1E-01	-4.8E+00	5.3E-01	-3.9E+00	1.3E+00
EES_1200m2	c. 1200m	Secondary forest	0.0E+00	-2.4E-01	-1.6E+00	4.9E-01	-4.5E+00	5.4E-01	-4.3E+00	6.7E-01
EES_1200m3	c. 1200m	Secondary forest	0.0E+00	-2.4E-01	-1.6E+00	5.1E-01	-4.3E+00	5.8E-01	-3.6E+00	1.3E+00
EES_1800m1	d. 1800m	Secondary forest	0.0E+00	-2.5E-01	-2.6E+00	5.2E-01	-4.6E+00	5.0E-01	-3.9E+00	1.2E+00
EES_1800m2	d. 1800m	Secondary forest	0.0E+00	-2.5E-01	-2.6E+00	5.1E-01	-4.3E+00	4.4E-01	-3.7E+00	1.0E+00
EES_1800m3	d. 1800m	Secondary forest	0.0E+00	-2.5E-01	-2.6E+00	5.1E-01	-4.2E+00	4.3E-01	-3.8E+00	8.0E-01
EES_2400m1	e. 2400m	Primary forest	0.0E+00	-2.9E-01	-2.4E+00	5.0E-01	-4.3E+00	4.6E-01	-4.1E+00	6.3E-01
EES_2400m2	e. 2400m	Primary forest	0.0E+00	-2.5E-01	-2.3E+00	5.1E-01	-4.4E+00	4.9E-01	-4.0E+00	9.5E-01
EES_2400m3	e. 2400m	Primary forest	0.0E+00	-2.5E-01	-2.3E+00	5.0E-01	-4.5E+00	4.9E-01	-4.2E+00	7.6E-01
EES_3000m1	f. 3000m	Primary forest	0.0E+00	-2.5E-01	-2.4E+00	5.1E-01	-4.0E+00	4.5E-01	-3.9E+00	6.1E-01
EES_3000m2	f. 3000m	Primary forest	0.0E+00	-2.5E-01	-2.5E+00	5.2E-01	-4.7E+00	5.2E-01	-4.4E+00	7.3E-01
EES_3000m3	f. 3000m	Primary forest	0.0E+00	-2.4E-01	-2.1E+00	5.2E-01	-4.5E+00	5.4E-01	-5.1E+00	6.4E-02
EES_Roth.	g.Roth.	Grassland	0.0E+00	-2.4E-01	-2.2E+00	5.2E-01	-4.2E+00	5.1E-01	-5.9E+00	1.2E+00

(2/2)

ermA/16S	S.D.ermA	sulI/16S	S.D.sulI	tetA(G)/16S	S.D.tetG	tetA(H)/16S	S.D.tetH	int1/16S	S.D.int1	trfA/16S	S.D.trfA
-2.0E+00	9.3E-01	-1.4E+00	8.5E-02	-3.4E+00	7.7E-01	-4.9E+00	2.4E-01	-5.1E+00	6.6E-01	-8.9E-01	8.1E-01
-1.9E+00	9.0E-01	-2.1E+00	1.0E-01	-3.6E+00	7.8E-01	-4.6E+00	2.2E-01	-4.6E+00	5.9E-01	-3.2E+00	1.0E+00
-2.7E+00	1.0E+00	-2.1E+00	9.9E-02	-3.9E+00	7.9E-01	-5.4E+00	2.5E-01	-4.4E+00	5.4E-01	-2.2E+00	8.8E-01
-3.0E+00	1.1E+00	-1.9E+00	9.3E-02	-3.2E+00	7.3E-01	n.d.	n.d.	-4.6E+00	5.7E-01	-2.9E+00	9.7E-01
-2.9E+00	1.0E+00	-2.0E+00	9.3E-02	-3.2E+00	7.2E-01	-4.8E+00	2.2E-01	-4.7E+00	5.7E-01	-2.2E+00	8.6E-01
-1.7E+00	1.0E+00	-5.5E-01	8.8E-02	-2.1E+00	7.4E-01	-3.5E+00	2.2E-01	-3.3E+00	5.6E-01	-1.4E+00	9.2E-01
-2.7E+00	9.9E-01	-2.1E+00	9.6E-02	-3.3E+00	7.2E-01	-4.7E+00	2.2E-01	-4.7E+00	5.7E-01	-2.9E+00	9.6E-01
-3.2E+00	1.1E+00	-1.3E+00	8.7E-02	-3.1E+00	7.5E-01	-4.1E+00	2.1E-01	-5.0E+00	6.6E-01	-4.1E+00	1.2E+00
-2.4E+00	1.1E+00	-1.1E+00	9.4E-02	-2.3E+00	7.2E-01	-3.3E+00	1.9E-01	-4.0E+00	5.9E-01	-2.7E+00	1.1E+00
-2.9E+00	1.0E+00	-2.1E+00	9.6E-02	-3.3E+00	7.3E-01	-5.0E+00	2.3E-01	-4.4E+00	5.4E-01	-3.0E+00	9.7E-01
-2.8E+00	9.9E-01	-2.1E+00	9.4E-02	-4.5E+00	8.5E-01	-5.1E+00	2.3E-01	-4.5E+00	5.4E-01	-2.4E+00	8.8E-01
-2.9E+00	1.0E+00	-2.3E+00	1.0E-01	-3.3E+00	7.1E-01	-5.4E+00	2.4E-01	-4.4E+00	5.3E-01	-2.7E+00	9.0E-01
-3.2E+00	1.1E+00	-2.0E+00	9.4E-02	-3.2E+00	7.2E-01	-5.0E+00	2.3E-01	-4.2E+00	5.2E-01	-3.3E+00	1.0E+00
-2.9E+00	1.0E+00	-2.0E+00	9.5E-02	-3.2E+00	7.2E-01	-5.1E+00	2.4E-01	-4.3E+00	5.4E-01	-2.5E+00	9.1E-01
-3.1E+00	1.1E+00	-2.2E+00	1.0E-01	-3.2E+00	7.2E-01	-5.5E+00	2.5E-01	-4.4E+00	5.4E-01	-3.3E+00	1.0E+00
-3.0E+00	1.0E+00	-2.2E+00	9.8E-02	-3.5E+00	7.5E-01	-5.2E+00	2.4E-01	-4.8E+00	5.8E-01	-3.0E+00	9.6E-01
-3.4E+00	1.1E+00	-2.3E+00	1.0E-01	-3.4E+00	7.4E-01	-4.8E+00	2.2E-01	-5.1E+00	6.3E-01	-2.6E+00	9.2E-01
-3.6E+00	1.2E+00	-1.8E+00	9.9E-02	-3.2E+00	7.6E-01	-4.6E+00	2.3E-01	-5.3E+00	7.0E-01	-3.5E+00	1.1E+00
-4.1E+00	1.4E+00	-2.0E+00	1.0E-01	-3.6E+00	8.0E-01	-4.5E+00	2.4E-01	-6.6E+00	8.9E-01	-5.5E+00	1.6E+00

Table S3 - Absolute quantification results from qPCR antibiotic-resistance and mobile elements markers genes normalized by 16S rRNA gene copy number in DNA sample with respective 95% CI.

# Peer-reviewed publications.

**Article - Nesme, J., and Simonet, P.** *The soil resistome: a critical review on origins, ecology and dissemination potential of antibiotic resistance in telluric bacteria.* **Environmental Microbiology**, (accepted for publication in september 2014).

**Article - Nesme, J., Cécillon, S., Delmont, T. O., Monier, J.-M., Vogel, T. M., and Simonet, P.** (2014). Large-Scale Metagenomic-Based Study of Antibiotic Resistance in the Environment. **Current Biology**, 24, 1096–1100

1 **Large-scale metagenomic-based study of antibiotic resistance in the environment.**

2 **AUTHORS:** Joseph Nesme<sup>1</sup>, Sébastien Cécillon<sup>1</sup>, Tom O. Delmont<sup>1</sup>, Jean-Michel Monier<sup>1\*</sup>,  
3 Timothy M. Vogel<sup>1</sup>, Pascal Simonet<sup>1,2</sup>

4 <sup>1</sup> Environmental Microbial Genomics, Laboratoire Ampère, CNRS UMR 5005, Ecole Centrale de Lyon,  
5 Université de Lyon, 36 Avenue Guy de Collongue, 69134 Ecully, France.

6 \* J-M. Monier current address: Enoveo, 7 place Antonin Poncet, 69002 Lyon, France.

7  
8 <sup>2</sup> Corresponding author: pascal.simonet@ec-lyon.fr, Tel.: + (33) 4 72 18 60 92  
9

10 **SUMMARY:** Antibiotic resistance, including multi-resistance acquisition and dissemination  
11 by pathogens, is a critical healthcare issue threatening our management of infectious  
12 diseases [1-3]. Rapid accumulation of resistance phenotypes implies a reservoir of  
13 transferable antibiotic resistance gene determinants (ARGDs) selected in response to  
14 inhibiting antibiotic concentrations as found in hospitals [1, 3-5]. Antibiotic resistance genes  
15 were found in environmental isolates, soil DNA [4-6], secluded caves [6, 7] and permafrost  
16 DNA [7, 8]. Antibiotics target essential and ubiquitous cell functions and resistance is a  
17 common characteristic of environmental bacteria [8-11]. Environmental ARGDs are an  
18 abundant reservoir of potentially transferable resistance for pathogens [9-12]. Using  
19 metagenomic sequences, we show that ARGDs can be detected in all ( $n=71$ ) environments  
20 analyzed. Soil metagenomes had the most diverse pool of ARGDs. The most common types  
21 of resistances found in environmental metagenomes were efflux pumps and genes  
22 conferring resistance to vancomycin, tetracycline or *beta*-lactam antibiotics used in  
23 veterinary and human healthcare. Our study describes the diverse and abundant antibiotic  
24 resistance genes in non-clinical environments and shows that these genes are not randomly  
25 distributed among different environments (*e.g.*, soil, oceans or human feces).

26 **HIGHLIGHTS:**

- 27 1. First large-scale environmental antibiotic resistance DNA sequence screening.  
28 2. Antibiotic resistance genes detected in all ecosystems  
29 3. ARGD distribution in the environment is ecosystem specific.

30 **KEYWORDS:** antibiotic resistance, resistome, antibiotic resistance gene determinants, ARGD,  
31 metagenomic, next-generation sequencing, BLASTx.

62 Relative antibiotic resistance gene determinant (ARGD) abundance ranged from 0.05% in  
63 chicken gut (metagenome accession number 4440284.3) to 5.6% in human feces  
64 (4440942.3) and differed significantly (at 95% CI) among the different environmental  
65 datasets (Figure 1, see also Table S1). Datasets from environments with the highest  
66 proportion of ARGD-annotated sequence reads were the three human feces datasets from  
67 *Japan 29 Healthy Human Gut* project (5.6% to 5.3% of all reads; Sanger sequencing), mouse  
68 gut (4.6% and 4.3% of reads; Sanger sequencing) and activated sludge (4.3% to 4.2% of  
69 reads; Sanger sequencing) (Figure 1). Human feces metagenomes obtained with Roche 454  
70 pyrosequencing technology had less than 1% ARGD relative abundance. Sequencing  
71 technology can significantly influence annotation and need to be taken into consideration  
72 when looking at metagenomic data. The longer reads obtained by Sanger were easier to  
73 annotate, thereby, increasing their relative annotated read frequency, but such datasets  
74 have lower read counts compared to high-throughput next generation sequencing  
75 technologies, which are harder to annotate due to their relatively shorter reads such as  
76 those obtained from Roche 454 [1, 3, 16, 17]. Still, Sanger-sequenced human feces datasets  
77 contained relatively more ARGD-annotated reads than environmental datasets from oceans  
78 or farm soil also obtained with Sanger. All three Sanger-sequenced datasets had higher  
79 relative numbers of ARGD-annotated reads than cow gut, chicken gut and six ocean  
80 datasets, which were obtained by Roche 454. These Roche 454-sequenced datasets are also  
81 those with the shortest mean read length (about 100bp) among the studied metagenomes.  
82 The ARGD relative diversity found in each environment was estimated as a percentage of the  
83 2999 different sequences from the ARDB detected in the corresponding sequence datasets  
84 (Figure 1). Overall, the percentage of different ARGD sequences detected ranged from 0.97%  
85 in cow gut (4441680.3) to 32.78% in Rothamsted soil E41. Our results showed that soil had  
86 the highest level of ARGD diversity with an average of  $28.8\% \pm 3.14\%$  of the database  
87 sequences found in the soil datasets (max. 32.8% in Rothamsted E41 and min. 22.5% in  
88 Waseca Farm Soil) (Figure 1). Datasets from cow rumen, chicken cecum and arctic snow had  
89 the lowest number of different ARGD sequences with only 1.3% of the ARDB database  
90 sequences detected in the cow rumen, chicken cecum and arctic snow metagenomic  
91 datasets. Metagenomic datasets from human feces from different origins and sequencing  
92 technologies had an average of  $12.7\% \pm 2.1\%$  of the different ARDB database sequences  
93 (max. 16.17% in Twin Study 4440616.3 and min. 9.3% in Twin Study 4440614.3). This

94 observation is consistent with the general trend of a reduced diversity level of antibiotic  
95 resistance genes in non soil environments compared to soil (Figure 1).

96 **Most prevalent resistance types in the environment.**

97 The *macab* and *bcr* efflux pumps accounted for 18.9% and 8.5% of the total ARGD hits in all  
98 environments, respectively (Table 1, details available as supplementary Table S2 and Figure  
99 S1). Most ARGD-annotated reads detected in metagenomic datasets using our stringent  
100 annotation cut-off were related to various types of efflux-pump components: membrane  
101 fusion protein from resistance-nodulation cell division efflux pumps (RND) and ATP-binding  
102 cassette (ABC) domains related to bacitracin efflux with 27.9% and 20.8% of the total hits of  
103 ARGD annotated sequences in all metagenomes, respectively (Table 1, see also  
104 supplemental Table S2). Tetracycline ribosomal protection protein encoding genes (*tet\_rpp* ;  
105 5.7%) and vancomycin resistance related genes (*van a* : 3.6% ; *van b* : 2.9% ; *van d* : 3.5% ;  
106 *van e* 2.3% ; *van g* 2.8% ; *van c* 1.8%) were also found. Penicillin binding protein (*Pbp*)  
107 accounted for 5.4% of annotated reads.  $\beta$ -lactamases were less frequently detected (class A  
108  $\beta$ -lactamases, *bla\_a* : 3.4%; class B  $\beta$ -lactamases, *bla\_b*: 0.2% ; class C  $\beta$ -lactamases, *bla\_c*:  
109 0.42% ; class D  $\beta$ -lactamases, *bla\_d*: 0.04%) (Table S2).

110

Resistance Class	Mechanism of resistance	Antibiotic specificity
mexef, ceo, mexvw, acr, mexhi, mexcd, mexab, mdtnop, amr, adeabc, smeabc, smedef, mdtef, mexxy, mdtk	RND class transporter	Multi-drug resistance efflux
macab	RND class transporter : Macrolide	
bcr, bcr_mfs	ABC class transporter system : Bacitracin	
mls_abc	ABC class transporter : Macrolide	
mls_mfs, mls_hdr	MFS class transporter : Macrolide	
cml	MFS class transporter : Chloramphenicol	
rosab	Potassium antiporter system	
mepa, norm	MATE transporter	
tcma, mdr, qac	MFS transporter	
vana, vanb, vanc, vand, vane, vang	Vancomycin resistance operon genes (vanH, vanS, vanR, vanX, vanY) for each vancomycin resistance operons : VanA- ;VanB- ;VanC- ;VanD- ;VanE- and VanG-types	Vancomycin
tet_rpp	Tetracycline ribosomal protection protein	Tetracycline
tet_efflux	Tetracycline-specific efflux pump	
tet_flavo	Flavoproteins resistance to Tetracycline	
bla_a, bla_b	Class A and Class B $\beta$ -lactamases	$\beta$ -lactams
pbp	Penicillin-binding protein	
baca	Bacitracin resistance	Bacitracin
cata	Chloramphenicol acetyltransferase	Chloramphenicol
ksga	Kasugamycin resistance	Kasugamycin
arna	Polymixin resistance	Polymixin
pur8	Puromycin resistance	Puromycin
vat	Virginiamycin resistance	Streptogramin
sul	Sulfonamide resistance	Sulfonamide
dfra	Trimethoprim resistance	Trimethoprim

111 **Table 1. – Details for the 50 most prevalent resistance types found in all metagenomes.**

112 Most prevalent resistance type found in our analysis are MDR efflux pumps, vancomycin resistance,  
 113 tetracycline resistance,  $\beta$ -lactams resistance and resistance to other antibiotic molecules: Polymixin ;  
 114 Bacitracin ; Sulfonamide ; Kasugamycin ; Puromycin ; Trimethoprim ; Streptogramin ;  
 115 Chloramphenicol. Resistance mechanisms involved and antibiotic molecule targets are specified for  
 116 each group of the 50 resistance classes found and their respective abundance is indicated in  
 117 supplemental data table S2 and figure S1.  
 118

119 **ARGD distribution shows distinct environmental clusters.**





120 In order to overcome the bias related to the sequencing method for ecosystem comparison,  
121 an additional ARGD distribution analysis based on a clustering method was carried out on a  
122 restricted subset of metagenomic sequences obtained by Roche 454 pyrosequencing. Sanger  
123 datasets were insufficiently homogeneous, and therefore, were discarded from the analysis.  
124 The dendrogram obtained was plotted using Ward hierarchical classification of the datasets  
125 Mountford index (see supplemental Experimental Procedures). The Mountford index reflects  
126 similarities between samples based on ARGD sequence presence-absence for each  
127 metagenome datasets (*i.e.* ARGD  $\beta$ -diversity is defined as a function of alpha- and gamma-  
128 diversities). The different datasets used in the analysis were clustered in three different  
129 groups that matched their environmental origin: ocean group, soil group and human feces  
130 group (Figure 2). Soil datasets were similar regardless of the soil geographical origin or usage  
131 (a prairie soil in England or a forest soil in Puerto Rico) whereas two separate clusters were  
132 identified for ocean datasets. Open ocean datasets grouped separately from coastal ocean  
133 datasets. The Venn diagram of resistance class level annotations showed a common set of  
134 antibiotic resistance genes shared between all three environmental clusters (Figure 3, see  
135 also Table S3). Among all different ARGD types (n=94), a majority (n=50) was common to all  
136 environments and less than 10% (n=6) were not found anywhere. The six types of ARGD not  
137 detected in any cluster were: *dfrb* (group B trimethoprim resistant dihydrofolate reductase),  
138 *fusb* (fusaric acid resistance), *mph* (macrolide phosphotransferase); *rif* (viral rifampin  
139 resistance), *vph* (viomycin phosphotransferase) and *tet\_other* (tetracycline resistance genes  
140 that are not efflux pumps or ribosomal binding proteins). Ocean and human feces clusters  
141 did not share any ARGD type. Multidrug efflux pump class *Imrp* was found only in the human  
142 feces cluster. ARGDs shared between soil and ocean clusters (n=3) were genes conferring  
143 resistance to fosfomycin, puromycin and quinolones. A large subset (n=24) of ARGD classes  
144 was shared between human feces and soil clusters. This subset includes all classes of  $\beta$ -  
145 lactamases, resistances to aminoglycoside through antibiotic molecules modification  
146 (phosphotransferases, acetyltransferases or nucleotidyltransferases), various multidrug  
147 efflux pumps and flavoproteins conferring resistance to tetracycline.

