Characterization of antibiotic resistance genes in microbiomes from different ecosystems

Concepcion Sanchez-Cid Torres

▶ **To cite this version:**

Concepcion Sanchez-Cid Torres. Characterization of antibiotic resistance genes in microbiomes from different ecosystems. Genomics [q-bio.GN]. Université de Lyon, 2021. English. NNT : 2021LYSE1056 . tel-03827338

HAL Id: tel-03827338
https://theses.hal.science/tel-03827338

Submitted on 24 Oct 2022

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.

N°d'ordre NNT : 2021LYSE1056

THESE de DOCTORAT DE L'UNIVERSITE DE LYON
opérée au sein de
l'Université Claude Bernard Lyon 1

Ecole Doctorale N° accréditation 160
ELECTRONIQUE, ELECTROTECHNIQUE, AUTOMATIQUE

Spécialité de doctorat : Ingénierie du vivant

Soutenue publiquement le 30/03/2021, par :
Concepcion SANCHEZ-CID TORRES

Caractérisation de la diversité des gènes de résistance à des antibiotiques dans les microbiotes de différents écosystèmes environnementaux

Devant le jury composé de :

Gillings, Michael	Professeur - Macquarie University, Australie	**Rapporteur**
Wellington, Elizabeth	Professeure - University of Warwick, GB	**Rapporteuse**
Favre-Bonté, Sabine	Maître de conférences, HDR – Université Claude Bernard Lyon 1	**Examinatrice**
Nicol, Graeme	Directeur de recherche – Centre National de la Recherche Scientifique	**Président du jury**
Smalla, Kornelia	Professeure - Julius Kühn Institut, Allemagne	**Examinatrice**
Vogel, Timothy M.	Professeur – Université Claude Bernard Lyon 1	**Directeur de thèse**
Delaurière, Laurence	Docteur - Société Promega France	**Invitée**

N°d'ordre NNT: 2021LYSE1056

DOCTORAL THESIS at L'UNIVERSITE DE LYON
(University of Lyon)
Delivered by
l'Université Claude Bernard Lyon 1

Graduate School N° 160
ELECTRONICS, ELECTROTECHNIC, CONTROL SYSTEMS

Doctoral Specialty: BIO-ENGINEERING

Thesis defended on the 30th of March, 2021:
Concepcion Sanchez-Cid Torres

Characterization of Antibiotic Resistance Genes in Microbiomes from Different Ecosystems

Thesis committee:

Michael Gillings	Professor - Macquarie University, Australie	**Reporter**
Elizabeth Wellington	Professor - University of Warwick, GB	**Reporter**
Sabine Favre-Bonté	Associate Professor – Université Claude Bernard Lyon1	**Examiner**
Graeme Nicol	Research Director– Centre National de la Recherche Scientifique CNRS)	**Jury president**
Kornelia Smalla	Professor - Julius Kühn Institut, Allemagne	**Examiner**
Timothy M. Vogel	Professor – Université Claude Bernard Lyon1	**Thesis Director**
Laurence Delaurière	PhD - Promega France	**Invited Member**

Acknowledgements

First of all, I would like to thank my supervisor, Timothy M. Vogel, for leading and inspiring me during this three-year journey. Thank you for all our discussions about science and life, for your great sense of humor and for trusting and encouraging me all along this PhD. From all the decisions that I have not made in my life, being your PhD student with you was the best one. Thank you for always being there when I needed your help, for all the fun we've had in conclaves, conferences, wine tastings and lunches and, more importantly, for always having chocolate and/or cookies in your office.

I also would like to thank Trista Schagat, Céline Menager, Pascal Simonet and all those who trusted me in the first place and organized this collaboration between Promega and the Laboratoire Ampère. I thank the ANRT and its CIFRE program, as well as Promega, for funding this work. I will always be grateful to my colleagues in Promega for their support over this journey and before it started, and particularly to Laurence Delaurière, who has always been there to help me do this PhD under the best conditions. I particularly enjoyed working with you in the supervision of a master student, Alexandre Baldi, it was a great experience that taught me very important lessons.

I could not have hoped for better colleagues to accompany me during this adventure. Those who know me would be outraged if I said that it was a pleasure waking up in the morning to go to work, but working with you made it worth it. Thanks to all the PhD students who shared pain, laughs, knowledge, adventures, coffee and beers with me: Rose, Romie, Mia, Arthur, Benoit and Adrien, I am glad we went through this together. I also want to thank Catherine for trusting me to work with her, for our insightful discussions about science, for all the fun we have had together and for being so fabulous. Thanks to Christoph for his help, his sense of humor and his patience when a nagging Spanish student had mental breakdowns with lab work or sequencing data. Thanks to Graeme for his jokes, kindness and help, for his epic contributions to karaoke around the globe and for moving my ski glasses just before the lab group photo that everybody uses for their presentations, it really helps me look smart and serious. I could never forget Edith's invaluable help during these three years and her constant smile and cheerfulness. Thanks to Laure for her amazing work at organizing the lab, for her help during this journey and for her patience when things get messy, and to Christina for providing me with a lot of soil and for all the good moments that we've shared. And last but not least, thanks to Alice for being so kind, adorable and funny, and for opening the door for me every time that I forgot my badge.

I would like to thank the members of the jury, Elizabeth M. Wellington (University of Warwick, UK), Michael Gillings (Macquarie University, Australia), Kornelia Smalla (Julius Kühn Institut, Germany) and Sabine Favre-Bonté (UCBL, Lyon), for having accepted to evaluate this work and for being a source of inspiration.

I have been lucky to cross paths with some wonderful people who have worked with me during this PhD. Thanks to the mountain fun team from the Polish Academy of Science, Bartek, Adam, Karol and Lukasz, for the amazing week we spent doing fieldwork in the Sudety Mountains in March 2020, for your patience with my poor skiing skills and for organizing our epical scape from Poland before borders got closed. I would like to thank Professor Tong Zang in the University of Hong Kong and his wonderful students for welcoming me in their lab for a month and a half, it was an amazing experience that I will not easily forget. Finally, I would like to thank Alexandre Guironnet, Laure Wiest and Emmanuelle Vulliet for measuring gentamicin concentrations in soil and water using HPLC-MS/MS.

I could not finish without thanking my lovely and supporting parents, my family in Seville and my family in Lyon. This experience would not have been the same without the support and love I have received from Lucas, Clàudia, Manel, Jorge, Cris and all the great friends that I have met here.

Résumé (français)

Depuis la découverte des antibiotiques au cours du XXème siècle, l'antibiothérapie a considérablement réduit la mortalité causée par les bactéries pathogènes. Cependant, au cours des dernières décennies, la communauté humaine est continuellement confrontée à de nouveaux cas cliniques de résistance aux antibiotiques et à l'émergence de clones multirésistants dans le monde entier, ce qui peut entraîner l'échec de l'antibiothérapie et le début d'une ère post-antibiotiques. Les antibiotiques et la résistance aux antibiotiques sont sur terre depuis des milliers d'années avant la découverte des antibiotiques. Cependant, l'industrialisation et l'utilisation extensive des antibiotiques chez l'homme et l'animal ont imposé une pression sélective sans précédent sur les communautés bactériennes, accélérant le développement de la résistance aux antibiotiques à l'échelle mondiale. Les antibiotiques sont utilisés dans le monde entier non seulement pour traiter les maladies causées par des agents pathogènes humains, mais aussi à des fins thérapeutiques et de stimulation de la croissance dans les fermes, l'aquaculture et l'agriculture. En raison des activités anthropiques, des concentrations résiduelles d'antibiotiques, des gènes de résistance aux antibiotiques et des bactéries résistantes aux antibiotiques atteignent l'environnement principalement via les stations d'épuration des eaux usées, la fertilisation du fumier ou des déchets des usines de production d'antibiotiques. Cela peut entraîner le développement et la sélection de résistance aux antibiotiques dans l'environnement et la dissémination des gènes de résistance aux antibiotiques et des bactéries résistantes aux antibiotiques de l'environnement aux microbiomes animaux et humains. Néanmoins, l'amplitude de ce phénomène reste inconnue.

L'objectif de cette thèse était d'évaluer la réponse du microbiome et du résistome environnementaux à la pollution chimique par des antibiotiques, ainsi qu'à la pollution biologique causée par les activités humaines. Dans un premier temps, une étude méthodologique comparant cinq méthodes d'extraction d'ADN et quatre méthodes d'extraction d'ARN a été réalisée pour évaluer l'effet de l'extraction d'ADN et de la profondeur de séquençage sur la découverte de la richesse taxonomique, fonctionnelle et de gènes de résistance aux antibiotiques du sol. Alors que la profondeur de séquençage avait un impact plus fort que l'extraction d'ADN sur la découverte de la richesse taxonomique et fonctionnelle, une grande variabilité dans la découverte de la richesse de gènes de résistance aux antibiotiques a été observée entre les triplicats, quelle que soit la profondeur de séquençage. En outre, certaines méthodes semblaient mesurer une richesse de gènes de résistance aux antibiotiques plus élevée que d'autres. Sur la base de cette étude, une méthode semi-automatisée a été sélectionnée pour l'extraction d'ADN à partir de sols pollués par des antibiotiques.

Dans un second temps, deux études utilisant des microcosmes ont été menées pour évaluer les effets de la pollution par des antibiotiques sur le microbiome et le résistome d'un sol agricole de la Côte de Saint André (France) et de l'eau du Rhône (France) en utilisant une combinaison des cultures et d'approches métagénomiques/qPCR. Les concentrations sous-inhibitrices et inhibitrices de gentamicine pour les bactéries du sol et de l'eau enrichies en milieu de culture ont été déterminées en évaluant la croissance bactérienne. Les microcosmes du sol ont été pollués avec une gamme de concentrations inhibitrices de gentamicine, tandis que les microcosmes de l'eau ont été pollués avec deux concentrations sous-inhibitrices et une concentration inhibitrice, afin d'établir des liens entre la dose de gentamicine et l'ampleur de la réponse dans le microbiome et le résistome environnementaux à différents temps d'exposition. Ces deux études illustrent comment les effets du même antibiotique sur différents environnements sont fortement dépendants des facteurs environnementaux et des