148

149 **DISCUSSION:** The objective of our study was to evaluate the quantity and diversity of  
150 antibiotic resistance genes from different environments using the metagenomic datasets  
151 from various biomes (soils, ocean, human gut, etc.). Metagenomic sequences were

152 annotated as ARGD if their sequence matches a set of known antibiotic resistance genetic  
153 determinants (ARDB). Our results show that soil environments have the largest apparent  
154 diversity of ARGDs while gut-associated microbiota have a higher proportion of ARGD-  
155 annotated reads, which may be due to exposure to antibiotic molecules that increased the  
156 overall antibiotic resistance of the commensal flora [1, 3, 17, 18].

157 Efflux pumps dominated the best-hit-only annotation (*i.e.*, not removing best-hits  
158 below a bit-score and percent identity threshold of respectively above 60% and 35%).  
159 Antibiotic-specific efflux pumps share homologies at the amino-acid level with other pump  
160 genes and share functional domains with every efflux pump protein (*e.g.*, membrane-fusion  
161 protein (MFP) domain). Efflux pumps are present in both antibiotic-susceptible and  
162 antibiotic-resistant bacteria and can be poor markers of resistance phenotype. Efflux-pump-  
163 related resistance is often conferred by point-mutations in their regulatory genes [5, 18, 19].  
164 However, in our analysis, most detected efflux-pump reads were related to efflux of  
165 antibiotic or toxic compounds even when re-examined with NCBI\_nr (Figure S3). This implied  
166 that the antibiotic resistance hit was not superseded by a hit with generalist database. Some  
167 ARGD protein families with many different variants, such as  $\beta$ -lactamases or efflux pumps,  
168 were more abundantly referenced in our database and that might have increased detection  
169 potential compared to ARGD with few representatives in the ARDB. ARGDs, such as  
170 Vancomycin resistance operons,  $\beta$ -lactamases or tetracycline ribosomal protection proteins,  
171 usually confer the resistance phenotype directly when translated. This makes these ARGD-  
172 annotated reads better markers of potential antibiotic resistance phenotypes. ARGDs were  
173 systematically detected in every environment screened despite all the described limitations.  
174 ARGDs are probably even more abundant in all of the ecosystems than what has been  
175 observed with sequence-based analysis. Thus, we believe our method retrieved many *bona*  
176 *fide* ARGDs from the environmental resistome.

177 Metagenomes obtained with the same technology form three distinct environmental  
178 groups (ocean, soil and human feces) on the basis of presence/absence of annotated ARGD  
179 sequences (Figure 2). While many ARGDs were shared between all environments (Figure 3),  
180 soil and human feces shared more (24 out of 94) ARGD classes (*i.e.* groups of sequences with  
181 shared antibiotic resistance-conferring mechanisms as defined in ARDB [13]) than any other  
182 pairs of ecosystems. This heightened degree of shared ARGDs between soil and human feces  
183 datasets has been previously explored by functional metagenomics and lead to the

184 conclusion that gene flow between these environments is likely [4, 5, 19]. Soil is probably  
185 the major source of antibiotic-producing soil microorganisms (*e.g.*, *Streptomyces*  
186 *chrysogenum* producing penicillin) and they live in apparently the most diverse environment  
187 in terms of ARGDs (Figure 1). Considering this diversity, the environment and more  
188 specifically the soil reservoir of antibiotic resistance genes could be major contributors to  
189 antibiotic resistance mechanism diversification in pathogen populations by horizontal gene  
190 transfer [1].

191         The potential ecological pressure that low natural concentrations of antibiotics might  
192 exert has important implications concerning the role of antibiotics molecules and ARGDs in  
193 bacterial ecology. In environments without anthropogenic amendment of antibiotics  
194 molecules, such as the Rothamsted soil or the Alaskan permafrost, antibiotic molecule  
195 concentrations were not detected, yet functional antibiotic resistance genes were always  
196 detected [4, 5, 20, 21]. In addition to their antibacterial activity, antibiotic molecules have  
197 been found to act as signal molecules between bacterial cells [8, 20-24] and many antibiotic  
198 resistance genes are themselves involved in diverse bacterial physiological functions [8, 22-  
199 27]. Their possible multifunctional role and their low fitness costs (and eventual fitness  
200 benefits independent of antibiotic resistance [25]) might explain the prevalence of most  
201 ARGDs in the environment. Indeed, maintenance in their host may not require a constant  
202 selective pressure, making environmental bacteria a persistent source of ARGDs for  
203 pathogens [25-27]. Human feces datasets showed that antibiotic resistance genes were  
204 present in human gut resident bacteria with as much as 5.6% ARGD-annotated reads. In  
205 human gut, other mechanisms of antibiotic resistance gene selection besides inhibition  
206 resistance could be involved. The most abundant ARGD-annotated reads detected in human  
207 datasets were vancomycin resistance genes even though vancomycin is a strongly regulated  
208 and rarely used "last resort" antibiotic in medicine. In contrast, Amoxicillin related resistance  
209 genes ( $\beta$ -lactamases) were much less abundant in human feces, although this drug is  
210 commonly used.

211         The ecological role of every ARGD in the environment has yet to be elucidated, but  
212 some ARGDs have been shown to play roles, such as inter-cellular signaling, other than  
213 antibiotic protection for their original (*i.e.* not clinical) saprophyte host [2, 25-27]. Antibiotic  
214 molecules and antibiotic resistance are an inherent part of common bacterial eco-physiology  
215 in the environment. This ARGD resource is selected in pathogens when they acquire this trait

216 and enhance their fitness in the case of antibiotherapy. Yet, not all common pathogens are  
217 resistant to antibiotics. With the efficiency of bacteria to adapt to the antibiotic challenge via  
218 mutations, DNA recombination, plasmid- and integron- mediated gene transfer, they are  
219 outpacing the health industry in terms of development and innovation. Further studies on  
220 the mechanism of antibiotic resistance dispersion (*e.g.*, role of mobile genetic elements) are  
221 critical for enhancing antibiotic clinical lifetimes. Antibiotic resistance genes in the  
222 environment (and clinical antibiotic resistance) are not likely to disappear in the near future,  
223 although selection mechanisms seem to limit dissemination of antibiotic resistance.  
224 Deciphering these mechanisms may provide the tools for enhancing the long-term efficiency  
225 of currently available and future antibiotics.  
226

227 **ACKNOWLEDGMENTS:**

228 We thank Dr. Annie Chaboud for critical reading of the revised manuscript. We thank Penny  
229 Hirsch and Ian Clark from Rothamsted Research (UK) for the soil used in the ANR project  
230 METASOIL. This work was financed by grants from the French Agency for the Environment  
231 and Energy Management (ADEME – project “Générique”) and from French Agency for Food,  
232 Environmental and Occupational Health & Safety (ANSES - project “Impactance”,  
233 convention n° EST-2012/2/28). The French Ministry of Higher Education and Research  
234 funded JN. TOD was supported by a doctoral fellowship from the Rhône Alpes Région

## 235 REFERENCES :

- 236 1. Ochman, H., Lawrence, J. G., and Groisman, E. A. (2000). Lateral gene transfer and the  
237 nature of bacterial innovation. *Nature* 405, 299–304.
- 238 2. Cantón, R., and Morosini, M.-I. (2011). Emergence and spread of antibiotic resistance  
239 following exposure to antibiotics. *FEMS Microbiol Rev* 35, 977–991.
- 240 3. Lawrence, J. G., and Roth, J. R. (1996). Selfish operons: horizontal transfer may drive  
241 the evolution of gene clusters. *Genetics* 143, 1843–1860.
- 242 4. Allen, H. K., Moe, L. A., Rodbumrer, J., Gaarder, A., and Handelsman, J. (2009).  
243 Functional metagenomics reveals diverse beta-lactamases in a remote Alaskan soil.  
244 *ISME J* 3, 243–251.
- 245 5. Demanèche, S., Sanguin, H., Poté, J., Navarro, E., Bernillon, D., Mavingui, P., Wildi, W.,  
246 Vogel, T. M., and Simonet, P. (2008). Antibiotic-resistant soil bacteria in transgenic  
247 plant fields. *Proc Natl Acad Sci USA* 105, 3957–3962.
- 248 6. Bhullar, K., Waglechner, N., Pawlowski, A., Koteva, K., Banks, E. D., Johnston, M. D.,  
249 Barton, H. A., and Wright, G. D. (2012). Antibiotic resistance is prevalent in an isolated  
250 cave microbiome. *PLoS ONE* 7, e34953.
- 251 7. D'Costa, V. M., King, C. E., Kalan, L., Morar, M., Sung, W. W. L., Schwarz, C., Froese, D.,  
252 Zazula, G., Calmels, F., Debruyne, R., et al. (2011). Antibiotic resistance is ancient.  
253 *Nature* 477, 457–461.
- 254 8. Martinez, J. L., Fajardo, A., Garmendia, L., Hernandez, A., Linares, J. F., Martínez-  
255 Solano, L., and Sánchez, M. B. (2009). A global view of antibiotic resistance. *FEMS*  
256 *Microbiol Rev* 33, 44–65.
- 257 9. Humeniuk, C., Arlet, G., Gautier, V., Grimont, P., Labia, R., and Philippon, A. (2002).  
258 Beta-lactamases of *Kluyvera ascorbata*, probable progenitors of some plasmid-  
259 encoded CTX-M types. *Antimicrob Agents Chemother* 46, 3045–3049.
- 260 10. Toleman, M. A., Spencer, J., Jones, L., and Walsh, T. R. (2012). BlaNDM-1 is a chimera,  
261 likely constructed in *Acinetobacter baumannii*. *Antimicrob Agents Chemother*.
- 262 11. Aminov, R., and Mackie, R. (2007). Evolution and ecology of antibiotic resistance  
263 genes. *FEMS Microbiol Lett* 271, 147–161.
- 264 12. Eisen, J. A. (2007). Environmental Shotgun Sequencing: Its Potential and Challenges  
265 for Studying the Hidden World of Microbes. *PLoS Biol* 5, e82.
- 266 13. Liu, B., and Pop, M. (2009). ARDB--Antibiotic Resistance Genes Database. *Nucleic*  
267 *Acids Res* 37, D443–7.
- 268 14. Looft, T., Johnson, T. A., Allen, H. K., Bayles, D. O., Alt, D. P., Stedtfeld, R. D., Sul, W. J.,  
269 Stedtfeld, T. M., Chai, B., Cole, J. R., et al. (2012). In-feed antibiotic effects on the

- 270 swine intestinal microbiome. *Proc Natl Acad Sci USA* *109*, 1691–1696.
- 271 15. Overbeek, R. A., Begley, T., Butler, R. M., Choudhuri, J. V., Chuang, H.-Y., Cohoon, M.,  
272 de Crécy-Lagard, V., Diaz, N., Disz, T., Edwards, R. A., et al. (2005). The subsystems  
273 approach to genome annotation and its use in the project to annotate 1000 genomes.  
274 *Nucleic Acids Res* *33*, 5691–5702.
- 275 16. Wommack, K. E., Bhavsar, J., and Ravel, J. (2008). Metagenomics: read length matters.  
276 *Appl Environ Microbiol* *74*, 1453–1463.
- 277 17. Forslund, K., Sunagawa, S., Roat Kultima, J., Mende, D., Arumugam, M., Typas, A., and  
278 Bork, P. (2013). Country-specific antibiotic use practices impact the human gut  
279 resistome. *Genome Res*.
- 280 18. Piddock, L. J. V. (2006). Clinically relevant chromosomally encoded multidrug  
281 resistance efflux pumps in bacteria. *Clin Microbiol Rev* *19*, 382–402.
- 282 19. Forsberg, K. J., Reyes, A., Wang, B., Selleck, E. M., Sommer, M. O. A., and Dantas, G.  
283 (2012). The Shared Antibiotic Resistome of Soil Bacteria and Human Pathogens.  
284 *Science* *337*, 1107–1111.
- 285 20. Aminov, R. I. (2009). The role of antibiotics and antibiotic resistance in nature. *Environ*  
286 *Microbiol* *11*, 2970–2988.
- 287 21. Fajardo, A., and Martínez, J. L. (2008). Antibiotics as signals that trigger specific  
288 bacterial responses. *Curr Opin Microbiol* *11*, 161–167.
- 289 22. Wellington, E. M. H., Boxall, A. B., Cross, P., Feil, E. J., Gaze, W. H., Hawkey, P. M.,  
290 Johnson-Rollings, A. S., Jones, D. L., Lee, N. M., Otten, W., et al. (2013). The role of the  
291 natural environment in the emergence of antibiotic resistance in gram-negative  
292 bacteria. *Lancet Infect Dis* *13*, 155–165.
- 293 23. Alonso, A., Sanchez, P., and Martinez, J. L. (2001). Environmental selection of  
294 antibiotic resistance genes. *Environ Microbiol* *3*, 1–9.
- 295 24. Allen, H. K., Donato, J., Wang, H. H., Cloud-Hansen, K. A., Davies, J., and Handelsman,  
296 J. (2010). Call of the wild: antibiotic resistance genes in natural environments. *Nat Rev*  
297 *Microbiol* *8*, 251–259.
- 298 25. Andersson, D. I., and Hughes, D. (2010). Antibiotic resistance and its cost: is it possible  
299 to reverse resistance? *Nat Rev Microbiol* *8*, 260–271.
- 300 26. Andersson, D. I., and Levin, B. R. (1999). The biological cost of antibiotic resistance.  
301 *Curr Opin Microbiol* *2*, 489–493.
- 302 27. Hall, A. R., Iles, J. C., and MacLean, R. C. (2011). The fitness cost of rifampicin  
303 resistance in *Pseudomonas aeruginosa* depends on demand for RNA polymerase.  
304 *Genetics* *187*, 817–822.

305

306 **FIGURE LEGENDS:**

307

308 **Figure 1. – Relative diversity and abundance of ARGD between metagenomes.**

309 Barplot showing the relative abundance of reads annotated as ARGD (gray bars), expressed as  
 310 number of ARGD hits divided by total number of reads in the dataset and ARGD diversity (white  
 311 bars), expressed as the percentage of the 2999 unique sequences in ARDB database found in each  
 312 dataset. Shaded boxes indicate a different environmental origin with from left to right: Antarctic  
 313 lakes; Arctic snow; chicken gut; cow gut; deep oceans; human feces; microbial fuel-cell; mouse gut;  
 314 ocean; halophile sediment; activated sludge and soil. Black dots indicate datasets obtained by Sanger  
 315 sequencing method, others are obtained by Roche 454 sequencing.

316

317 **Table 1. – Details for the 50 most prevalent resistance classes found in all metagenomes.**

318 Resistance classes are groups of genes implicated in the same resistance phenotype as defined in  
 319 ARDB (ardb.cbcb.umd.edu). Most prevalent resistance classes found in our analysis are MDR efflux  
 320 pumps, vancomycin resistance, tetracycline resistance,  $\beta$ -lactams resistance and resistance to other  
 321 antibiotic molecules: Polymixin ; Bacitracin ; Sulfonamide ; Kasugamycin ; Puromycin ; Trimethoprim ;  
 322 Streptogramin ; Chloramphenicol. Resistance mechanisms involved and antibiotic molecule targets  
 323 are specified for each group of the 50 resistance classes found and their respective abundance is  
 324 indicated in supplemental data table S2 and figure S1.

325

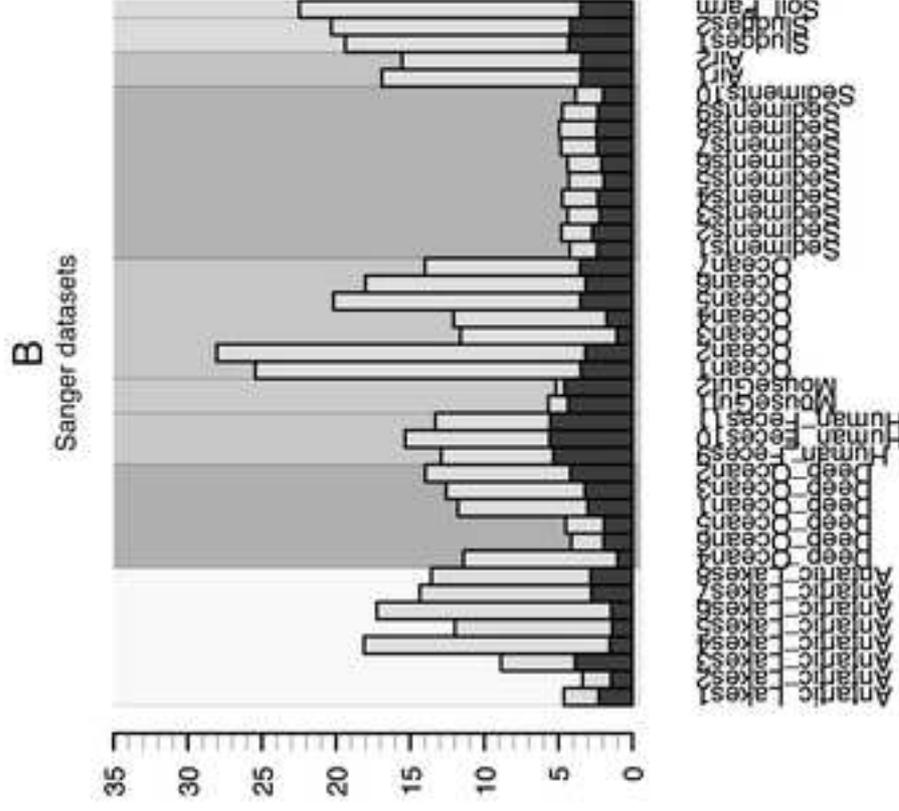
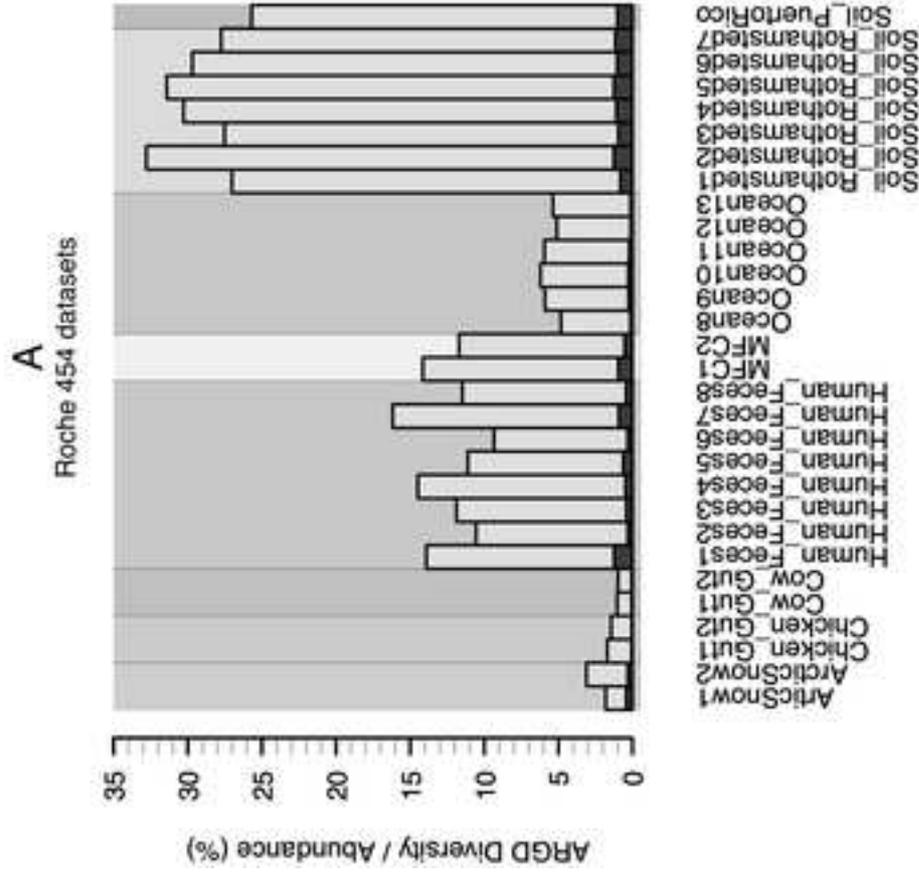
326 **Figure 2. – Hierarchical clustering based on Mountford diversity index of ARGD among soil, ocean  
and human feces Roche 454 metagenomic datasets.**

327 Mountford index has been computed from the distribution of ARGD hits in each metagenomic  
 328 dataset considering each different ARGD as a different species. Hierarchical clustering based on ward  
 329 method has been performed on Mountford index matrix. Environmental clusters are highlighted with  
 330 colors: open ocean (dark blue); coastal ocean (light blue); soil (brown); human feces (red).