propriétés physico-chimiques. Alors que des concentrations inhibitrices allant jusqu'à 1 mg de gentamicine par gramme de sol étaient fortement adsorbées sur les particules du sol et n'avaient pas d'effets significatifs sur le microbiome ni le résistome du sol après 8 jours d'exposition, une concentration sous-inhibitrice de 50 ng de gentamicine par ml d'eau pendant 2 jours d'exposition impactait la composition des communautés bactériennes totales et actives et l'abondance et l'expression des gènes de résistance à la gentamicine. Ces résultats supportent la crainte que les concentrations sous-inhibitrices d'antibiotiques puissent sélectionner des résistances dans l'environnement et méritent donc plus d'attention lors de l'évaluation des risques associés à la pollution environnementale par des antibiotiques. De plus, cette recherche souligne les limites des termes «sous-inhibiteur» et «inhibiteur» dans des environnements complexes et l'importance d'effectuer des études en utilisant des microcosmes et des études de terrain afin d'évaluer les effets de la pollution par des antibiotiques sur le résistome environnemental.

Pour terminer, une étude de terrain a été menée sur des échantillons de neige obtenus des Sudety Mountains (Pologne) soumis à différentes expositions aux activités humaines et à différentes quantités de végétation autour, afin d'évaluer l'impact de facteurs anthropiques et environnementaux sur le microbiome et le résistome de la neige en utilisant une approche métagénomique / qPCR. Cette recherche fournit des éléments qui supportent l'hypothèse selon laquelle des facteurs environnementaux et anthropiques ont un impact sur l'écologie de la neige et induisent des changements dans le microbiome et le résistome de la neige en fournissant aux communautés bactériennes des niveaux plus élevés de carbone organique et d'autres nutriments. Cela favoriserait la croissance d'une communauté bactérienne plus abondante, qui à son tour augmenterait l'abondance des gènes de résistance aux antibiotiques et pourrait stimuler la compétition et la prolifération de ces gènes dans la neige. Étant donné que les activités anthropiques induisent des modifications sur le résistome de la neige en ayant un impact limité sur les composants principaux du microbiome, les effets de cette pollution sont probablement générés par un apport accru de matière organique provenant de déchets générés par les humains plutôt que par l'apport direct de gènes de résistance aux antibiotiques et de bactéries résistantes aux antibiotiques du microbiome humain. Cette pollution organique pourrait stimuler le développement d'une résistance aux antibiotiques dans le microbiome de la neige qui pourrait ensuite être disséminée dans l'atmosphère ou à travers la fonte des neiges. Les recherches menées dans cette étude mettent en évidence la nécessité de faire des études sur le développement de la résistance aux antibiotiques dans des environnements pollués par les activités humaines et la prise en compte des sources organiques de pollution en plus des polluants biologiques (gènes de résistance aux antibiotiques et bactéries résistantes aux antibiotiques).

Mots clés : résistance aux antibiotiques, résistome environnemental, gènes de résistance aux antibiotiques, pollution anthropique, développement de résistances, extraction d'ADN, gentamicine, concentrations sous-inhibitrices, concentrations inhibitrices, métagénomique, séquençage Illumina MiSeq, séquençage Oxford Nanopore, qPCR / RT-qPCR.

Abstract

Since the discovery of antibiotics during the 20th century, antibiotic therapy has drastically reduced the mortality caused by bacterial pathogens. However, during the last few decades, the human community is continuously facing new cases of antibiotic resistance in clinics and the emergence of multidrug resistant (MDR) clones worldwide that result in antibiotic therapy failure and the beginning of a post-antibiotic era. Antibiotics and antibiotic resistance predate the discovery of antibiotics by millions of years. However, the industrialization and extensive use of antibiotics in humans and animals has imposed a selective pressure without precedent on bacterial communities and accelerated the development of antibiotic resistance at a global scale. Antibiotics are used worldwide to treat diseases caused by human pathogens and to promote therapeutic and growth in farms, aquacultures and crops. As a consequence of anthropogenic activities, residual concentrations of antibiotics, antibiotic resistance genes (ARGs) and antibiotic resistant bacteria (ARB) reach the environment mainly through wastewater treatment plants, manure fertilization or release from antibiotic producing plants. This may result in the development and selection of antibiotic resistance in environmental settings and the dissemination of ARB and ARGs from the environment to animal and human microbiomes. Nonetheless, the scope of this phenomenon remains unclear.

The goal of this thesis was to evaluate the response of the environmental microbiome and resistome to chemical pollution with antibiotics and to biological pollution caused by anthropogenic activities. A methodological study compared five DNA extraction methods and four RNA extraction methods to evaluate the effect of DNA extraction and sequencing depth on taxonomic, functional and ARG richness discovery in soil. Whereas sequencing depth had a stronger impact than DNA extraction on taxonomic and functional richness discovery, a high variability in ARG richness discovery was observed between triplicates regardless of sequencing depth. Furthermore, some methods arguably measured a higher ARG richness than others. Based on this study, a semi-automated method was selected for DNA extraction from antibiotic polluted soils.