331

332 **Figure 3. – Common ARGD class between four different environmental clusters based on Roche  
454 Datasets.**

333 Numbers inside each circle intersection are indicating number of common resistance classes  
 334 between environmental datasets. **Only in soil ( $\alpha$ ):** *mls\_hdr*, *pmra*, *qacab*, *qnr*, *tmr*, *ble*, *bmr*, *emea*,  
 335 *ere*, *fush*. **Common to Soil and Ocean only  $\beta$ :** *fos* (fosfomycin resistance) ; *pac* (puromycin resistance)  
 336 ; *mfpa* (quinolones resistance). **Common to Soil, Ocean and Human Feces ( $\gamma$ ):** *acr*, *adeabc*, *amr*,  
 337 *macab*, *mdtef*, *mdtk*, *mdtm*, *mdtnop*, *mexab*, *mexcd*, *mexef*, *mexhi*, *mexvw*, *mexxy*, *smeabc*, *smedef*  
 338 (RND efflux pumps), *arna* (polymixin resistance),  *Baca*, *bcr*, *bcr\_mfs* (bacitracin resistance), *catb*, *ceo*,  
 339 *cml* (chloramphenicol resistance), *dfra* (group A trimethoprim resistant dihydrofolate reductase),  
 340 *emrd*, *emre*, *mdr*, *nora*, *norm*, *rosab*, *tcm* (MFS transporters), *erm*, *lsa* (macrolide resistance), *ksga*  
 341 (kasugamycin resistance), *meca*, *pbp* ( $\beta$ -lactams resistance), *mls\_abc* (ABC family transporter), *pur8*  
 342 (puromycin resistance), *qac* (quaternary ammonium compound efflux pump), *sul* (sulfonamide-  
 343 resistant dihydropteroate synthase), *tet\_efflux*, *tet\_xprt* (tetracycline efflux pumps), *tet\_rpp*  
 344 (tetracycline ribosomal protection protein), *vana*, *vanb*, *vanc*, *vand*, *vane*, *vang* (vancomycin  
 345 resistance), *vat* (virginiamycin A acetyltransferase). **Common to Soil and Human Feces ( $\delta$ ):** *aac*, *ant*,  
 346 *aph* (aminoglycoside transferases), *bla\_a*, *bla\_b*, *bla\_c*, *bla\_d* ( $\beta$ -lactamases), *blt*, *cata*  
 347 (chloramphenicol acetyltransferase), *lnu* (lincomycin nucleotidyltransferase) , *mdfa*, *mdtg*, *mdth*,  
 348 *mdtl*, *mepa*, *mls\_mfs*, *ykk* (Multidrug efflux pump), *mecr1* (methicillin resistance), *sta*, *str*  
 349 (streptomycin resistance), *tet\_flavo*, *tet\_mod* (resistance to tetracycline), *vanz* (*vancomycin*  
 350 *resistance*) and *tsnr* (*thiosptrepton* resistance). **Only in Human Feces ( $\epsilon$ ):** *lmp* (MFS transporter). **Not**  
 351 **detected ( $\zeta$ ):** *dfrb* (group B trimethoprim resistant dihydrofolate reductase), *fusb* (fusaric acid  
 352 resistance), *mph* (macrolide phosphotransferase); *rif* (viral rifampin resistance), *vph* (viomycin  
 353 phosphotransferase) and *tet\_other* (tetracycline resistance genes that are not efflux pumps or  
 354 ribosomal binding proteins).  
 355  
 356



# Cluster Dendrogram

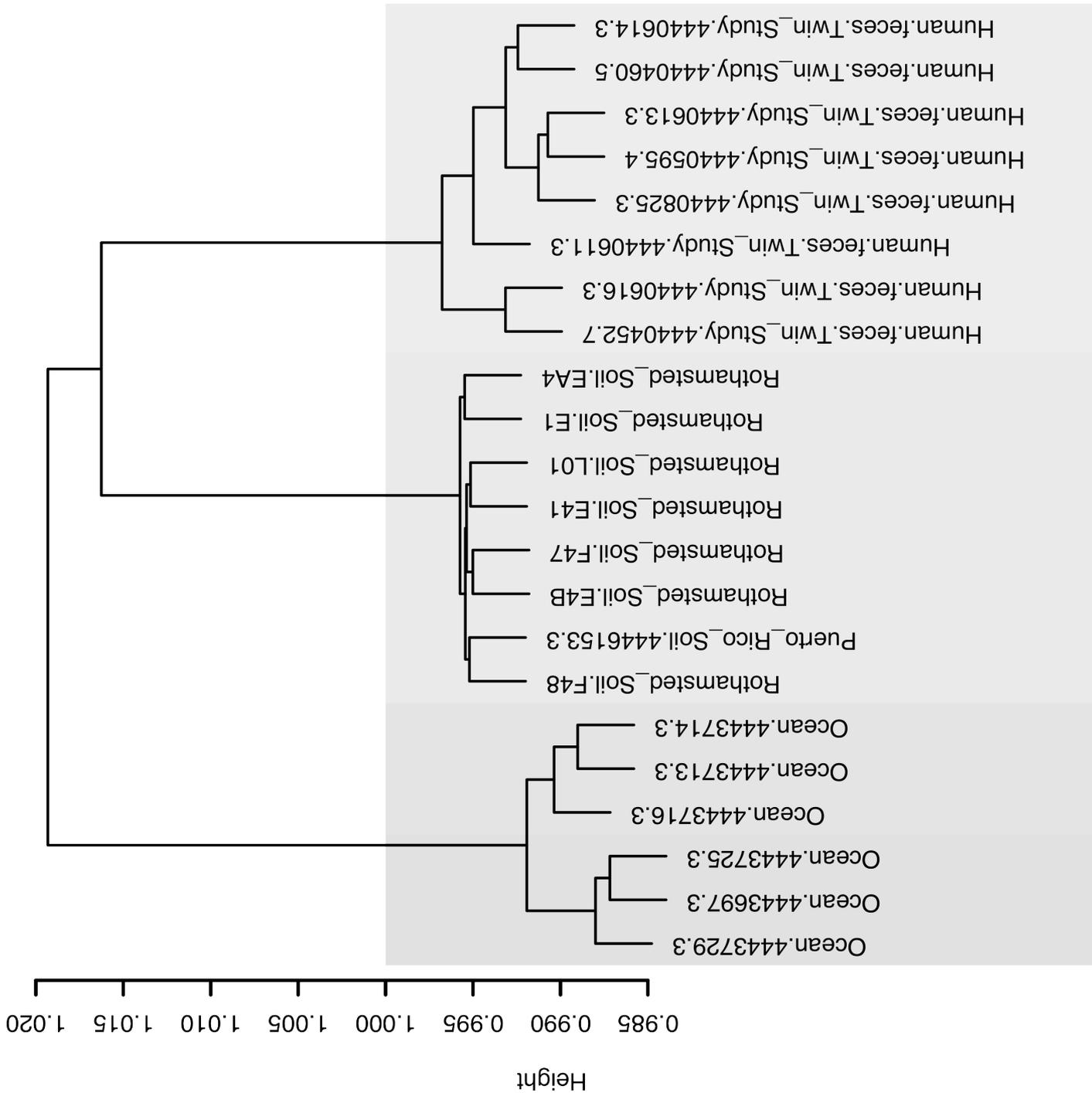
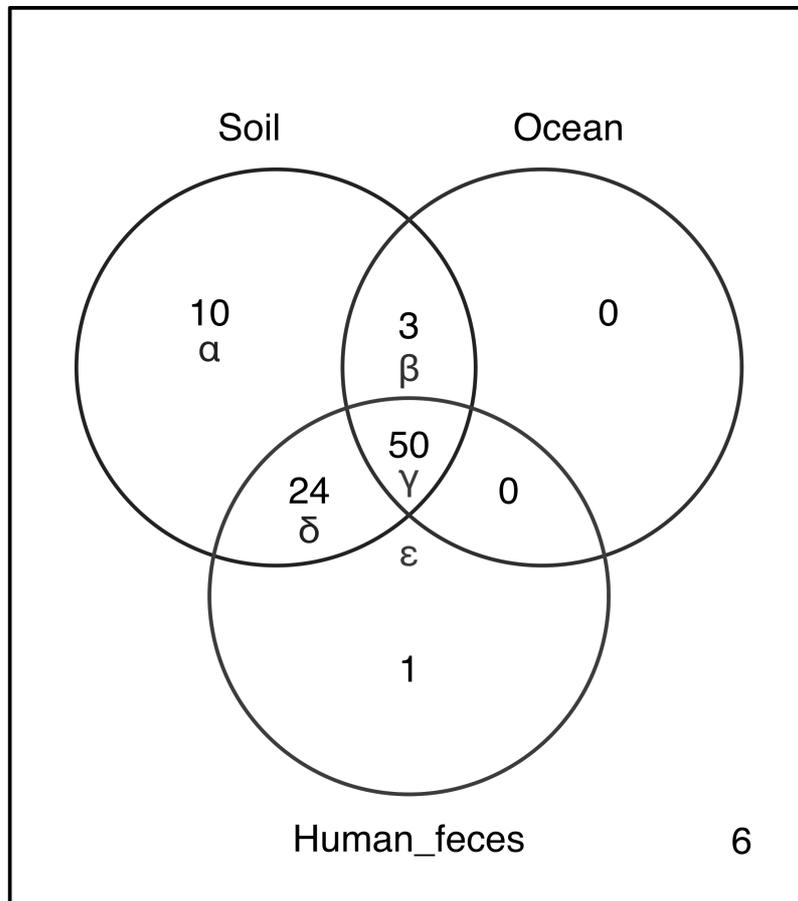


Figure 3





**The soil resistome: a critical review on antibiotic resistance origins, ecology and dissemination potential in telluric bacteria.**

Journal:	<i>Environmental Microbiology and Environmental Microbiology Reports</i>
Manuscript ID:	EMI-2014-0821
Manuscript Type:	EMI - Minireview
Journal:	Environmental Microbiology
Date Submitted by the Author:	23-Jun-2014
Complete List of Authors:	Nesme, Joseph; Ecole Centrale de Lyon, Laboratoire Ampere Simonet, Pascal; Ecole Centrale de Lyon, Laboratoire Ampère, CNRS
Keywords:	antibiotic resistance, horizontal gene transfer, metagenomics/community genomics, resistome, soil

SCHOLARONE™  
Manuscripts

Wiley-Blackwell and Society for Applied Microbiology

**The soil resistome: a critical review on antibiotic resistance origins, ecology and dissemination potential in telluric bacteria.**

Joseph Nesme and Pascal Simonet

Environmental Microbial Genomics, Laboratoire Ampère, CNRS, Ecole Centrale de Lyon, Université de Lyon, 36 Avenue Guy de Collongue, 69134 Ecully, France.

**SUMMARY:** Soil is a large reservoir of microbial diversity and the majority of antimicrobial compounds used today in human and veterinary healthcare have been isolated from soil microorganisms. The Darwinian hypothesis of an “arms-shields race” between antibiotic producers and resistant strains is often cited to explain antibiotic resistance gene determinants (ARGD) origins and diversity. ARGD abundance and antibiotic molecule exposure are, however, not systematically linked and many other factors can contribute to resistance gene emergence, selection and dissemination in the environment. Soil is a heterogeneous habitat and represents a broad spectrum of different ecological niches. Soil harbors a large genetic diversity at small spatial scale, favoring exchange of genetic materials by means of horizontal gene transfer (HGT) that will contribute to ARGD dissemination between bacteria and eventually acquisition by pathogen genomes therefore threatening antibiotic therapies. Our current knowledge on the extent of the soil resistome abundance and diversity has been greatly enhanced with since metagenomic revolution and the help of high-throughput sequencing technologies. Different ecological hypotheses explaining their high prevalence in soil and questioning their transfer rate to pathogens, in respect to these recent experimental results will be discussed in the present review.

**KEYWORDS:** antibiotic resistance; horizontal gene transfer; metagenomic; resistome; soil.

## Introduction

Since their introduction in medicine to treat millions of people, antibiotics have considerably reduced infectious diseases casualties and may account for the biggest medical revolution of the 20<sup>th</sup> century. Penicillin, one of the first widely used antibiotics, was isolated in 1929 by Alexander Fleming (Fleming, 1929) and introduced onto the market in 1938, in time to significantly reduce the number of deaths caused by infected wounds during World War II (Quinn, 2013). The antibiotics golden age lasted from the 1940s to the 1990s when most antibiotic molecules still used today were discovered and introduced into the medical market. Most of antibiotics are natural bioactive compounds synthesized, either by Bacteria (mostly Actinobacteria such as *Streptomyces* spp.) or Fungi, with for most of them a soil origin (Figure 1) (Waksman, 1941; Bérdy, 2005). A more fundamental aspect of soil microbial ecology concerning studies of antagonist interactions between soil microorganisms was at the root of pioneers' researches on antibiotics drug discovery (I. M. Lewis, 1929; Waksman and Woodruff, 1940; Allen et al., 2010). The immediate success of these molecules in the early age of antibiotic therapy was directly linked to pathogen sensitivity that dramatically reduced infectious diseases mortality and morbidity (Florey, 1946; Smith, 2000). The increasing resistance of pathogenic strains seen in hospitals, however, very quickly hindered this victory: first observation of penicillin-resistant *Staphylococcus* spp. occurred as early as 1945 (Plough, 1945). More concerning, resistance seems unavoidable since resistant pathogens are always quickly observed in clinical environment when introducing new antibiotics molecules (K. Lewis, 2013). But little is known on antibiotic resistance genes' ecology *in situ*, in their natural hosts. According to the World Health Organization (WHO), antibiotic resistance is defined by "an increase in the minimum inhibitory concentration of a compound for a previously sensitive strain" (WHO

Fact Sheet N°194). This definition of antibiotic resistance is *ad-hoc* for clinical microbiology but neglects concepts of intrinsic resistance (Cox and Wright, 2013) or resistance genes naturally occurring in the environment, such as in antibiotic producers for example (Benveniste and Davies, 1973; Allen et al., 2010).