Two microcosm studies were carried out to evaluate the effects of antibiotic pollution on the microbiome and resistome of an agricultural soil from La Côte de Saint André (France) and of water from the Rhône river (France) using a combination of culture-based and metagenomics/qPCR approaches. Gentamicin sub-inhibitory and inhibitory concentrations in soil and water bacteria enriched in culture media were determined by evaluating bacterial growth. Soil microcosms were polluted with a range of inhibitory concentrations of gentamicin, whereas water microcosms were polluted with two sub-inhibitory concentrations and an inhibitory concentration, in order to establish links between gentamicin dose and the magnitude of the response in the environmental microbiome and resistome at different exposure times. These two studies illustrate how the effects of the same antibiotic on different environments are strongly dependent of environmental factors and physicochemical properties. Whereas inhibitory concentrations up to 1 mg of gentamicin per gram of soil were highly adsorbed onto soil particles and did not have significant effects on the soil microbiome nor resistome after 8-days exposure, a sub-inhibitory concentration of 50 ng of gentamicin per ml of water impacted in the composition of bacterial total and active communities and the abundance and expression of gentamicin resistance genes during 2-days exposure. These findings support the concern that sub-inhibitory concentrations of antibiotics may select for resistance in the environment and therefore deserve more attention when assessing the risks associated to antibiotic environmental pollution. In addition, this research underlines the limitations of the terms “sub-inhibitory” and “inhibitory” in complex environments and the importance of microcosm and field studies for the evaluation of the effects of antibiotic pollution on the environmental resistome.

Finally, a field study was carried out with snow from the Sudety Mountains (Poland) with a range of exposure to human activities and surrounding vegetation in order to evaluate the impact of both anthropogenic and environmental factors on the snow microbiome and resistome using a metagenomics and qPCR approach. This research supports the hypothesis that both environmental and anthropogenic factors impact snow ecology and induce changes in the snow microbiome and antibiotic resistome by providing bacterial communities with higher levels of organic carbon and other nutrients. This would support the growth of a more abundant bacterial community, which in turn increases the abundance of the antibiotic resistome and could stimulate competition and ARG proliferation in snow. Since anthropogenic activities induce changes in the snow resistome without having a strong impact on the core microbiome, the effects of this pollution are probably caused by an increased input of organic matter from anthropogenic waste rather than by human microbiome ARB and ARG. This organic pollution could stimulate the development of antibiotic resistance in the snow microbiome that might be subsequently disseminated through the atmosphere or snow melting. The research carried out in this study highlights the need for survey of antibiotic resistance development in anthropogenic polluted sites and the consideration of organic sources of pollution in addition to biological pollutants (ARB and ARGs).

Keywords : antibiotic resistance, environmental resistome, ARGs, anthropogenic pollution, resistance development, DNA extraction, gentamicin, sub-inhibitory concentrations, inhibitory concentrations, metagenomics, Illumina MiSeq sequencing, Oxford Nanopore sequencing, qPCR/RT-qPCR.

Index of contents

Acknowledgements **1**

Résumé (français) **3**

Abstract **5**

List of publications **10**

Synthèse (français) **11**

Chapitre I. Introduction et hypothèses 11

Chapitre II. Optimisation de l'extraction des acides nucléiques du sol 14

Chapitre III. Evaluation des effets de la pollution du sol par de la gentamicine sur le mcirobiome et le résistome du sol 16

Chapitre IV. Evaluation de la réponse du microbiome et résistome de l'eau à des concentrations sous-inhibitrices de gentamicine 18

Chapitre V. Evaluation des effets des facteurs anthropiques et environnementaux sur le microbiome et le résistome de la neige 20

Chapitre VI. Conclusions et perspectives 22

Bibliographie 25

Chapter I. A review on the effects of anthropogenic environmental pollution on antibiotic resistance development and dissemination **29**

Abstract 29

A brief introduction of antibiotics and antibiotic resistance 30

Antibiotic resistance dissemination: Horizontal Gene Transfer (HGT) and Mobile Genetic Elements (MGE) 33

Human, animal and environmental microbiomes are interconnected 36

Consequences of anthropogenic activity on resistance development and dissemination in the environment 39

PhD objectives, approach and working hypothesis 42

References 44

Chapter II. Optimization of nucleic acid extraction from soil. Evaluation of the effect of nucleic acid extraction and sequencing depth on taxonomic, functional and ARG discovery **55**

Abstract 55

Introduction 55

Materials and methods 56

Results 59

Discussion 66

References 69

Chapter III. Gentamicin adsorption onto soil particles prevents it from having overall short-term effects on the soil microbiome and resistome **71**

Abstract 71

Introduction 71

Materials and methods 72

Results 75

Discussion 81

References 83

Chapter IV. Gentamicin at sub-inhibitory concentrations induces a response in the environmental microbiome and resistome **86**

Abstract 86

Introduction 87

Materials and methods 88

Results 92

Discussion 98

Additional research 101

References 106

Chapter V. Environmental and anthropogenic factors shape the snow microbiome and resistome **110**

Abstract 110

Introduction 110

Materials and methods 111

Results 114

Discussion 121

References 123

Chapter VI. Conclusions and perspectives **126**

Effect of DNA/RNA extraction and sequencing depth on richness discovery 126

Impact of antibiotic pollution on resistance development in the environment 127

Impact of anthropogenic activity on resistance development in the environment 130