### **What is the extent of ARGD diversity in nature and what is the share of currently known ones?**

In a broader view of microbial ecology, we still have to elucidate the origins and selection processes involved in dissemination of the diverse mechanisms of antibiotic resistance that we are currently observing in clinical and environmental settings. Findings on antibiotic resistance genes ecology *in situ* could have considerable impact on clinical microbiology as well as being a fantastic context for studying traits-adaptation in bacterial population. Are resistance genes a result of rapid adaptation through gene shuffling or mutation? Or are they rather acquired through horizontal gene transfer (HGT)? Are most ARGD acquired by pathogens from non-pathogenic saprophyte soil microbial bacteria lineages? This evolutionary scenario is commonly cited and admitted as one of the most probable source of antibiotic resistance introduction in the clinic (Fleming, 1932; Benveniste and Davies, 1973; Martinez, 2008). Can this chain of events be affected by other factors than the sole antibiotic molecule selective pressure? Urbanization factors such population density and building density are for example shown to be positively correlated in soil samples with integron-related genes frequently implied in transfer of antibiotic resistance genes (Nardelli et al., 2012; Gillings, 2013). What is the fitness-benefit for producing antibiotic molecules, or resisting it, in natural environments with intense competition such as soil? Secreting antibiotic substances for interference competition that is the specific inhibition of

neighboring cells feeding on the same substrate as a producer for example, is indeed often hypothesized to be a strongly selected adaptive trait of bacteria living in oligotrophic environment, such as soil. Data gathered in environmental microbiology studies show that ARGD have been found in diverse environmental samples, such as soil (Demanèche et al., 2008; Allen et al., 2009) or oceanic cold-seep sediments (J. S. Song et al., 2005) but also from pristine environment predating the antibiotic era (D'Costa et al., 2011; Bhullar et al., 2012). More recently, we have found that all environments contain genes with significant similarities with what is known in the clinic as antibiotic resistance genes, illustrating the ubiquity of ARGD in environmental samples (Nesme et al., 2014). In this study, soil is found to be the most prominent reservoir harboring as much as 30% of all different known ARGD found in sequence databases and those different environments have a different distribution of resistance genes. All these data nonetheless support the hypothesis of an extremely abundant environmental reservoir, especially in soil, but are also emphasizing on antibiotic resistance genes importance for bacterial eco-physiology at the ecosystem level. In other terms, these results suggest that antibiotic resistance and antibiotic production genes are essential for many bacteria, whatever their intrinsic function may be and despite we currently do not understand precisely which selective forces are at play in their dissemination, explaining their presence at significant abundance in all environments studied so far. Reservoirs of resistance also have critical implications for clinical microbiology and our understanding of such gene dissemination to pathogens. Studying resistant environmental bacteria with functional metagenomic methods (i.e.: expression of metagenome DNA sequence in a heterologous host), Forsberg *et al.* showed that a significant amount of resistance genes are shared between the soil and the gut microbiota and can confer resistance to a previously sensitive *Escherichia coli* host (D'Costa et al., 2011;

Forsberg et al., 2012). This shared resistome is also observed with metagenome sequence reads annotations described previously (Allen et al., 2009; Forsberg et al., 2012; Nesme et al., 2014). While not indicating the direction of transfer these studies show that commensal and pathogenic bacteria share antibiotic resistance genes with soil bacteria. Tracking directions of gene transfer is certainly from the most interesting research topics in evolutionary microbial ecology and such fundamental findings have direct implications for the future of antibiotic resistance management. Antibiotic resistance gene acquisition by lateral gene transfer, from different environments such as soil, oceans, etc., to pathogens genomes is still the most probable origin of resistance genes emergence in clinical environment. Experimental studies tracking confidently antibiotic-resistance gene reservoirs are still scarce and most often limited to cultivable microorganisms and/or well-defined classes of resistance genes. Such studies, while tracking down the most probable route from the environmental reservoir to pathogen host and subsequent dissemination are limited by prior knowledge of resistance in the clinic and will therefore miss any genes not similar to known antibiotic resistance genes and tested strains cultivability. In these studies, authors are searching for a discrepancy between gene localization (*i.e.* plasmid/other mobile element versus chromosome), habitat (*i.e.* pathogens or commensal versus soil or water-borne) for a given gene sequence (*i.e.* an allele of the *AmpC* gene family) in several hosts, considering these examples as signature of, respectively, clinical or environmental antibiotic-resistance. If authors could be considered lucky to perform such observations, considering the wide majority of not yet cultured microorganisms, these results obtained by cultivation approaches emphasize on the very high prevalence of antibiotic resistant microorganisms in the environment. In this respect, studying the environmental reservoir of antibiotic-resistance genes and its links to the clinic is of crucial importance and stresses on

the need for thorough environmental surveys of this situation in natural environments; but also in environments impacted by human-activities (Heuer and Smalla, 2007; Gaze et al., 2011; Nardelli et al., 2012).

The fast pace of antibiotic resistance increase in clinical environment (*i.e.*: every known bacterial pathogen has developed resistance to one or more antimicrobial molecules used in the clinic) suggests a pre-existing pool of antibiotic resistance genes in “natural” environmental reservoirs and a transferability of these genes. In this context, a major challenge of microbiologists is to track reservoirs of resistance genes and limits their dissemination to hosts where their expression is problematic, such as in pathogens infecting humans. Understanding factors driving selection and dissemination of environmental antibiotic resistance genes will considerably enhance our ability to prevent dissemination of antibiotic-resistance genes in the clinic. It is admitted that horizontal gene-transfer is a major mechanism of bacteria adaptation to clinical-antibiotic concentrations either through gene acquisition or modulation of resistance-mechanisms expression. This is even more evident when considering that the most potent antibiotic-resistance genes in pathogens are most often encoded on mobile genetic elements (Schlüter et al., 2007; Stokes and Gillings, 2011; Djordjevic et al., 2013).

#### **Antibiotic resistance, production and dissemination: old concepts and new tools.**

Deciphering the different chains of gene transfer events leading to antibiotic-resistance phenotype acquisition in previously sensitive pathogenic bacteria is a major challenge for microbial ecology. Untangling this network requires to carefully take into account its different aspects; from gene source to drastic host's microbiome population selection following antibiotic therapy. Obviously, natural selection is not acting on these traits with

the same fitness cost depending on gene host or environmental factors (Cohan, 2001). In addition, mobile genetic elements are not transferring similar gene loads or affects the same populations, for example due to plasmid incompatibility groups, phage specificity or the different nature of transfers (transformation, conjugation or transduction) as well as the environment considered, not sharing the same pool of resistance genes (Halary et al., 2010; Sen et al., 2011; Nesme et al., 2014). These theoretical statements can now be more thoroughly tested experimentally using recent advances in molecular biology, bioinformatics and increased sequence data-throughput, as in the above-cited papers. These technologies reaffirmed clearly that complex environmental matrix such as soil harbors a diversity far greater than what was previously expected; with more than thousands of different species found in a gram of soil (Torsvik, 2011). Combining data provided with these new tools and previously developed statistical algorithms for community ecology enabled a vast field of experimentations for microbial ecologists. In the context of antibiotic-resistance, a trait selected for resistance in the clinic may be costly for the same organism in the environment, in the absence of high antibiotic-molecule selective pressure, for example when such traits are encoded on high-copy number plasmids (De Gelder et al., 2008; Martinez, Fajardo, et al., 2009; Andersson and D. Hughes, 2010). *A contrario*, the same genes will be strongly selected in the clinic, due to strong and constant antibiotic molecules usage for antibiotic therapy increasing its relative abundance and lowering relative antibiotic resistance gene diversity. This explains why only a small, though largely disseminated, proportion of potent antibiotic resistance genes are found in clinical environments compared to soil or oceanic samples (J. S. Song et al., 2005; Bhullar et al., 2012; Nesme et al., 2014; Sommer, 2014). This is underlying the major role of horizontal gene transfers in the frame of antibiotic resistance genes ecology. Gene fluxes between

different hosts strongly shape microbial population structures and have considerable effects on general microorganisms ecology, adaptation and genetic diversification (Ochman et al., 2000; Abby et al., 2012). Antagonistic relationships driven by antibiotic-molecule secretion, and their respective resistances, are also shown to be a strong cohesion factor between related *Vibrio* spp. populations (Cordero et al., 2012). *Vibrio* spp. bacteria have complex integron machinery promoting genetic intra-population variability that can be mediated by antibiotic molecules, emphasizing on the strong impact of antibiotics resistance and production in this genus. (Cordero et al., 2012; Baharoglu et al., 2012). Such scenarios can explain how sympatric speciation of closely related bacteria sharing a common habitat may occur in the soil environment. Mutual exclusion of co-specific cells mediated by a different antibiotic sensitivity and production, leading to different ecotypes (ecologically distinct population), ultimately resulting in two separate species sharing a same habitat but occupying different ecological niches (Cohan, 2001). Both from their practical and fundamental aspects, studies on antibiotic resistance are of the utmost interest and promise to bring considerable new experimental data. Such results are at our reach, considering the recent technological advances cited above (*i.e.*: molecular biology, bioinformatics and increased sequence data-throughput), and appropriate experimental planning and analysis tools from theoretical ecology. Soil is the original habitat of most currently known bacteria producing antibiotic molecules and certainly one of the richest habitats in term of microbial diversity. Methods to study environmental bacteria in complex ecosystems such as soil are still limited and only recent advances in sequencing technology and molecular biology have started to fill the gap. A majority of environmental microbes resist cultivation with standard laboratory techniques and the few cultured ones do not accurately represent overall biodiversity. These uncultured microorganisms remained mostly a black-box until the

metagenomic revolution (Handelsman et al., 1998; C. Walsh, 2000; Dechesne et al., 2003; Schloss and Handelsman, 2005; Bais et al., 2006; Mummey et al., 2006; Kümmerer, 2009). The principle of metagenomics (from “meta-“ = *all* and genomics) *sensu lato* is to use standard molecular techniques on DNA extracted directly from environmental samples, such as a gram of soil, alleviating cultivation biases. Development of metagenomics jointly with next-generation sequencing technologies alleviated a number of previous critical technical limitations. Details on metagenomics methods developed to study antibiotic resistance in soil are available elsewhere (Kakirde et al., 2010; Monier et al., 2011).

### **The environmental resistome.**

The complex antibiotic resistome ecology can be deciphered only if taking into account its environmental aspect (*i.e.*: outside hospitals' walls) and considering the clinic as an environment among others with its own specific properties (D'Costa et al., 2007; Allen et al., 2010). In contrast to a strict clinical definition of antibiotic resistance, the resistome includes every resistance gene: precursors with low resistance or affinity for antibiotic molecules; cryptic resistance genes with no or low expression of resistance phenotype in their hosts, and clinical resistance genes such as extended spectrum  $\beta$ -lactamases (ESBL) conferring high antibiotic concentration resistance (Wright, 2007). Such genes, defined as antibiotic resistance when found in pathogens, may serve a very different function for their host in the environment. Taking into account the importance of host and genetic context for resistance phenotype expression is therefore crucial. This is illustrated in Figure 2, representing the environmental resistome, with overlapping circles: the same gene (*i.e.*: genetic sequence), depending on its host and ecological context may belong to several subgroups of the environmental resistome. A potent antibiotic-resistance conferring gene can be for example

silenced in a given genome or overexpressed in another one. This makes this gene a cryptic antibiotic resistance in one case, and for example a problematic antibiotic-resistance phenotype if effectively expressed in a pathogen that will resist antibiotic treatment. Most antibiotic molecules target generalist bacterial functions such as cell wall synthesis (*e.g.*:  $\beta$ -lactams, vancomycin) or protein translation (*e.g.*: tetracyclines, aminoglycosides) that are essential for most bacteria. It is likely that producer strains need to shield themselves against their own antibiotics in order to avoid suicide (Benveniste and Davies, 1973; Cundliffe, 1989; Ranjard and Richaume, 2001; Nunan et al., 2003; Cundliffe and Demain, 2010) conveniently explaining origin of antibiotic resistance genes in soil saprophyte microorganisms. Yet, the role of antibiotics as specific inhibitors involved in competition for resource at the micro-scale level is not so clear in many cases. From response to antibiotic production to core cell physiological functions, antibiotic resistance genes are playing a number of different roles in soil saprophytic bacteria yet to determine in most cases. Experimental studies trying to address these alternative hypotheses about the cause and consequences of antibiotic resistance and antibiotic production diversity in soil in a global view of antibiotic resistance will be reviewed in this chapter in an effort to better understand the different aspects of antibiotic production and resistance genes evolution in this environment.

### **On the origin of soil biodiversity and the struggle for existence.**

Understanding the specific features of a complex environment such as the soil is critical to determine how all the components of this biota, from cells to molecules, are interacting (Figure 2). Biologically, soil is an incredibly rich environment, both in terms of microbial abundance and species diversity. A gram of soil may contain from  $10^6$  to  $10^9$  bacterial cells

of  $10^3$  to  $10^6$  different bacterial species, depending on its location (*i.e.* forest or prairie soil, tropical or polar soil) (Girvan et al., 2003; Peay et al., 2007; Torsvik, 2011). These numbers may vary when considering the method used (Kirk et al., 2004; Roesch et al., 2007; Turnbaugh et al., 2009; Delmont et al., 2011; Forsberg et al., 2012) or how bacterial species or Operational Taxonomic Units (OTU) are defined (Benveniste and Davies, 1973; Demanèche et al., 2008; Goldstein and DeSalle, 2011; Forsberg et al., 2012; Zhang et al., 2013). Soil is formed with aggregates of solid organic matter (SOM) and mineral particles of different sizes maintained together by plant roots, fungi mycelium and bacteria forming sticky exopolysaccharide (EPS) biofilms (Benveniste and Davies, 1973; Burmølle et al., 2007). Spatially, cell density varies in soil making it a patchy environment (Figure 3). Most bacterial cells are living in biofilms colonies, packed in small pores formed between mineral particles or in the rhizosphere under influence of plant root exudates, where nutrients are abundant and the soil biotic capacity is significantly increased. Inversely, large pores of different sizes between mineral particles are mostly empty of bacteria (Figure 3) (Bais et al., 2006; Demanèche et al., 2008; Forsberg et al., 2012). In addition to local biome climate, a whole range of factors will modify a general model as the one presented on Figure 3. A large range of physicochemical gradients forms in soil through physical diffusion. Environmental gradients are crucial for ecological niche differentiation, which increases the range of different habitats in soil. As a consequence, soil is patchy and composed of heterogeneous aggregates where nutrients availability is limited. Ecological theory suggests a model where ecosystem species richness is dependent on the number of different niches available (J. S. Song et al., 2005; Soberon, 2005; Prosser et al., 2007; Allen et al., 2009). Soil spatial heterogeneity maintains a tremendous bacterial genetic diversity: microbial cells are struggling for existence in this habitat and try to colonize the microscale habitat best suited

to their ecotype (Cohan, 2001; J. S. Song et al., 2005; Demanèche et al., 2008; Allen et al., 2009; Philippot et al., 2010). Environmental factors (*e.g.*: particle size, pH, water availability, vegetation cover etc.) are correlated to bacterial or fungi species diversity distribution in soil (Ranjard and Richaume, 2001; Nunan et al., 2003; K. Lewis, 2013) (Peay et al., 2007; Davies, 2011). Could this constant interference competition that is the inhibition of neighboring cells by secretion of toxic compounds, between antibiotic producers and cells living nearby, explains why antibiotic producers and resistant strains are so prevalent in soil? Is this antibiotic resistance genetic diversity a problem in term of healthcare? Soil is an open and connected ecosystem interacting constantly with all compartments of the biosphere that makes exchange of genetic material possible between ecologically distinct lineages usually found in other ecosystems, such as commensal bacteria inhabiting the human gut for example where expression of antibiotic-resistance may be problematic (Turnbaugh et al., 2009; Davies and Ryan, 2012; Forsberg et al., 2012).

### **The soil parvome: living in the chemical arsenal.**

If antibiotic production is a common feature for a large diversity of soil microbes, a proportional diversity of antibiotic resistance genes can be expected in bacterial populations confronted to such pressure in soil. Pathogens' antibiotic resistance genes are sharing striking nucleotide identities with genes encoding the same antibiotic-resistance genes in the environment (Benveniste and Davies, 1973; Bérdy, 2005; Demanèche et al., 2008; Forsberg et al., 2012). Soil is likely be the main reservoir of ARGDs transferred to human pathogens or commensal bacteria. The same resistance enzymes (Benveniste and Davies, 1973; Davies, 2011) or even allelic variants of antibiotic resistance genes (Benveniste and Davies, 1973; Demanèche et al., 2008; Martinez, 2008; Monier et al., 2011; Forsberg et al.,

2012; Nesme et al., 2014) are found both in soil antibiotic producers and pathogenic bacteria. Genes encoding common clinical resistance enzymes such as  $\beta$ -lactamases exhibit more diversity in a soil sample than the same genes in the clinical environment (J. S. Song et al., 2005; Allen et al., 2009; Kallifidas et al., 2012). These results are consistent with the hypothesis of a shared resistome between soil bacteria and pathogens, however, limited in diversity. Gene allele variations shared between pathogens for a given resistance protein are far less diverse than those found in environmental samples (J. S. Song et al., 2005; Demanèche et al., 2008; Allen et al., 2009; D'Costa et al., 2011). This shared resistome has important implications for the future of antibiotics and drug development. Are all environmental resistance genes equally susceptible to jump into pathogen genomes and thwart present and future antibiotic therapies in a matter of time? Soil is the natural habitat for many microbial antibiotic producers. *Actinobacteria* and other soil microorganisms provided most of the currently-used pharmaceutical drugs, which were developed from the 1940s to 1980s (Bhullar et al., 2012; K. Lewis, 2013). High-throughput screening of actinomycetes and related microbes has led to isolation of a wide range of antibiotic molecule classes with very different properties and mechanisms of action. Davies (2011) describes these bioactive compounds as the soil "parvome". The parvome is consisting of all small molecules produced by living organisms as secondary metabolites (from "parv-" = small, "-ome" = group) (Marshall et al., 1998; D'Costa et al., 2011; Davies, 2011). These products encompass, but are not limited to, all secondary metabolites and serve various functions for their producers, from microbes to plants, insects and animals (Davies and Ryan, 2012; Liu et al., 2013). In the context of soil and antibiotic production, the soil parvome seems to be incredibly abundant and still largely unknown. Judged from what has been currently investigated among the whole tree of life, a major part of bioactive

compounds has been identified in microorganisms. Amongst 16,500 different antibiotic molecules produced by microbes: 29.7% are synthesized by fungi; 52.73% by actinobacteria and 17.58% by other bacteria phyla (Figure 1) (Bérdy, 2005; Fajardo and Martínez, 2008). Microorganisms secrete an even greater repository of bioactive compounds, such as antibiotics (C. Walsh, 2000; Davies, 2011) but molecules isolated from microorganisms during the golden age of antibiotic molecules drug discovery may likely come from a rather "easy to cultivate" part of the overall microbial antibiotic producer diversity (Figure 3). Few new natural antimicrobial drugs have been discovered recently as the antibiotic discovery pipeline is drying out and needs to be fixed. This implies an even more diverse production *in-situ*, especially when keeping in mind the low amount of cultured bacteria from environmental samples by traditional laboratory measures. Considering currently available technologies such as single cell-sequencing, individual cell-sorting, more and more metagenomes and genomes sequencing and screening throughput, a large array of antibiotics drug discovery platforms could be readily developed and enrich an already very diverse library of discovered antibiotic compounds in soil microbes. In this respect, new concepts such as global metagenomic approaches and development of high-throughput sequencing technologies can now try to circumvent this well-admitted cultivation bias for the study of antibiotic resistance in the environment (Benveniste and Davies, 1973; Martinez, 2008; Monier et al., 2011; Bhullar et al., 2012; A. R. Hall, 2013; Nesme et al., 2014). Still, fishing in such a consequent diversity remains a considerable challenge. Screening of antibiotic molecules by functional metagenomic usually relies on expression of environmental DNA, in a heterologous host such as *E. coli* or *B subtilis*, that is subsequently tested then for its potency to produce antibiotic molecules. This requires a successful cloning step, a successful expression of antibiotic production genes step and finally a

successful screening of the heterologous protein. These three steps all have a limited success rate that is overcome by dealing with very large numbers of transformed bacteria. Great numbers, higher screening throughput and of course, more time... It is, however, possible to limit this screening by targeting already known antibiotic compounds biosynthesis pathways and screening only DNA fragments likely to bear such antibiotic properties. With such a rationalized approach Sean Brady and his group (Kallifidas et al., 2012) was able to isolate DNA fragment from soil that can be used to synthesize a new antibiotic molecule: Tetracycline A. This new antibiotic is notably active against methicillin-resistant *Staphylococcus aureus* strains (MRSA), a very problematic multi-drug resistant pathogen, (Marshall et al., 1998; Laskaris et al., 2010).