References 131

Annex. Supplementary Information **132**

Chapter II 132

Chapter III 157

Chapter IV 162

Chapter V 166

List of publications

Publications

Concepcion Sanchez-Cid, Alexandre Guironnet, Laure Wiest, Emmanuelle Vulliet, Timothy M. Vogel. 2021. **Gentamicin adsorption onto soil particles prevents it from having overall short-terms effects on the soil microbiome and resistome.** *Antibiotics* 10:191. https://doi.org/10.3390/antibiotics10020191

Manuscripts *in prep*

Concepcion Sanchez-Cid, Alexandre Guironnet, Christoph Keuschnig, Laure Wiest, Emmanuelle Vulliet, Tong Zang, Timothy M. Vogel. **Gentamicin at sub-inhibitory concentrations induces a response in the environmental microbiome and resistome.**

Concepcion Sanchez-Cid, Christoph Keuschnig, Karol Torzewski, Lukasz Stachnik, Bartek Luks, Adam Nawrot, Timothy M. Vogel, Catherine Larose. **Environmental and anthropogenic factors shape the snow microbiome and antibiotic resistome.**

Concepcion Sanchez-Cid, Romie Tignat-Perrier, Alexandre Baldi, Laure Franqueville, Laurence Delaurière, Trista Schagat, Timothy M. Vogel. **Sequencing depth has a stronger effect than nucleic acid extraction on soil bacterial richness discovery.**

Side publications

Alexandre Guironnet, **Concepcion Sanchez-Cid,** Laure Wiest, Timothy M. Vogel, Emmanuelle Vulliet. 2021. **Aminoglycosides analysis optimization using Ion pairing Liquid Chromatography coupled to tandem Mass Spectrometry and application on wastewater samples.** *Journal of Chromatography A*, Elsevier, pp.462133. https://doi.org/10.1016/j.chroma.2021.462133

Concepcion Sanchez-Cid, Benoit Bergk-Pinto, Christoph Keuschnig, Cédric Malandain, Sébastien Kaskassian, Timothy M. Vogel. **Metagenomic support of chlorinated solvent contaminated site remediation.** Manuscript *in prep.*

Synthèse (français)

Chapitre I. Introduction et hypothèses

Depuis la découverte des antibiotiques au cours du XXème siècle, l'antibiothérapie a considérablement réduit la mortalité causée par les bactéries pathogènes. Cependant, au cours des dernières décennies, la communauté humaine est continuellement confrontée à de nouveaux cas cliniques de résistance aux antibiotiques[1,2] et à l'émergence de clones multirésistants dans le monde entier, ce qui peut entraîner l'échec de l'antibiothérapie et le début d'une ère post-antibiotiques. Les antibiotiques et la résistance aux antibiotiques sont sur terre depuis des milliers d'années avant la découverte des antibiotiques[3,4]. Cependant, l'industrialisation et l'utilisation extensive des antibiotiques chez l'homme et l'animal ont imposé une pression sélective sans précédent sur les communautés bactériennes, accélérant le développement de la résistance aux antibiotiques à l'échelle mondiale[5]. Les antibiotiques sont utilisés dans le monde entier non seulement pour traiter les maladies causées par des agents pathogènes humains, mais aussi à des fins thérapeutiques et de stimulation de la croissance dans les fermes, l'aquaculture et l'agriculture. En raison des activités anthropiques, des concentrations résiduelles d'antibiotiques, des gènes de résistance aux antibiotiques et des bactéries résistantes aux antibiotiques atteignent l'environnement principalement via les stations d'épuration des eaux usées, la fertilisation du fumier ou les déchets des usines de production d'antibiotiques. Cela peut entraîner le développement et la sélection de résistance aux antibiotiques dans l'environnement et la dissémination des gènes de résistance aux antibiotiques et des bactéries résistantes aux antibiotiques de l'environnement aux microbiomes animaux et humains. Néanmoins, l'amplitude de ce phénomène reste inconnue.

Les bactéries peuvent être intrinsèquement résistantes à un antibiotique[6] ou acquérir une résistance par mutations de leurs gènes ou par transfert horizontal de gènes[7]. Des mutations dans les gènes bactériens qui ont un impact sur l'activité antibiotique peuvent se produire pendant la réplication cellulaire, être sélectionnées et générer des clones par transfert vertical[8]. La souche porteuse de cette mutation peut proliférer sous la pression des antibiotiques, ce qui lui confère un grand avantage sélectif. Cependant, les mutations génèrent une diminution de la forme physique et sont souvent perdues en l'absence de pression sélective[9]. D'autre part, la résistance aux antibiotiques peut se propager à partir de bactéries contenant un gène de résistance vers des bactéries sensibles partageant la même niche au moyen du transfert horizontal de gènes. Il a été démontré que les événements de transfert horizontal des gènes contribuent au développement de la majorité des résistances aux antibiotiques actuellement trouvées chez les pathogènes[10–13]. Cependant, la contribution de l'évolution verticale à la menace de résistance aux antibiotiques ne doit pas être négligée, car plusieurs souches mutantes ont abouti à des pathogènes d'importance clinique[14–16]. Les trois principaux mécanismes de HGT sont la conjugaison, la transduction et la transformation.

Malgré le fait que la production d'antibiotiques et la résistance aux antibiotiques se produisent naturellement dans l'environnement, l'utilisation d'antibiotiques dans la thérapie humaine et la production alimentaire génère un flux de concentrations résiduelles d'antibiotiques, ainsi que des bactéries résistantes aux antibiotiques et des gènes de résistance aux antibiotiques vers l'environnement[17,18] (Figure 1). Ces polluants peuvent atteindre l'environnement non seulement par contact direct avec les humains, mais aussi par les effluents des hôpitaux, les usines de traitement des eaux usées, par contact direct avec les cultures contaminées, l'urine et les matières fécales des

animaux ou par des applications de fumier pour la fertilisation des sols. Les antibiotiques des usines de production peuvent également être rejetés accidentellement dans l'environnement. À l'échelle mondiale, les polluants biologiques (bactéries résistantes aux antibiotiques et gènes de résistance aux antibiotiques) peuvent intégrer le résistome environnemental, d'où ils peuvent se disséminer vers le microbiome humain. De plus, la pression sélective imposée par des concentrations résiduelles d'antibiotiques sur les microbiomes humains, animaux et environnementaux peut favoriser le développement des résistances aux antibiotiques[19,20], le transfert horizontal de gènes de résistance aux antibiotiques aux bactéries sensibles[21] et la dissémination des bactéries résistantes aux antibiotiques entre les trois compartiments. Les bactéries résistantes aux antibiotiques et les gènes de résistance aux antibiotiques environnementaux peuvent disséminer aux microbiomes animaux par contact direct, et au microbiome humain par contact direct ou par l'eau potable[22,23], les légumes ou les vecteurs animaux. Par conséquent, la pollution de l'environnement par des antibiotiques et des résidus issus des activités humaines pourrait entraîner le développement de nouvelles résistances dans les cliniques.