#### **The soil resistome diversity: surviving the chemical arsenal.**

From a broader perspective, this large chemical cocktail suggests a reservoir of interacting genes with matching proportions. In terms of microbial ecology, soil microorganisms interact with (and eventually resist) an impressively diverse chemical arsenal, whose function is yet to be determined in most cases but may likely be very old. This emphasizes the ecological significance of antibiotic resistance. Experimental studies also show that potent antibiotic resistance genes can be detected in very old environmental samples. Exploring a 30,000-year-old Alaskan permafrost soil sample, D'Costa *et al.* (D'Costa et al., 2011) extracted DNA molecules with a strong nucleotide homology to antibiotic resistance genes isolated from clinical settings. These genes match common resistance genes, such as *Streptomyces* spp.  $\beta$ -lactamases *bla*, several *tetM*-related tetracycline ribosomal protection proteins, and more surprisingly, a complete *vanHAX* vancomycin resistance operon that is similar to current clinical resistance genes. Expression of the "old" *vanA* gene variants from

the permafrost cores in a heterologous *Escherichia coli* host produced a vancomycin resistant phenotype. This is consistent with those ancestral genes being functional (Haas and Défago, 2005; A. R. Hall, 2013). With a similar kind of approach, Bhullar *et al.* (2012) isolated antibiotic resistant bacteria from the Lechuguilla Cave (Carlsbad Caverns National Park, USA) that have been kept apart from the outside world for 4 to 7 million years (*i.e.* before *Homo sapiens* appeared about 2.3 million years ago). A plethora of bacterial strains from this human-less environment resisted as many as 14 different antibiotic molecules (Bhullar *et al.* 2012). Some resistance mechanisms identified in these strains are similar to those for known antibiotic resistance genes, but many of the genes (and their mechanisms of resistance) identified in this study were unknown beforehand (Mummey *et al.*, 2006; Laskaris *et al.*, 2010; Bhullar *et al.*, 2012).

The diversity of antibiotic resistance genes found in very different bacterial lineages could then be explained as countermeasures to inhibitory concentration of antibiotics secreted by producers. This *arms-shield* hypothesis could explain the origin of clinical antibiotic resistance genes as corroborated by experimental studies (Marshall *et al.*, 1998; Delmont *et al.*, 2011; D'Costa *et al.*, 2011), but other hypotheses could also explain the high diversity and abundance of resistance genes in natural environments, such as inhibitors of quorum-sensing (Dechesne *et al.*, 2003; Seoane *et al.*, 2011; Liu *et al.*, 2013; Musovic *et al.*, 2014), translation factors (Thiele Bruhn, 2003; Haas and Défago, 2005; Fajardo and Martínez, 2008) or a common origin with essential bacterial components (C. Walsh, 2000; Dechesne *et al.*, 2003; Mummey *et al.*, 2006; Kümmerer, 2009), explaining their high prevalence in environmental bacteria genomes. If antibiotic resistance genes are indeed solely due to antibiotic molecule inhibition, this is supposing that antibiotic molecules are active and often found at inhibitory concentrations in soil ecosystems. Thus, resistance genes would

significantly increase fitness in order to be effectively selected (Thiele Bruhn, 2003; Bhullar et al., 2012; A. R. Hall, 2013). This arms-shield race will therefore leave footprints in genomes involved. This was for example shown in populations of *Streptomyces* spp. producers-resistant and resistant only genomes that have undergone co-evolution attested by frequent genomic recombination at antibiotic resistance or production genes loci (Marshall et al., 1998; Thiele Bruhn, 2003; Moreno-Bondi et al., 2009; Laskaris et al., 2010). While being non-exclusive, the role of antibiotics as inhibitors of microbial growth in soil has been experimentally studied. Soil microbiologists have long since tried to use bio-control agents of phytopathogens after observing competitive interactions occurring in the rhizosphere, where competition for resources is intense. Some rhizobacteria, such as *Pseudomonas fluorescens* or *Bacillus subtilis*, produce antibiotic molecules targeting soil-born plant-pathogens also colonizing the rhizosphere. Therefore, antibiotic molecules and their subsequent resistance genes are likely to be found in rhizospheric soil. Observations of antagonistic interactions that often involve plant-pathogens are experimented and reviewed (Tolls, 2001; Haas and Défago, 2005; Kümmerer, 2009; A. R. Hall, 2013; Leal et al., 2013) and demonstrate that antibiotic molecules can be potent in soil. This has been extensively studied in the case of phytopathogens or plant-growth promoting rhizobacteria interactions occurring in soil under the influence of plant's roots exudates leading to a fierce competition for resources. New concepts and methods to carefully model the 3D structure of bacterial population at the micro-aggregate scale are currently under development and will hopefully enhance our understanding of the micrometer-scale interactions occurring in soil (Thiele Bruhn, 2003; Mummey et al., 2006; Moreno-Bondi et al., 2009; Laskaris et al., 2010; Girardi et al., 2011; Braschi et al., 2013). Considering that soil DNA extraction is a major source of variations (Tolls, 2001; Thiele Bruhn, 2003; Delmont et al., 2011; Leal et al.,

2013), new mathematical models are designed for a better and more even sampling between sites, making inter-sample comparison with minimal errors (Thiele Bruhn, 2003; Dechesne et al., 2003; Chee-Sanford et al., 2009; Seoane et al., 2011; Girardi et al., 2011; Braschi et al., 2013; Musovic et al., 2014). These authors propose a more rigorous sampling of soil micro-scale functional diversity and should be considered in the frame of antibiotic resistance studies in soil, together with a better understanding of different antibiotic molecule activities on different bacterial populations. How can samples be representative of the soil structural composition when micro-aggregates range from  $10^{-6}$  to  $10^{-3}$  meters and sampling is done at the  $10^{-2}$  meter, for DNA extraction whatever amount, or when results are used to infer biology at the kilometer scale such as in bio-geographic studies?

#### **Dosing small molecules in soil at the soil micro-aggregate scale.**

Another very important point to consider is the concentration at which antibiotic molecules and microbes interact. Measuring complex and small molecules such as antibiotics in very heterogeneous matrix like soil is subject to technical limitations and we are currently unable to accurately measure trace amounts of antibiotics. A relatively recent compilation of studies measuring *in-situ* concentration and degradation kinetic of pharmaceuticals compounds, including antibiotics, in soil or water environments are extensively reviewed elsewhere (Thiele Bruhn, 2003; Haas and Défago, 2005; Willingham and Werth, 2008).

These papers report that the highest concentrations are found either in production animal feces (pig and chicken farms) or in antibiotic molecules production sites or hospital sewers outlets. In comparison, “natural” soil (*i.e.*: not obviously exposed to antibiotic pollution) is not accumulating antibiotics and concentrations measured are in the  $\mu\text{g.kg}^{-1}$  range,  $10^2$  to  $10^3$  fold less concentrated than the  $\mu\text{g.g}^{-1}$  range of concentrations used for medical

antibiotic therapy, despite high producer abundance found in soil (Figure 1). The most widely used analytical techniques for antibiotic molecules are liquid chromatography separation coupled with mass spectrometry (LC-MS) and tandem mass-spectrometry (LC-MS/MS), which are highly sensitive for a large range of compounds in the  $\mu\text{g}\cdot\text{kg}^{-1}$  range (Thiele Bruhn, 2003). Antibiotics are largely variable in chemical structure and extraction yield from the soil matrix can vary considerably between two compounds. Therefore, sample preparation prior to analysis is critical and results are highly influenced by this step (Thiele Bruhn, 2003; Moreno-Bondi et al., 2009; Sun et al., 2013). Besides solvent choice, antibiotics are showing a broad range of adsorption behaviors on small mineral particles, such as negatively charged clay (la Torre et al., 2011). Antibiotics can be either tightly immobilized on soil particles or washed away by water runoff depending on their chemical properties (Tolls, 2001; Fajardo and Martínez, 2008; Kümmerer, 2009; Tsuey-Ching et al., 2012; Sun et al., 2013; Leal et al., 2013). In addition to soil sorption, antibiotics may be degraded by biotic or abiotic mechanisms at varying rates (Thiele Bruhn, 2003; Bahl et al., 2004; Moreno-Bondi et al., 2009; Aminov, 2009; Girardi et al., 2011; Braschi et al., 2013). All these factors explain why antibiotics are hardly detected in soil and do not seem to accumulate significantly in a bioactive form, except in specific **highly polluted environments where the highest concentrations are measured, such as in fresh manure or antibiotic producing plants sewage waters, densely populated waste-water treatment plants for example** (M. Roberts, 1996; Tolls, 2001; Thiele Bruhn, 2003; M. C. Roberts, 2005; Leal et al., 2013). Moreover, measuring chemicals may be counter-intuitive in some cases as a detectable compound may not be accessible for bacteria in a bioactive form. Modeling the *in situ* effects of chemical perturbation on bacterial communities with *in vitro* devices, such as microfluidic flow-cells modeling the soil porous environment (Massova and Mobashery,

1998; Garau et al., 2005; Willingham and Werth, 2008) or exposure to chemical gradients (Chuanchuen et al., 2003; Meroueh et al., 2003; J. L. Song et al., 2014; Li et al., 2014; Wang et al., 2014) in an attempt to decompose each factors' effects on either pure or mixed microbial cultures. Such devices are still in active development and therefore have not yet been used in thorough environmental microbiology studies. Another way around is to make use of a proxy to measure antibiotic-concentration. With the help of a reporter strain, Bahl *et al.* (2004) were able to measure that 0.4% of the total tetracycline ingested by rats is bioavailable to their gut microbiota (Bahl et al., 2004; Martinez and Sánchez, 2009; Chee-Sanford et al., 2009; Reith and Mayer, 2011; Johnson et al., 2013). These studies are still scarce and few have been successfully applied to the soil environment (Meroueh et al., 2003; Thiele Bruhn, 2003; J. S. Song et al., 2005; Sun et al., 2013). Such approaches have however the clear advantage of measuring bioavailability instead of total antibiotic concentrations, a determinant factor in selection of antibiotic-resistance genes.

Quantification of antibiotic concentrations at the soil micro-aggregate scale, where antibiotic antagonists interaction are the much likely to occur, would be necessary in order to understand at which concentrations most bacteria are exposed. The dose-response curve observed in antibiotic therapy, where aim is limitation of microbial growth, may not be the same at low concentration at which bacteria sense antibiotics in situ. The same antibiotic molecule has been shown to act on the same strain both as an activator and an inhibitor of transcription depending on its relative concentration (Fajardo and Martínez, 2008; Hibbing et al., 2009; Tsuey-Ching et al., 2012; Sun et al., 2013).

### **An ecosystem vision of antibiotic resistance and production in soil.**

Describing soil microorganisms' complex ecology and physiology *in-situ* will hopefully help us to preserve current and future antibiotic efficiency. Antibiotic resistance is as an ecosystem feature prevalent in the environment and especially in soil based on our current knowledge. Defining a common core of functions conferred by antibiotic resistance genes or the role of antibiotic molecules in nature seems irrelevant when looking at the incredible diversity of mechanisms already described. What we are commonly labeling as antibiotics, however, share common general features such as being small molecules and having interactions with essential bacterial components. Could antibiotic molecules and their corollary antibiotic resistances genes (or rather interacting genes) serve different purposes than interference competition in the environment, such as metabolism regulation, quorum sensing/quenching or even inter-domain communication (Bahl et al., 2004; Davies, 2006; Allen et al., 2009; Aminov, 2009; Martinez and Sánchez, 2009)? The diversity of molecular mechanisms displayed by microorganisms is tremendous and we only know a fraction of it. Several alternative evolutionary scenarios relying on other selection effects beside competition have been suggested as origins for at least part of the antibiotic resistance gene diversity.

A well-studied example of secondary adaptation leading to antibiotic resistance is multi-drug resistance (MDR) efflux pumps. MDR efflux pumps form a membrane-anchored protein family whose role is to regulate cell homeostasis by selective diffusion of compounds in or out the cytoplasm. Some pumps have a very specific affinity for antibiotics and can confer a high level of resistance to their hosts, such as tetracycline resistances conferred by specific efflux pumps (M. Roberts, 1996; M. C. Roberts, 2005; Fenical and Jensen, 2006; Reith and Mayer, 2011; Johnson et al., 2013). Drug-specific and multi-drug resistance efflux pumps are very similar in structure to other efflux pumps and may have, therefore, been selected for

cations such as  $\text{Na}^+$  or  $\text{K}^+$ , used to exchange protons, regulate cell homeostasis, excrete toxic metabolic compounds and even trigger cell communication (Shen, 2003; J. S. Song et al., 2005; Aminov, 2009; Martinez and Sánchez, 2009). Bacterial metabolism and physiology are quite plastic and a single gene sequence may give rise to a broad spectrum of different functions depending on the cell context. A gene that was previously selected for homeostasis regulation or heavy metal excretion may become well suited in the clinical context and be selected rather for antibiotic resistance than its ancestral bacteria physiological feature. Another example of secondary-adaptation following gene-duplication is the penicillin binding protein and  $\beta$ -lactamases protein families that display respectively a transpeptidase and a  $\beta$ -lactam ring hydrolyzing activity.  $\beta$ -lactams (*i.e.*: penicillin and its derivatives) are the most widely used antibiotic drugs in human medicine.  $\beta$ -lactams inhibit bacterial growth by sequestering transpeptidase enzymes known as penicillin-binding proteins (PBP). Penicillin-binding-proteins (PBP) are responsible for amino-acid crosslinking of peptidoglycan layers and cell-wall formation. When inhibited by  $\beta$ -lactams antibiotics, inactivated PBP will cause improper cell division. Interestingly, PBP (the antibiotic target) and  $\beta$ -lactamases (the resistance genes) share sequence similarity at the amino-acid level, similar protein fold, same structural motifs and similar active sites when considering structural biology results (Massova and Mobashery, 1998; Garau et al., 2005). Only slight divergences between the two protein families led to considerable diversification of function. Still, PBP and  $\beta$ -lactamases are related from an evolutionary perspective and ancient penicillin-binding proteins are thought to have given rise to those of  $\beta$ -lactamases (Chuanchien et al., 2003; Meroueh et al., 2003; Fenical and Jensen, 2006; Boucher et al., 2009). After a PBP gene duplication releases selection pressure on an essential locus such as PBP encoding genes,  $\beta$ -lactamases may have been selected for cell-wall related functions,

such as peptidoglycan recycling, and subsequently for resistance to  $\beta$ -lactam inhibition (Meroueh et al., 2003; Reith and Mayer, 2011). Furthermore, the  $\beta$ -lactamase genetic diversity encountered in the clinic is also found, together with many other allelic variants, in Edison seamount deep-sea sediments. These observations clearly favor a transfer hypothesis of antibiotic resistance genes from soil to commensal and pathogenic bacteria. Antibiotic resistance would then in many cases be more a by-product of bacterial genetic evolution (secondary adaptation) of general bacterial functions such as cytoplasm homeostasis, cell partition or cell-wall assembly (Meroueh et al., 2003; Martinez, Sánchez, et al., 2009; Reith and Mayer, 2011; Liu et al., 2013). Antibiotic resistance is an advantageous trait that may often be a secondary adaptation to very specific environmental conditions where antibiotic concentrations produce stressful conditions, such as clinical environment in a patient's gut undergoing antibiotic therapy (Looft et al., 2012; Forslund et al., 2013), or in the rhizosphere, where competition for resources is fierce and where interference competition can be strongly selected (Haas and Défago, 2005; Oskay, 2009). Antibiotic resistance and production genes are found in environment where degradation and dilution of secreted compounds are likely to occur at a rapid rate, such as in high water content environments. In this context, secreting antibiotics for local competition seems to be a rather costly behavior. Yet, marine specific clades of *Actinobacteria* can be recovered from ocean samples and are potent antibiotic producers (Fenical and Jensen, 2006; Reith and Mayer, 2011; Toleman et al., 2012; Heuer and Smalla, 2012; Bogaerts et al., 2013; Johnson et al., 2013), which indicates the ubiquity of antibiotic production and resistance genes in all environments. Biosynthesis of many small bioactive metabolites, including many antibiotics, is performed by polyketide synthases. Polyketide synthases (PKS) are specific multifunctional enzymes similar to fatty acid biosynthesis pathway enzymes. A large number

of PKS modules encode modular enzymes responsible for the synthesis of commonly used antibiotics such as macrolides or peptide antibiotics, that are commonly used in medicine (Stein, 2005; Jenke-Kodama and Dittmann, 2009; Pidot et al., 2014). This complex genetic system of multi-modular enzymes is conserved in bacteria, fungi, plants and some animal lineages (Ginolhac et al., 2004). Genetic modules encoding PKS are organized in clusters where homologous recombination is likely to occur and generate an impressive diversity of possible enzymes, and therefore, bioactive compounds synthesis pathways (Sidhu et al., 2002; Shen, 2003; J. S. Song et al., 2005; Yong et al., 2009; Carattoli et al., 2012). PKS genetic diversity associated with their multifunctional properties enables soil microorganisms to produce an incredible diversity of different antibiotics molecules (Figure 1) (Weldhagen, 2004; Bérdy, 2005; Heuer and Smalla, 2012). This fantastic resource of genetically encoded bioactive compounds secreted by soil bacteria is a fantastic source of novel antibiotic drugs. A considerable challenge to now overcome is to prevent dissemination of resistance genes, already existing in the environment, before these new drugs become themselves ineffective to cure infections by pathogens.

### **Dissemination of antibiotic resistance: from horizontal gene transfer to superbugs.**

A major challenge for antibiotic resistance research studies is to address the question of catastrophic antibiotic resistance increase in clinic leading to a significant increase in infectious diseases mortality, that were before treatable. Yet, antibiotic resistance is a common ecological feature in soil bacteria, as is antibiotic molecules production.

Experimental observations and evolutionary studies indicate that antibiotic resistance genes are ancient and ubiquitous and represent a reservoir of transferable genetic material. From an epidemiology perspective, antibiotic resistance is more and more prevalent in the clinic.