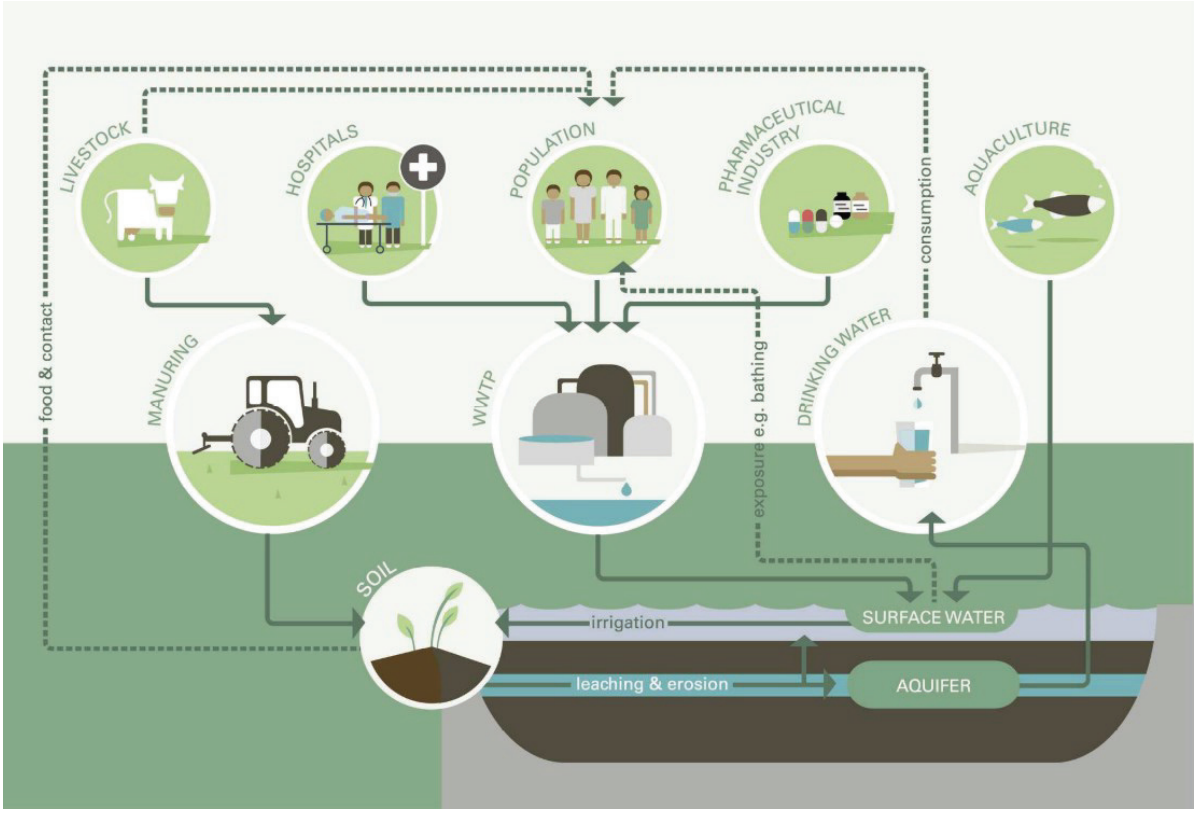


Figure 1. Voies de diffusion des antibiotiques, des bactéries résistantes aux antibiotiques et des gènes de résistance aux antibiotiques entre les humains, les animaux et l'environnement. En raison des activités anthropiques, les concentrations résiduelles d'antibiotiques, de bactéries résistantes aux antibiotiques et de gènes de résistance aux antibiotiques atteignent l'environnement principalement par les effluents hospitaliers, les stations d'épuration des eaux usées, la fertilisation du fumier, les aquacultures ou les déchets des usines de production d'antibiotiques. Cela peut entraîner le développement et la sélection de la résistance aux antibiotiques dans l'environnement et la diffusion des bactéries résistantes aux antibiotiques et de gènes de résistance aux antibiotiques entre différents compartiments environnementaux et de l'environnement vers les microbiomes animaux et humains, soit par contact direct avec des animaux ou des environnements contaminés, soit par la consommation alimentaire ou la consommation d'eau. Figure extraite de Vikesland et al, 2017 118.

L'objectif de cette thèse était d'évaluer les effets de la pollution anthropique de l'environnement sur le développement et la dissémination des résistances aux antibiotiques dans l'environnement. Deux types de polluants ont été analysés séparément. Dans un premier temps, les effets de la contamination chimique par des antibiotiques ont été évalués dans les microcosmes du sol et de l'eau. L'hypothèse principale était que la gentamicine serait fortement adsorbée sur les particules du sol et sa biodisponibilité serait réduite peu de temps après l'administration. Ceci ajouté à la faible activité des bactéries du sol et à la probabilité plus faible de contact dans une matrice solide entraînerait un manque d'impact visible de la gentamicine, même à des concentrations élevées, sur le microbiome du sol ou la composition de son résistome à grande échelle. Si la gentamicine avait un effet sur le génome bactérien, celui-ci serait couvert par la grande diversité du microbiome du sol et la faible abondance de la plupart de ses composants. En ce qui concerne les effets des concentrations sous-inhibitrices de gentamicine sur le résistome de l'eau, le contact accru entre les bactéries de l'eau -par rapport aux bactéries du sol- et la plus grande disponibilité de la gentamicine dans l'eau entraînerait un environnement plus propice pour favoriser le développement et la dissémination de la résistance. Ainsi, des concentrations sous-inhibitrices de gentamicine induiraient des changements dans la communauté bactérienne et augmenteraient l'abondance et l'expression des gènes de résistance aux antibiotiques sans inhiber la croissance globale des communautés bactériennes dans l'eau.

Deuxièmement, des échantillons de neige polluée et non-polluée par des humains ont été analysés pour déterminer si les activités anthropiques ont provoqué une augmentation des polluants biologiques dans la neige. L'hypothèse principale de cette recherche était que les facteurs environnementaux et anthropiques ont un impact sur l'écologie de la neige et induisent des changements dans le microbiome de la neige et son résistome en fournissant aux communautés bactériennes des niveaux plus élevés de carbone organique et d'autres nutriments. Cela favoriserait la croissance d'une communauté bactérienne plus abondante, ce qui à son tour augmenterait l'abondance des gènes de résistance aux antibiotiques. Des recherches antérieures menées par notre groupe montrent que des niveaux accrus d'acides organiques induisent une compétition et augmentent l'abondance des ARG dans la neige, et qu'ils favorisent également l'établissement d'une communauté plus active métaboliquement. Par conséquent, les sources anthropiques et environnementales de carbone organique pourraient stimuler la concurrence et la prolifération des ARG dans la neige.

Chapitre II. Optimisation de l'extraction des acides nucléiques du sol

Une étude méthodologique était nécessaire afin d'évaluer les biais liés à l'extraction de l'ADN et de l'ARN du sol et le séquençage avant évaluation des effets de la pollution antibiotique. L'écosystème du sol abrite l'une des plus grandes diversités de micro-organismes de tous les écosystèmes[24]. Il est essentiel d'accéder à la totalité des communautés et fonctions du microbiome du sol pour mieux comprendre le rôle de ces communautés microbiennes dans le fonctionnement du sol et son écosystème. L'utilisation des techniques de séquençage de nouvelle génération (NGS ou Next Generation Sequencing) a amélioré notre accès aux communautés microbiennes présentes dans le sol, en particulier à la grande proportion de micro-organismes non cultivés[25]. Néanmoins, chaque étape méthodologique, de l'échantillonnage du sol à l'annotation des séquences, présente des biais inhérents qui limitent la profondeur et la fiabilité des analyses du microbiome du sol[26]. Parmi tous ces biais, ceux associés à l'extraction d'ADN ont été particulièrement mis en évidence pour leurs effets[27,28]. L'ADN peut être adsorbé par des composés du sol tels que l'argile, qui, lorsqu'ils sont combinés à la présence de bactéries récalcitrantes de lyse, réduisent l'efficacité d'extraction de l'ADN. De plus, la matière organique et les acides humiques connus pour inhiber potentiellement les réactions enzymatiques sont souvent co-extraits avec l'ADN. Ces biais s'appliquent également à l'extraction d'ARN[29], qui est de plus en plus utilisée pour identifier des communautés et des gènes potentiellement actifs. De nombreuses études ont comparé les méthodes d'extraction de l'ADN et de l'ARN et documenté les biais imposés par les procédures de lyse. Cela a suscité des propositions de différentes méthodes d'extraction et de purification des acides nucléiques du sol au cours des dernières décennies dans le but d'obtenir une image impartiale de la biodiversité du microbiome du sol. Néanmoins, aucune méthode ne permet de surmonter tous les biais décrits ci-dessus et le débat sur le choix des méthodes d'extraction nucléique est toujours en cours.