But very problematic clinic-related antibiotic resistance concerns only a limited number of different bacterial phyla. Such as the ESKAPE pathogens: *Enterococcus faecium*, *Staphylococcus aureus* belong to Firmicutes; *Enterobacteriaceae* spp. (*i.e.* *Klebsiella*, *Escherichia* species), *Acinetobacter baumannii* and *Pseudomonas aeruginosa* belong to gamma-Proteobacteria (*i.e.* Gram negative bacteria) (Fenical and Jensen, 2006; Boucher et al., 2009; Jolivet-Gougeon et al., 2011). Resistance to antibiotic treatment in the clinic is much higher than the rate of discovery of new antibiotic resistance genes or variants in pathogen genomes, coherent with a dissemination of selected antibiotic resistance genes (clonal expansion after selection) rather than an increase in new antibiotic-resistance gene recruitment from environmental reservoirs. ARGDs themselves do not seem, however, to be strongly selected in the environment but on the other hand can disseminate at an incredibly fast pace in clinical settings, such as the New-Dehli metallo- $\beta$ -lactamase bla<sub>NDM-1</sub>. This gene was first identified in a New-Delhi hospital in India in 2009 and spread worldwide in very short amount of time (Boucher et al., 2009; Yong et al., 2009; T. R. Walsh et al., 2011; Toleman et al., 2012; Dortet et al., 2012; Bogaerts et al., 2013). Shortly after this first identification, it was detected worldwide in hospitals, indicating its prevalence in *Enterobacteriaceae* pathogenic strains (Ho et al., 2011; P et al., 2011; Forsberg et al., 2012; Ho et al., 2012). The bla<sub>NDM-1</sub> gene has also been detected in *Acinetobacter baumannii* isolates, which is a soil dwelling bacteria responsible for opportunistic nosocomial infections. Comparative genomics analysis of bla<sub>NDM-1</sub> locus in different genomes suggests that this gene has emerged by chromosomal recombination in *Acinetobacter* strains (Toleman et al., 2012; Bogaerts et al., 2013). This gene is often found on plasmids belonging to different incompatibility groups such as IncN and IncA/C in *Enterobacteriaceae* strains (Villa et al., 2012; Carattoli et al., 2012). Such plasmids have a broad-host range and can be

replicated in different bacterial lineages. IncN and IncA/C are also conjugative and possess all the genes required for their autonomous transfer (Carattoli, 2009). Genes carried by these genetic elements can be disseminated at a rapid rate in distant bacterial lineages. Bacteria possess myriads of molecular mechanisms developed for rapid adaptation to their changing environment from accessory genetic elements such as plasmids (Smillie et al., 2010) or transposons (Liebert et al., 1999; A. P. Roberts and Mullany, 2009) to hypermutator phenotypes (Turrientes et al., 2010). Each of these different mechanisms have implications at different population levels and have all been shown to mediate antibiotic resistance gene transfer by mobilizing resistance genes carried on their accessory module, but are also induced by exposure to antibiotic molecules (Jolivet-Gougeon et al., 2011; Blázquez et al., 2012; Baharoglu et al., 2012). All these observations are emphasizing on the duality (first and second orders) of natural selection processes involved in antibiotic resistance gene dissemination in pathogens taking place in a more complex and general ecology of microbial genome dynamics. Gene *bla<sub>NDM-1</sub>* is now spread worldwide due to a succession of distinct HGT events: emergence of the *bla<sub>NDM-1</sub>* gene by genomic recombination within *Acinetobacter spp.* population; transfer on a mobile genetic integrative element, such as a transposon; transposon insertion on a broad-host range plasmid; and dissemination in different bacterial lineages including pathogenic *Enterobacteriaceae*. Experimental studies on the *bla<sub>NDM-1</sub>* gene determinant of antibiotic resistance tracked its dissemination from an environmental reservoir (*Acinetobacter spp.* to large dissemination on plasmids. Such an evolutionary scenario illustrates the distinction between efficient resistance genes (*i.e.*: conferring resistance above a therapeutic antibiotic molecule concentration) and efficient dissemination of antibiotic resistance genes that can both lead to an increased incidence of multi-drug resistant pathogens and involve different selection processes. A large range of

other studies have so far enabled to trace the evolutionary and hosts route of several types of antibiotic-resistance genes such as: plasmid-mediated *qnrA* (quinolones resistance) linked to a possible origin in environmental *Vibrio* spp. strains (Poirel and Rodriguez-Martinez, 2005; Nordmann, 2005); *bla<sub>OXA</sub>* (Barlow and B. G. Hall, 2002; Poirel et al., 2004); *bla<sub>CTX-M</sub>* (Cantón et al., 2012; Villa et al., 2012) and *bla<sub>AmpC</sub>* (Humeniuk et al., 2002) types of beta-lactamases or tetracyclines ribosomal-protection resistance genes (Aminov et al., 2001). These diverse experimental results are coherent with previously established models on antibiotic resistance gene dissemination following antibiotic treatment introduction to cure infection by pathogenic bacteria but also illustrate the complexity and diversity of mechanisms and bacterial populations involved in antibiotic transfer from the environmental resistome. Mobile genetic elements are found in all microbial genomes and antibiotic-resistance genes are often associated with such genetic structures in pathogen genomes (Stokes and Gillings, 2011; Nardelli et al., 2012; Djordjevic et al., 2013). These molecular mechanisms participate in the generation of genetic diversity and horizontal gene transfer between and within populations (Ochman et al., 2000; Halary et al., 2010; Gophna and Ofan, 2011). Such events cannot be prevented, but their rate is quite low at the population level because any mutant emerging individually within a population is likely to be outcompeted by neighboring cells or lost due to genetic drift before this newly acquired trait is fixed in the population. Most of antibiotic resistance increase in the clinic is more likely to be the result of successful mutant clonal expansion rather than massive exchanges between soil and clinical resistomes, as observed with experimental results investigating sequence similarities shared between soil antibiotic resistance genes and those found in the human-gut or pathogens (Benveniste and Davies, 1973; Cheng et al., 2012). In a recent study, the same authors show that only a limited number of resistance genes are shared

between human-gut and soil bacteria. Forsberg et al. (Forsberg et al., 2012) illustrated this by showing that only few sequences found in both environments and conferring antibiotic resistance phenotype possess nucleotide similarity above 90%. They are also showing that there is a coherence between differences in taxonomic profile and differences in antibiotic resistance gene distribution found in soil or human feces samples (Forsberg et al., 2014).

Authors are suggesting that this is due to a limited rate of successful transfer events between these two environments. These need to be elucidated with a better understanding of lateral gene transfer events occurring in soil in order to decipher the relationship between transfer of antibiotic resistance and gene transferability via mobile genetic elements (Thomas and Nielsen, 2005; Dantas and Sommer, 2012; Sommer, 2014). Indeed, it was found as early as 1983 that conjugative plasmids, often responsible of resistance gene transfer in Enterobacteriaceae pathogens, are common in these strains before the antibiotic era showing that transfer mechanisms are abundant in bacteria and selected independently of antibiotic resistance genes (V. M. Hughes and Datta, 1983; Perry and Wright, 2013).

### **Concluding remarks.**

This is the increasing problem of antibiotic resistance in clinical environments that lead microbiologists to question the origin of antibiotic resistance and hypothesized the pre-existence of an environmental resistome. This assumption is supported by the very rapid acquisition of resistance mechanisms by pathogens, previously sensitive, upon introduction of new antibiotics. At first, only exposure of pathogens to antibiotics was considered as a selective factor eliciting antibiotic-resistance genes and increasing their dissemination rate. This is only since the 1970's that the environmental reservoir of antibiotic-resistance genes is considered as source of antibiotic-resistance for pathogens and plays a major role in their

dissemination. There are still substantial limitations for the dissemination of antibiotic resistance phenotypes to pathogen populations. This is revealed by a striking discrepancy between diversity of antibiotic resistance mechanisms that can be found in the environment when compared to those found on clinical isolates. Soil was the first environment to be considered, due to origins of antibiotics molecules, isolated from soil microorganisms and when it was found that producers themselves encode antibiotic-resistance genes identical to those found in the clinic. Major results concerning the antibiotic-resistance problem were obtained in the field of environmental microbiology with the notable help of molecular biology applications directly on DNA extracted from soil samples. These methodological and conceptual revolutions unraveled the incredible diversity of soil microorganisms diversity, orders of magnitude higher than previously expected based on results obtained by cultivation or microscopic magnification of soil samples. While very interesting from a fundamental aspect, such discoveries should lead to the re-examination of theories on the origins of antibiotic-resistance, notably considering the impressive abundance and genetic diversity of antibiotic-resistance genes found in the environment. Sequence-based analyses show that a whole continuum of antibiotic-resistance genes diversity, coherent with species phylogeny, is present in the environment. Antibiotic-resistance confers resistance, obviously, but antibiotic molecule targets are ubiquitous in bacterial physiology and the molecular biology dogma: replication, transcription, translation or cell partitioning; and interactions with antibiotics-resistance genes seems much more complex than previously thought in a strict arms/shield selection race. Bacteria evolved a large range of mechanisms to exchange genes, remodel their genome and ultimately adapt to changing environmental conditions. These molecular mechanisms, such as integrons, super-integrons, transposons, plasmids and general transduction via virus particles (*i.e.*: mobile genetic elements); but also

insertions sequences, recombination or gene duplications and mutations are all generating genetic diversity at different rates. Studies on the mobile genetic pool exchanged by bacteria will hopefully help to tackle the problem of antibiotic-resistance dissemination in the clinic. But these studies are also considerably contributing to our understanding of fundamental microbiology and ecology. Current topics and issues, such as bioremediation of soils, discovery of new enzymatic activity with industrial applications, and hopefully new pharmaceuticals compounds resources, will largely benefit from results obtained by studies on soil bacterial diversity and adaptive potential. Considering the limited resolution of current sequencing techniques available to study the incredible microorganisms' genetic diversity found in environmental samples one of way could be to fractionate this diversity. Such targets in frame of antibiotic resistance dissemination studies could be DNA molecules from mobile genetic elements, such as viruses or plasmids likely responsible of rapid gene dissemination within distant microorganisms' lineages. A method to fractionate such diversity could be for example some of their specific properties: viruses are small particles that can be filtered from bigger cells. Plasmid DNA can also be separated specifically from chromosomal DNA, based on their respective molecular configuration. Chromosomal DNA molecules are very long (>1Mbp for free-living bacteria) when compared even to plasmid molecules (from 1kbp to 0.5Mbp). Large DNA molecules are sheared due to mechanical stress and found as linear DNA when bacterial cells are lysed. Plasmids, with proper lysis and handling procedures, can conserve their supercoiled molecular conformation and will have a higher buoyant density than linear DNA. Linear chromosomal and supercoiled plasmid DNA can therefore be separated, using for example centrifugation gradients. Such fractions of DNA molecules, named either "viromes" or "plasmidomes", represent a smaller fraction of overall genetic diversity found in environmental samples and can be hopefully assembled in

full molecules with more success than a full environmental metagenome. Such assemblies will nonetheless improve gene annotation accuracy but will at last bridge dissemination vectors phylogeny (using such markers found in viruses, or plasmid backbones) and genes they vehicle. This is only with a better insight on antibiotic resistance gene flow dynamic that we will be able to make substantial improvement for the sustainability of antibiotic molecules approved for healthcare.

### **Acknowledgements.**

The French Ministry of Higher Education and Research (MESR) funds JN. The authors declare there is no conflict of interest.

### **References.**

- Abby, S.S., Tannier, E., Gouy, M., and Daubin, V. (2012) Lateral gene transfer as a support for the tree of life. *Proc Natl Acad Sci USA* **109**: 4962–4967.
- Allen, H.K., Donato, J., Wang, H.H., Cloud-Hansen, K.A., Davies, J., and Handelsman, J. (2010) Call of the wild: antibiotic resistance genes in natural environments. *Nat Rev Microbiol* **8**: 251–259.
- Allen, H.K., Moe, L.A., Rodbumrer, J., Gaarder, A., and Handelsman, J. (2009) Functional metagenomics reveals diverse beta-lactamases in a remote Alaskan soil. *ISME J* **3**: 243–251.
- Aminov, R.I. (2009) The role of antibiotics and antibiotic resistance in nature. *Environ Microbiol* **11**: 2970–2988.
- Aminov, R.I., Garrigues-Jeanjean, N., and Mackie, R.I. (2001) Molecular ecology of tetracycline resistance: development and validation of primers for detection of tetracycline resistance genes encoding ribosomal protection proteins. *Appl Environ Microbiol* **67**: 22–32.
- Andersson, D.I. and Hughes, D. (2010) Antibiotic resistance and its cost: is it possible to reverse resistance? *Nat Rev Microbiol* **8**: 260–271.
- Baharoglu, Z., Krin, E., and Mazel, D. (2012) Connecting environment and genome plasticity in the characterization of transformation-induced SOS regulation and carbon catabolite control of the *Vibrio cholerae* integron integrase. *J Bacteriol* **194**: 1659–1667.
- Bahl, M.I., Hansen, L.H., Licht, T.R., and Sørensen, S.J. (2004) In Vivo Detection and Quantification of Tetracycline by Use of a Whole-Cell Biosensor in the Rat Intestine. *Antimicrob Agents Chemother* **48**: 1112–1117.
- Bais, H.P., Weir, T.L., Perry, L.G., Gilroy, S., and Vivanco, J.M. (2006) The role of root exudates in rhizosphere interactions with plants and other organisms. *Annu Rev Plant Biol* **57**: 233–266.
- Barlow, M. and Hall, B.G. (2002) Phylogenetic Analysis Shows That the OXA b-Lactamase

- Genes Have Been on Plasmids for Millions of Years. *J. Mol. Evol.* **55**: 314–321.
- Benveniste, R. and Davies, J. (1973) Aminoglycoside antibiotic-inactivating enzymes in actinomycetes similar to those present in clinical isolates of antibiotic-resistant bacteria. *Proc Natl Acad Sci USA* **70**: 2276–2280.
- Bérdy, J. (2005) Bioactive microbial metabolites. *J. Antibiot.* **58**: 1–26.
- Bhullar, K., Waglechner, N., Pawlowski, A., Koteva, K., Banks, E.D., Johnston, M.D., et al. (2012) Antibiotic resistance is prevalent in an isolated cave microbiome. *PLoS ONE* **7**: e34953.
- Blázquez, J., Couce, A., Rodríguez-Beltrán, J., and Rodríguez-Rojas, A. (2012) Antimicrobials as promoters of genetic variation. *Curr Opin Microbiol* **15**: 561–569.
- Bogaerts, P., Huang, T.-D., Rezende de Castro, R., Bouchahrouf, W., and Glupczynski, Y. (2013) Could *Acinetobacter pittii* act as an NDM-1 reservoir for Enterobacteriaceae? *J Antimicrob Chemother* **68**: 2414–2415.
- Boucher, H.W., Talbot, G.H., Bradley, J.S., Edwards, J.E., Gilbert, D., Rice, L.B., et al. (2009) Bad Bugs, No Drugs: No ESKAPE! An Update from the Infectious Diseases Society of America. *Clin Inf Dis.*
- Braschi, I., Blasioli, S., Fellet, C., Lorenzini, R., Garelli, A., Pori, M., and Giacomini, D. (2013) Persistence and degradation of new  $\beta$ -lactam antibiotics in the soil and water environment. *Chemosphere* **93**: 152–159.
- Burmølle, M., Hansen, L.H., and Sørensen, S.J. (2007) Establishment and Early Succession of a Multispecies Biofilm Composed of Soil Bacteria. *Microb Ecol* **54**: 352–362.
- Cantón, R., González-Alba, J.M., and Galán, J.C. (2012) CTX-M Enzymes: Origin and Diffusion. *Front Microbiol* **3**: 110.
- Carattoli, A. (2009) Resistance Plasmid Families in Enterobacteriaceae.
- Carattoli, A., Villa, L., Poirel, L., Bonnin, R.A., and Nordmann, P. (2012) Evolution of IncA/C bla(CMY-2)-Carrying Plasmids by Acquisition of the bla(NDM-1) Carbapenemase Gene. *Antimicrob Agents Chemother* **56**: 783–786.
- Chee-Sanford, J.C., Mackie, R.I., Koike, S., Krapac, I.G., Lin, Y.-F., Yannarell, A.C., et al. (2009) Fate and transport of antibiotic residues and antibiotic resistance genes following land application of manure waste. *J Environ Qual* **38**: 1086–1108.
- Cheng, G., Hu, Y., Yin, Y., Yang, X., Xiang, C., Wang, B., et al. (2012) Functional screening of antibiotic resistance genes from human gut microbiota reveals a novel gene fusion. *FEMS Microbiol Lett.*
- Chuanchien, R., Karkhoff-Schweizer, R.R., and Schweizer, H.P. (2003) High-level triclosan resistance in *Pseudomonas aeruginosa* is solely a result of efflux. *American Journal of Infection Control* **31**: 124–127.
- Cohan, F.M. (2001) Bacterial species and speciation. *Systematic biology.*
- Cordero, O.X., Wildschutte, H., Kirkup, B., Proehl, S., Ngo, L., Hussain, F., et al. (2012) Ecological populations of bacteria act as socially cohesive units of antibiotic production and resistance. *Science* **337**: 1228–1231.
- Cox, G. and Wright, G.D. (2013) Intrinsic antibiotic resistance: Mechanisms, origins, challenges and solutions. *Int J Med Microbiol* **303**: 287–292.
- Cundliffe, E. (1989) How antibiotic-producing organisms avoid suicide. *Annu Rev Microbiol* **43**: 207–233.
- Cundliffe, E. and Demain, A.L. (2010) Avoidance of suicide in antibiotic-producing microbes. *J Ind Microbiol Biotechnol* **37**: 643–672.
- D'Costa, V.M., Griffiths, E., and Wright, G.D. (2007) Expanding the soil antibiotic resistome:

- exploring environmental diversity. *Curr Opin Microbiol* **10**: 481–489.
- D'Costa, V.M., King, C.E., Kalan, L., Morar, M., Sung, W.W.L., Schwarz, C., et al. (2011) Antibiotic resistance is ancient. *Nature* **477**: 457–461.
- Dantas, G. and Sommer, M.O.A. (2012) Context matters - the complex interplay between resistome genotypes and resistance phenotypes. *Curr Opin Microbiol* **15**: 577–582.
- Davies, J. (2006) Are antibiotics naturally antibiotics? *J Ind Microbiol Biotechnol* **33**: 496–499.
- Davies, J. (2011) How to discover new antibiotics: harvesting the parvome. *Curr Opin Chem Biol* **15**: 5–10.
- Davies, J. and Ryan, K.S. (2012) Introducing the parvome: bioactive compounds in the microbial world. *ACS Chem. Biol.* **7**: 252–259.
- De Gelder, L., Williams, J.J., Ponciano, J.M., Sota, M., and Top, E.M. (2008) Adaptive plasmid evolution results in host-range expansion of a broad-host-range plasmid. *Genetics* **178**: 2179–2190.
- Dechesne, A., Pallud, C., Debouzie, D., Flandrois, J.P., Vogel, T.M., Gaudet, J.P., and Grundmann, G.L. (2003) A novel method for characterizing the microscale 3D spatial distribution of bacteria in soil. *Soil Biology and Biochemistry* **35**: 1537–1546.
- Delmont, T.O., Robe, P., Cécillon, S., Clark, I.M., Constancias, F., Simonet, P., et al. (2011) Accessing the soil metagenome for studies of microbial diversity. *Appl Environ Microbiol* **77**: 1315–1324.
- Demanèche, S., Sanguin, H., Poté, J., Navarro, E., Bernillon, D., Mavingui, P., et al. (2008) Antibiotic-resistant soil bacteria in transgenic plant fields. *Proc Natl Acad Sci USA* **105**: 3957–3962.
- Djordjevic, S.P., Stokes, H.W., and Roy Chowdhury, P. (2013) Mobile elements, zoonotic pathogens and commensal bacteria: conduits for the delivery of resistance genes into humans, production animals and soil microbiota. *Front Microbiol* **4**: 86.
- Dortet, L., Nordmann, P., and Poirel, L. (2012) Association of the emerging carbapenemase NDM-1 with a bleomycin resistance protein in Enterobacteriaceae and Acinetobacter baumannii. *Antimicrob Agents Chemother* **56**: 1693–1697.
- Fajardo, A. and Martínez, J.L. (2008) Antibiotics as signals that trigger specific bacterial responses. *Curr Opin Microbiol* **11**: 161–167.
- Fenical, W. and Jensen, P.R. (2006) Developing a new resource for drug discovery: marine actinomycete bacteria. *Nat. Chem. Biol.* **2**: 666–673.
- Fleming, A. (1929) On the Antibacterial Action of Cultures of a Penicillium, with Special Reference to Their Use in the Isolation of B. Influenzae.
- Fleming, A. (1932) On the specific antibacterial properties of penicillin and potassium tellurite. Incorporating a method of demonstrating some bacterial antagonisms. *J. Pathol.* **35**: 831–842.
- Florey, H.W. (1946) The Use of Micro-organisms for Therapeutic Purposes. *The Yale journal of biology and medicine.*
- Forsberg, K.J., Patel, S., Gibson, M.K., Lauber, C.L., Knight, R., Fierer, N., and Dantas, G. (2014) Bacterial phylogeny structures soil resistomes across habitats. *Nature* **509**: 612–616.
- Forsberg, K.J., Reyes, A., Wang, B., Selleck, E.M., Sommer, M.O.A., and Dantas, G. (2012) The Shared Antibiotic Resistome of Soil Bacteria and Human Pathogens. *Science* **337**: 1107–1111.
- Forslund, K., Sunagawa, S., Roat Kultima, J., Mende, D., Arumugam, M., Typas, A., and Bork,

- P. (2013) Country-specific antibiotic use practices impact the human gut resistome. *Genome Res.*
- Garau, G., Di Guilmi, A.M., and Hall, B.G. (2005) Structure-Based Phylogeny of the Metallo-β-Lactamases. *Antimicrob Agents Chemother* **49**: 2778–2784.
- Gaze, W.H., Zhang, L., Abdousslam, N.A., Hawkey, P.M., Calvo-Bado, L., Royle, J., et al. (2011) Impacts of anthropogenic activity on the ecology of class 1 integrons and integron-associated genes in the environment. *ISME J* **5**: 1253–1261.
- Gillings, M.R. (2013) Evolutionary consequences of antibiotic use for the resistome, mobilome and microbial pangenome. *Front Microbiol* **4**: 4.
- Ginolhac, A., Jarrin, C., Gillet, B., Robe, P., Pujic, P., Tuphile, K., et al. (2004) Phylogenetic analysis of polyketide synthase I domains from soil metagenomic libraries allows selection of promising clones. *Appl Environ Microbiol* **70**: 5522–5527.
- Girardi, C., Greve, J., Lamshöft, M., Fetzer, I., Miltner, A., Schäffer, A., and Kästner, M. (2011) Biodegradation of ciprofloxacin in water and soil and its effects on the microbial communities. *Journal of Hazardous Materials* **198**: 22–30.
- Girvan, M.S., Bullimore, J., Pretty, J.N., Osborn, A.M., and Ball, A.S. (2003) Soil Type Is the Primary Determinant of the Composition of the Total and Active Bacterial Communities in Arable Soils. *Appl Environ Microbiol* **69**: 1800–1809.
- Goldstein, P.Z. and DeSalle, R. (2011) Integrating DNA barcode data and taxonomic practice: determination, discovery, and description. *Bioessays* **33**: 135–147.
- Gophna, U. and Ofran, Y. (2011) Lateral acquisition of genes is affected by the friendliness of their products. *Proc Natl Acad Sci USA* **108**: 343–348.
- Haas, D. and Défago, G. (2005) Biological control of soil-borne pathogens by fluorescent pseudomonads. *Nat Rev Microbiol* **3**: 307–319.
- Halary, S., Leigh, J.W., Cheaib, B., Lopez, P., and Bapteste, E. (2010) Network analyses structure genetic diversity in independent genetic worlds. *Proc Natl Acad Sci USA* **107**: 127–132.
- Hall, A.R. (2013) Genotype-by-environment interactions due to antibiotic resistance and adaptation in *Escherichia coli*. *J Evol Biol* **26**: 1655–1664.
- Handelsman, J., Rondon, M.R., Brady, S.F., Clardy, J., and Goodman, R.M. (1998) Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. *Chemistry & Biology* **5**: R245–R249.
- Heuer, H. and Smalla, K. (2007) Manure and sulfadiazine synergistically increased bacterial antibiotic resistance in soil over at least two months. *Environ Microbiol* **9**: 657–666.
- Heuer, H. and Smalla, K. (2012) Plasmids foster diversification and adaptation of bacterial populations in soil. *FEMS Microbiol Rev* **36**: 1083–1104.
- Hibbing, M.E., Fuqua, C., Parsek, M.R., and Peterson, S.B. (2009) Bacterial competition: surviving and thriving in the microbial jungle. *Nat Rev Microbiol* **8**: 15–25.
- Ho, P.L., Li, Z., Lai, E.L., Chiu, S.S., and Cheng, V.C.C. (2012) Emergence of NDM-1-producing Enterobacteriaceae in China. *J Antimicrob Chemother* **67**: 1553–1555.
- Ho, P.L., Lo, W.U., Yeung, M.K., Lin, C.H., Chow, K.H., Ang, I., et al. (2011) Complete Sequencing of pNDM-HK Encoding NDM-1 Carbapenemase from a Multidrug-Resistant *Escherichia coli* Strain Isolated in Hong Kong. *PLoS ONE* **6**.
- Hughes, V.M. and Datta, N. (1983) Conjugative plasmids in bacteria of the “pre-antibiotic” era. *Nature* **302**: 725–726.
- Humeniuk, C., Arlet, G., Gautier, V., Grimont, P., Labia, R., and Philippon, A. (2002) Beta-lactamases of *Kluyvera ascorbata*, probable progenitors of some plasmid-encoded CTX-

- M types. *Antimicrob Agents Chemother* **46**: 3045–3049.
- Jenke-Kodama, H. and Dittmann, E. (2009) Evolution of metabolic diversity: Insights from microbial polyketide synthases. *Phytochemistry* **70**: 1858–1866.
- Johnson, J.W., Fisher, J.F., and Mobashery, S. (2013) Bacterial cell-wall recycling. *Ann N Y Acad Sci* **1277**: 54–75.
- Jolivet-Gougeon, A., Kovacs, B., Le Gall-David, S., Le Bars, H., Bousarghin, L., Bonnaure-Mallet, M., et al. (2011) Bacterial hypermutation: clinical implications. *J Med Microbiol* **60**: 563–573.
- Kakirde, K., Parsley, L., and Liles, M. (2010) Size does matter: Application-driven approaches for soil metagenomics. *Soil Biology and Biochemistry*.
- Kallifidas, D., Kang, H.-S., and Brady, S.F. (2012) Tetracycline A, an MRSA-active antibiotic identified through induced expression of environmental DNA gene clusters. *J. Am. Chem. Soc.* **134**: 19552–19555.
- Kirk, J.L., Beaudette, L.A., Hart, M., Moutoglis, P., Klironomos, J.N., Lee, H., and Trevors, J.T. (2004) Methods of studying soil microbial diversity. *J Microbiol Methods* **58**: 169–188.
- Kümmerer, K. (2009) Antibiotics in the aquatic environment--a review--part I. *Chemosphere* **75**: 417–434.
- la Torre, de, A., Iglesias, I., Carballo, M., Ramírez, P., and Muñoz, M.J. (2011) An approach for mapping the vulnerability of European Union soils to antibiotic contamination. *Sci Total Environ* 1–8.
- Laskaris, P., Tolba, S., Calvo-Bado, L., Wellington, E.M., and Wellington, L. (2010) Coevolution of antibiotic production and counter-resistance in soil bacteria. *Environ Microbiol* **12**: 783–796.
- Leal, R.M.P., Alleoni, L.R.F., Tornisiello, V.L., and Regitano, J.B. (2013) Sorption of fluoroquinolones and sulfonamides in 13 Brazilian soils. *Chemosphere* **92**: 979–985.
- Lewis, I.M. (1929) BACTERIAL ANTAGONISM WITH SPECIAL REFERENCE TO THE EFFECT OF PSEUDOMONAS FLUORESCENS ON SPORE FORMING BACTERIA OF SOILS. *J Bacteriol* **17**: 89–103.
- Lewis, K. (2013) Platforms for antibiotic discovery. *Nat Rev Drug Discov* **12**: 371–387.
- Li, B., Qiu, Y., Glidle, A., McIlvenna, D., Luo, Q., Cooper, J., et al. (2014) Gradient microfluidics enables rapid bacterial growth inhibition testing. *Anal Chem* **86**: 3131–3137.
- Liebert, C., Hall, R., and Summers, A. (1999) Transposon Tn21, flagship of the floating genome. *Microbiology and Molecular Biology Reviews* **63**: 507–.
- Liu, Z., Wang, W., Zhu, Y., Gong, Q., Yu, W., and Lu, X. (2013) Antibiotics at subinhibitory concentrations improve the quorum sensing behavior of *Chromobacterium violaceum*. *FEMS Microbiol Lett* **341**: 37–44.
- Looft, T., Johnson, T.A., Allen, H.K., Bayles, D.O., Alt, D.P., Stedtfeld, R.D., et al. (2012) In-feed antibiotic effects on the swine intestinal microbiome. *Proc Natl Acad Sci USA* **109**: 1691–1696.
- Marshall, C.G., Lessard, I.A., Park, I., and Wright, G.D. (1998) Glycopeptide antibiotic resistance genes in glycopeptide-producing organisms. *Antimicrob Agents Chemother* **42**: 2215–2220.
- Martinez, J. and Sánchez, M. (2009) Functional role of bacterial multidrug efflux pumps in microbial natural ecosystems - Martinez - 2009 - FEMS Microbiology Reviews - Wiley Online Library. *FEMS microbiology* ....
- Martinez, J.L. (2008) Antibiotics and Antibiotic Resistance Genes in Natural Environments.

- Science* **321**: 365–367.
- Martinez, J.L., Fajardo, A., Garmendia, L., Hernandez, A., Linares, J.F., Martínez-Solano, L., and Sánchez, M.B. (2009) A global view of antibiotic resistance. *FEMS Microbiol Rev* **33**: 44–65.
- Martinez, J.L., Sánchez, M.B., Martínez-Solano, L., Hernandez, A., Garmendia, L., Fajardo, A., and Alvarez-Ortega, C. (2009) Functional role of bacterial multidrug efflux pumps in microbial natural ecosystems. *FEMS Microbiol Rev* **33**: 430–449.
- Massova, I. and Mobashery, S. (1998) Kinship and diversification of bacterial penicillin-binding proteins and beta-lactamases. *Antimicrob Agents Chemother* **42**: 1–17.
- Meroueh, S.O., Minasov, G., Lee, W., Shoichet, B.K., and Mobashery, S. (2003) Structural aspects for evolution of beta-lactamases from penicillin-binding proteins. *J. Am. Chem. Soc.* **125**: 9612–9618.
- Monier, J.-M., Demanèche, S., Delmont, T.O., Mathieu, A., Vogel, T.M., and Simonet, P. (2011) Metagenomic exploration of antibiotic resistance in soil. *Curr Opin Microbiol* **14**: 229–235.
- Moreno-Bondi, M.C., Marazuela, M.D., Herranz, S., and Rodriguez, E. (2009) An overview of sample preparation procedures for LC-MS multiclass antibiotic determination in environmental and food samples. *Anal Bioanal Chem* **395**: 921–946.
- Mummey, D., Holben, W., Six, J., and Stahl, P. (2006) Spatial stratification of soil bacterial populations in aggregates of diverse soils. *Microb Ecol* **51**: 404–411.
- Musovic, S., Klümper, U., Dechesne, A., Magid, J., and Smets, B.F. (2014) Long-term manure exposure increases soil bacterial community potential for plasmid uptake. *Env Mic Reports* **6**: 125–130.
- Nardelli, M., Scalzo, P.M., Ramírez, M.S., Quiroga, M.P., Cassini, M.H., and Centrón, D. (2012) Class 1 integrons in environments with different degrees of urbanization. *PLoS ONE* **7**: e39223.
- Nesme, J., Cécillon, S., Delmont, T.O., Monier, J.-M., Vogel, T.M., and Simonet, P. (2014) Large-Scale Metagenomic-Based Study of Antibiotic Resistance in the Environment. *Curr Biol*.
- Nordmann, P. (2005) Emergence of plasmid-mediated resistance to quinolones in Enterobacteriaceae. *J Antimicrob Chemother* **56**: 463–469.
- Nunan, N., Wu, K., Young, I.M., Crawford, J.W., and Ritz, K. (2003) Spatial distribution of bacterial communities and their relationships with the micro-architecture of soil. *FEMS Microbiol Ecol* **44**: 203–215.
- Ochman, H., Lawrence, J.G., and Groisman, E.A. (2000) Lateral gene transfer and the nature of bacterial innovation. *Nature* **405**: 299–304.
- Oskay, M. (2009) Antifungal and antibacterial compounds from Streptomyces strains. *African Journal of Biotechnology* **8**: 3007–3017.
- P, N., T, N., and L, P. (2011) Global Spread of Carbapenemase-producing Enterobacteriaceae. *Emerging Infect Dis* **17**: 1791–1798.
- Peay, K.G., Bruns, T.D., Kennedy, P.G., Bergemann, S.E., and Garbelotto, M. (2007) A strong species-area relationship for eukaryotic soil microbes: island size matters for ectomycorrhizal fungi. *Ecol Letters* **10**: 470–480.
- Perry, J.A. and Wright, G.D. (2013) The antibiotic resistance “mobilome”: searching for the link between environment and clinic. *Front Microbiol* **4**: 138.
- Philippot, L., Andersson, S.G.E., Battin, T.J., Prosser, J.I., Schimel, J.P., Whitman, W.B., and Hallin, S. (2010) The ecological coherence of high bacterial taxonomic ranks. *Nat Rev*

- Microbiol* **8**: 523–529.
- Pidot, S.J., Coyne, S., Kloss, F., and Hertweck, C. (2014) Antibiotics from neglected bacterial sources. *International Journal of Medical Microbiology* **304**: 14–22.
- Plough, H.H. (1945) Penicillin resistance of *Staphylococcus aur...* [Am J Clin Pathol. 1945] - PubMed - NCBI. *American journal of clinical pathology*.
- Poirel, L. and Rodriguez-Martinez, J.M. (2005) Origin of plasmid-mediated quinolone resistance determinant QnrA. *Antimicrobial agents ....*
- Poirel, L., Héritier, C., and Nordmann, P. (2004) Chromosome-encoded ambler class D beta-lactamase of *Shewanella oneidensis* as a progenitor of carbapenem-hydrolyzing oxacillinase. *Antimicrob Agents Chemother* **48**: 348–351.
- Prosser, J.I., Bohannan, B.J.M., Curtis, T.P., Ellis, R.J., Firestone, M.K., Freckleton, R.P., et al. (2007) The role of ecological theory in microbial ecology. *Nat Rev Microbiol* **5**: 384–392.
- Quinn, R. (2013) Rethinking antibiotic research and development: World War II and the penicillin collaborative. *Am J Public Health* **103**: 426–434.
- Ranjard, L. and Richaume, A. (2001) Quantitative and qualitative microscale distribution of bacteria in soil. *Res Microbiol* **152**: 707–716.
- Reith, J. and Mayer, C. (2011) Peptidoglycan turnover and recycling in Gram-positive bacteria. *Appl Microbiol Biotechnol* **92**: 1–11.
- Roberts, A.P. and Mullany, P. (2009) A modular master on the move: the Tn916 family of mobile genetic elements. *Trends Microbiol* **17**: 251–258.
- Roberts, M. (1996) Tetracycline resistance determinants: mechanisms of action, regulation of expression, genetic mobility, and distribution. *FEMS Microbiol Rev* **19**: 1–24.
- Roberts, M.C. (2005) Update on acquired tetracycline resistance genes. *FEMS Microbiol Lett* **245**: 195–203.
- Roesch, L.F.W., Fulthorpe, R.R., Riva, A., Casella, G., Hadwin, A.K.M., Kent, A.D., et al. (2007) Pyrosequencing enumerates and contrasts soil microbial diversity. *ISME J* **1**: 283–290.
- Schloss, P.D. and Handelsman, J. (2005) Metagenomics for studying unculturable microorganisms: cutting the Gordian knot. *Genome Biol* **6**: 229.
- Schlüter, A., Szczepanowski, R., Pühler, A., and Top, E.M. (2007) Genomics of IncP-1 antibiotic resistance plasmids isolated from wastewater treatment plants provides evidence for a widely accessible drug resistance gene pool. *FEMS Microbiol Rev* **31**: 449–477.
- Sen, D., Van Der Auwera, G.A., Rogers, L.M., Thomas, C.M., Brown, C.J., and Top, E.M. (2011) Broad-host-range plasmids from agricultural soils have IncP-1 backbones with diverse accessory genes. *Appl Environ Microbiol* **77**: 7975–7983.
- Seoane, J., Yankelevich, T., Dechesne, A., Merkey, B., Sternberg, C., and Smets, B.F. (2011) An individual-based approach to explain plasmid invasion in bacterial populations. *FEMS Microbiol Ecol* **75**: 17–27.
- Shen, B. (2003) Polyketide biosynthesis beyond the type I, II and III polyketide synthase paradigms. *Curr Opin Chem Biol*.
- Sidhu, M.S., Heir, E., Leegaard, T., Wiger, K., and Holck, A. (2002) Frequency of disinfectant resistance genes and genetic linkage with beta-lactamase transposon Tn552 among clinical staphylococci. *Antimicrob Agents Chemother* **46**: 2797–2803.
- Smillie, C., Garcillán-Barcia, M.P., Francia, M.V., Rocha, E.P.C., and la Cruz, de, F. (2010) Mobility of plasmids. *Microbiol Mol Biol Rev* **74**: 434–452.
- Smith, T. (2000) Antibiotics from soil bacteria. *Nat Struct Biol* **7**: 189–190.
- Soberon, J. (2005) Interpretation of models of fundamental ecological niches and species'

- distributional areas.
- Sommer, M.O.A. (2014) Microbiology: Barriers to the spread of resistance. *Nature* **509**: 567–568.
- Song, J.L., Au, K.H., Huynh, K.T., and Packman, A.I. (2014) Biofilm responses to smooth flow fields and chemical gradients in novel microfluidic flow cells. *Biotechnol Bioeng* **111**: 597–607.
- Song, J.S., Jeon, J.H., Lee, J.H., Jeong, S.H., and Jeong, B.C. (2005) Molecular characterization of TEM-type beta-lactamases identified in cold-seep sediments of Edison Seamount (south of Lihir Island, Papua New Guinea). *J Microbiol*.
- Stein, T. (2005) Bacillus subtilis antibiotics: structures, syntheses and specific functions - Stein - 2005 - Molecular Microbiology - Wiley Online Library. *Mol Microbiol*.
- Stokes, H.W. and Gillings, M.R. (2011) Gene flow, mobile genetic elements and the recruitment of antibiotic resistance genes into Gram-negative pathogens. *FEMS Microbiol Rev* **35**: 790–819.
- Sun, P., Barmaz, D., Cabrera, M.L., Pavlostathis, S.G., and Huang, C.-H. (2013) Detection and quantification of ionophore antibiotics in runoff, soil and poultry litter. *J Chromatogr A*.
- Thiele Bruhn, S. (2003) Pharmaceutical antibiotic compounds in soils—a review. *Journal of Plant Nutrition and Soil Science* **166**: 145–167.
- Thomas, C.M. and Nielsen, K.M. (2005) Mechanisms of, and barriers to, horizontal gene transfer between bacteria. *Nat Rev Microbiol* **3**: 711–721.
- Toleman, M.A., Spencer, J., Jones, L., and Walsh, T.R. (2012) bla<sub>NDM-1</sub> is a chimera, likely constructed in Acinetobacter baumannii. *Antimicrob Agents Chemother*.
- Tolls, J. (2001) Sorption of Veterinary Pharmaceuticals in Soils: A Review - Environmental Science & Technology (ACS Publications). *Environ Sci Technol*.
- Torsvik, V. (2011) Prokaryotic Diversity--Magnitude, Dynamics, and Controlling Factors. *Science* **296**: 1064–1066.
- Tsuey-Ching, Y., Hecht, D., Tsai, J.J.P., and Hu, R.-M. (2012) Cefoxitin is both an inhibitor of class A beta-lactamase of Xanthomonas campestris pv. campestris str. 17 and an inducer of its gene. *Res Microbiol* **163**: 550–556.
- Turnbaugh, P.J., Hamady, M., Yatsunencko, T., Cantarel, B.L., Duncan, A., Ley, R.E., et al. (2009) A core gut microbiome in obese and lean twins. *Nature* **457**: 480–484.
- Turrientes, M.C., Baquero, M.R., Sánchez, M.B., Valdezate, S., Escudero, E., Berg, G., et al. (2010) Polymorphic mutation frequencies of clinical and environmental Stenotrophomonas maltophilia populations. *Appl Environ Microbiol* **76**: 1746–1758.
- Villa, L., Poirel, L., Nordmann, P., Carta, C., and Carattoli, A. (2012) Complete sequencing of an IncH plasmid carrying the bla<sub>(NDM-1)</sub>, bla<sub>(CTX-M-15)</sub> and qnrB1 genes. *J Antimicrob Chemother* **67**: 1645–1650.
- Waksman, S.A. (1941) Antagonistic relations of microorganisms. *Bacteriological reviews*.
- Waksman, S.A. and Woodruff, H.B. (1942a) Selective Antibiotic Action of Various Substances of Microbial Origin. *J Bacteriol*.
- Waksman, S.A. and Woodruff, H.B. (1942b) Streptothricin, a New Selective Bacteriostatic and Bactericidal Agent, Particularly Active Against Gram-Negative Bacteria. *Experimental Biology and Medicine* **49**: 207–210.
- Waksman, S.A. and Woodruff, H.B. (1940) The Soil as a Source of Microorganisms Antagonistic to Disease-Producing Bacteria. *J Bacteriol* **40**: 581–600.
- Walsh, C. (2000) Molecular mechanisms that confer antibacterial drug resistance. *Nature* **406**: 775–781.

- Walsh, T.R., Weeks, J., Livermore, D.M., and Toleman, M.A. (2011) Dissemination of NDM-1 positive bacteria in the New Delhi environment and its implications for human health: an environmental point prevalence study. *Lancet Infect Dis* **11**: 355–362.
- Wang, X., Kang, Y., Luo, C., Zhao, T., Liu, L., Jiang, X., et al. (2014) Heteroresistance at the single-cell level: adapting to antibiotic stress through a population-based strategy and growth-controlled interphenotypic coordination. *MBio* **5**: e00942–13.
- Weldhagen, G.F. (2004) Integrons and beta-lactamases--a novel perspective on resistance. *Int J Antimicrob Agents* **23**: 556–562.
- Willingham, T.W. and Werth, C.J. (2008) Evaluation of the Effects of Porous Media Structure on Mixing-Controlled Reactions Using Pore-Scale Modeling and Micromodel Experiments - Environmental Science & Technology (ACS Publications). *Environmental science & ...*
- Wright, G.D. (2007) The antibiotic resistome: the nexus of chemical and genetic diversity. *Nat Rev Microbiol* **5**: 175–186.
- Yong, D., Toleman, M.A., Giske, C.G., Cho, H.S., Sundman, K., Lee, K., and Walsh, T.R. (2009) Characterization of a New Metallo-β-Lactamase Gene, blaNDM-1, and a Novel Erythromycin Esterase Gene Carried on a Unique Genetic Structure in *Klebsiella pneumoniae* Sequence Type 14 from India. *Antimicrob Agents Chemother* **53**: 5046–5054.
- Zhang, J., Kapli, P., Pavlidis, P., and Stamatakis, A. (2013) A general species delimitation method with applications to phylogenetic placements. *Bioinformatics*.

## Figure legends.

### Figure 1 - Stacked histogram showing the relative antibiotic and bioactive compounds percentage in known natural products from living organisms.

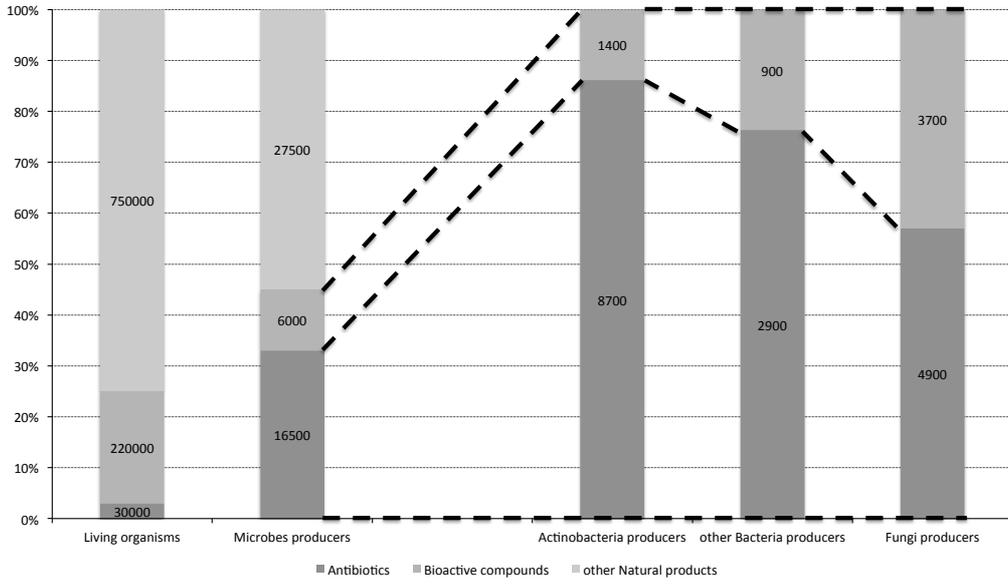
Stacked histograms are showing the proportion of known bioactive and antibiotics molecules (*i.e.*: compounds registered in BNPD; <http://naturalprod.ucsd.edu/>) produced by (from left to right): all living organisms from the three domain of Life; microorganisms only; Actinobacteria; other Bacteria and Fungi. All data are from Bérdy, J. (2005). Bioactive microbial metabolites. *Journal of Antibiotics* 58, 1–26.

### Figure 2 – The resistome: potential for antibiotic resistance.

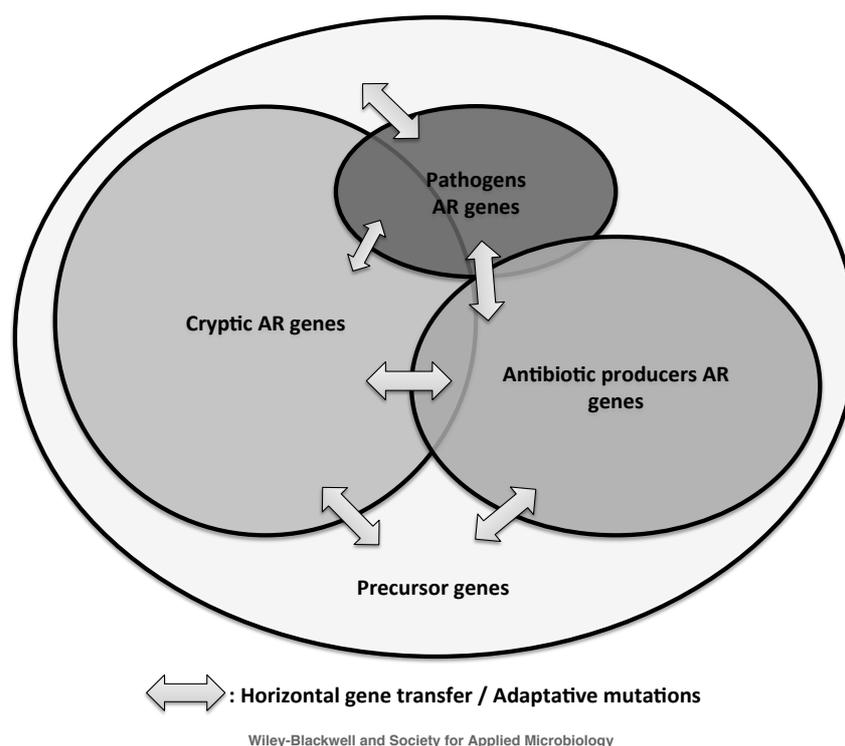
The resistome is all genes coding antibiotic resistance (AR) related proteins or similar protein likely to evolve toward potent antibiotic resistance. This encompass problematic AR gene identified in pathogenic bacteria, AR gene from antibiotic producers, cryptic AR genes (*i.e.* a resistance gene with no obvious association to antibiotic resistance) which may be or not expressed but could confer resistance in another genetic context for example and AR precursor genes coding protein with low level of affinity or resistance to antibiotic molecules. A significant part of those different genes subset overlaps due to sequence similarities and are most likely sharing a common evolutionary history. Genes can of course switch from one state to another through either through horizontal gene transfer (HGT) or point mutation or recombination. Adapted from Wright *et al.* Nat. Mic Review. (Wright,G.D, 2010).

### Figure 3 - Schematic representation of the soil ecosystem.

Soil is a complex and heterogeneous matrix with varying physicochemical conditions. From high to low nutrient or oxygen concentrations soil is structured in macro- and micro-aggregate structured by biotic and abiotic factors at different scale, forming a complex trophic network. While varying between soils, atmospheric gases are generally forming a diffusion gradient with highest concentrations found in topsoil while water is drained quickly and leaches soluble nutrients in a vertical gradient from top to bottom. Plants roots and fungi mycelium are colonizing spaces between mineral particles of varying size (from clay  $\leq 4\mu\text{m}$  to coarse pebbles  $< 16\text{mm}$ ), giving soil its structuration. Rhizosphere is defined as a soil under influence of plants roots exudates. This is in soil where the highest concentration of bacteria and organic nutrients are found. Soil is a rather patchy ecosystem in term of bacteria population with large areas of low bacterial cell density of and local hotspots of high cell density and high metabolic activity often associated with higher nutrient concentrations. Soil aggregates of different sizes are conditions are harboring different abundance and diversity of bacterial populations each struggling to find their ecological niche favoring a high overall genetic diversity in this environment. As a consequence, soil is considered as a significant reservoir of diverse antibiotic resistance and antibiotic production genes, among a whole range of other bioactive molecules. Adapted from Fortuna, A. (2012) The Soil Biota. *Nature Education Knowledge* 3(10):1.

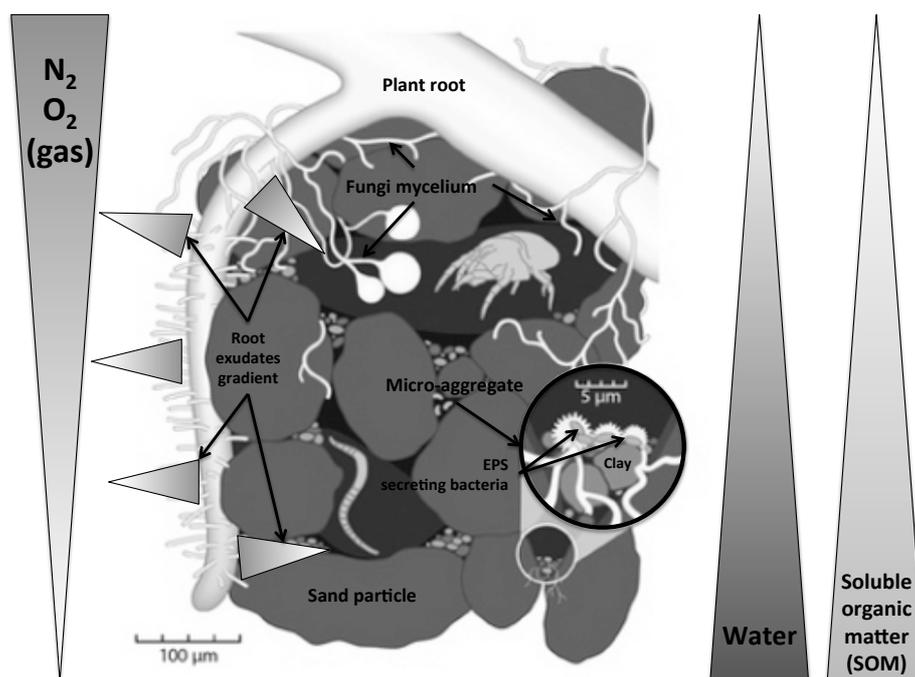


Wiley-Blackwell and Society for Applied Microbiology



**Figure 2 – The resistome: potential for antibiotic resistance.**

The resistome is all genes coding antibiotic resistance (AR) related proteins or similar protein likely to evolve toward potent antibiotic resistance. This encompass problematic AR gene identified in pathogenic bacteria, AR gene from antibiotic producers, cryptic AR genes (i.e. a resistance gene with no obvious association to antibiotic resistance) which may be or not expressed but could confer resistance in another genetic context for example and AR precursor genes coding protein with low level of affinity or resistance to antibiotic molecules. A significant part of those different genes subset overlaps due to sequence similarities and are most likely sharing a common evolutionary history. Genes can of course switch from one state to another through either through horizontal gene transfer (HGT) or point mutation or recombination. Adapted from Wright *et al.* Nat. Mic Review. (Wright,G.D, 2010).



Wiley-Blackwell and Society for Applied Microbiology

### Figure 3 - Schematic representation of the soil ecosystem.

Soil is a complex and heterogeneous matrix with varying physicochemical conditions. From high to low nutrient or oxygen concentrations soil is structured in macro- and micro-aggregate structured by biotic and abiotic factors at different scale, forming a complex trophic network. While varying between soils, atmospheric gases are generally forming a diffusion gradient with highest concentrations found in topsoil while water is drained quickly and leaches soluble nutrients in a vertical gradient from top to bottom. Plants roots and fungi mycelium are colonizing spaces between mineral particles of varying size (from clay  $\leq 4\mu m$  to coarse pebbles  $< 16mm$ ), giving soil its structuration. Rhizosphere is defined as a soil under influence of plants roots exudates. This is in soil where the highest concentration of bacteria and organic nutrients are found. Soil is a rather patchy ecosystem in term of bacteria population with large areas of low bacterial cell density of and local hotspots of high cell density and high metabolic activity often associated with higher nutrient concentrations. Soil aggregates of different sizes are conditions are harboring different abundance and diversity of bacterial populations each struggling to find their ecological niche favoring a high overall genetic diversity in this environment. As a consequence, soil is considered as a significant reservoir of diverse antibiotic resistance and antibiotic production genes, among a whole range of other bioactive molecules. Adapted from Fortuna, A. (2012) *The Soil Biota. Nature Education Knowledge* 3(10):1.





# Abstract.

Environmental bacteria and especially soil bacteria are active producers of antibiotic molecules and most drugs used nowadays are isolated from saprophytic soil bacteria and these microorganisms have also evolved numerous resistance pathways leading to an arsenal of Antibiotic Resistance Genes Determinants (ARGD) known as the environmental resistome. A survey of ARGD prevalence is required in order to characterize this natural phenomenon with critical implications in our current infectious diseases management. In order to perform such analysis we compiled a set of 71 metagenomic datasets from various environmental origins: soils, oceans, lakes, human feces, indoor air, etc., and compared their sequences with a database of known antibiotic resistance gene determinants (ARGD). ARGD-annotated reads are found in every environment analyzed confirming their ubiquity. Soil is found to be the richest and shares a large part of ARGD with the human gut microbiome, indicating ARGD transfers between these environments. Experiments using qPCR and metagenomic DNA sequencing on soil samples from two sites with known and distinct antibiotic pollution history were conducted to understand how ARGD abundance and diversity in soil are affected when impacted by antibiotic molecules. The first site is a reference soil from a long-term experiment without history of antibiotic pollution (Rothamsted Park Grass, UK). Soil microcosms are setup with addition of either antibiotic-containing animal manure or pure molecules and incubated for 6 months to monitor changes in ARGD concentration following these perturbations. Our second study-site is a very remote settlement in French Guiana where antibiotics are available since recently and may have impacted the local soil microbial community. Soil samples are taken following a line-transect going from the village (antibiotic source) to 3km deep in the forest in a gradient of human-impact. Our results all confirm prevalence of ARGD in soil at significant abundance but also that ARGD distribution is more correlated to environmental factors such as soil type, microbial taxonomy composition or microcosms incubation conditions than antibiotic molecules exposure in both sites. Pathogens ARGD diversity is far lower than ARGD diversity found in the environment and not all the soil resistome is readily accessible for transfer. In order to characterize the soil mobile gene pool, a strategy is proposed to isolate specifically mobile DNA directly from the environment for sequencing purposes. Better knowledge on the microbial ecology factors limiting ARGD transfers to pathogens may greatly help us reduce the current threat on our limited medical antibiotic molecules resource.