D'autre part, la profondeur de séquençage a également montré un effet important sur la découverte bactérienne. Plusieurs études ont montré que de faibles profondeurs de séquençage peuvent biaiser l'évaluation des communautés microbiennes, à la fois en termes de richesse et de diversité[30–32]. Cependant, la contribution relative de la profondeur de séquençage et de l'extraction d'ADN ou d'ARN à la découverte de la richesse bactérienne du sol n'a pas encore été évaluée. Dans cette étude, nous avons comparé la richesse phylogénétique et fonctionnelle détectée par deux nouvelles méthodes semi-automatisées d'extraction et de purification d'ADN à celle mesurée par deux kits commerciaux et la méthode phénol/chloroforme telle que décrite par Griffiths et al.[33] Puisque les fonctions de résistance aux antibiotiques sont rares dans le microbiome environnemental, la découverte de la richesse des gènes de résistance aux antibiotiques peut être plus inégale en fonction de la profondeur de séquençage et donc suivre un modèle différent de la découverte de la richesse fonctionnelle. Ainsi, la richesse en gènes de résistance aux antibiotiques détectée par ces cinq méthodes a également été évaluée. De plus, nous avons évalué l'effet de la profondeur de séquençage sur la richesse taxonomique mesurée par deux méthodes semi-automatisées pour l'extraction et la purification d'ARN à celle mesurée par un kit commercial et la méthode phénol/chloroforme. Les acides nucléiques ont été extraits de deux échantillons de sol et la région hypervariable V3-V4 du gène 16S de l'ARNr ou de son ADNc a été séquencée pour déterminer l'unité taxonomique opérationnelle (OTU ou Operational Taxonomic Unit). La richesse taxonomique mesurée par chaque méthode a été définie au niveau du genre pour chaque séquence en utilisant la base de données Ribosome Database Project (RDP) et le classificateur bayésien RDP[34]. Les séquences de métagénomique obtenues à partir d'échantillons d'ADN ont été annotées à l'aide de MEGAN6[35] et des sous-systèmes hiérarchiques SEED.

Le troisième niveau de la classification SEED a été sélectionné pour l'analyse de découverte de la richesse fonctionnelle. Enfin, les gènes de résistance aux antibiotiques ont été identifiés en utilisant la base de données CARD[36].

Le principal résultat de cette étude était que la profondeur de séquençage a un impact plus fort sur la découverte de la richesse taxonomique et fonctionnelle que le choix d'une méthode d'extraction, et que la raréfaction des séquences à une profondeur égale conduit à l'attribution de genres et de classes fonctionnelles uniques qui ont été réellement détectés par plus d'une méthode avant raréfaction. Par conséquent, le choix d'une méthode d'extraction d'ADN ou d'ARN semble moins critique que la nécessité d'optimiser les techniques de séquençage à haut débit pour étudier les communautés microbiennes du sol et leurs fonctions associées. En revanche, la découverte de la richesse des gènes de résistance aux antibiotiques est fortement impactée par la variation de la composition du résistome entre échantillons, et une seule méthode (Maxwell 1) a montré la découverte de richesse plus élevée et plus cohérente que les autres dans le sol de la Côte de Saint André, l'échantillon de sol qui allait être utilisé pour les expériences sur la pollution environnementale par des antibiotiques. Par conséquent, bien que cette richesse accrue ne soit sûrement pas exclusivement due à l'extraction d'ADN et que des précautions doivent être prises avant de préconiser cette méthode pour l’analyse du résistome du sol, compte tenu de la globalité des résultats obtenus dans cette étude, cette méthode a été choisie pour l'extraction d'ADN dans l'évaluation des effets de la pollution du sol par des antibiotiques.

Chapitre III. Evaluation des effets de la pollution du sol par de la gentamicine sur le microbiome et le résistome du sol

La plupart des antibiotiques actuellement utilisés en thérapie humaine et en production alimentaire ont été isolés à partir de bactéries et de champignons du sol[37]. Une conséquence naturelle de cette production d'antibiotiques par les bactéries du sol est le développement des résistances aux antibiotiques dans le sol. Le sol est considéré comme l'un des principaux réservoirs environnementaux des gènes de résistance aux antibiotiques[38]. Des gènes de résistance aux antibiotiques impliqués dans des cas de résistances dans les hôpitaux, ainsi que des nouveaux gènes de résistance aux antibiotiques, ont été identifiés même dans des sols quasiment pas soumis aux activités humaines[39–42], montrant que la résistance aux antibiotiques se produit dans le sol même en l'absence d'une forte pression sélective anthropique. De plus, les gènes de résistance aux antibiotiques présents dans le sol sont souvent associés à des éléments génétiques mobiles[43] et, par conséquent, peuvent être transférés à la fois à d'autres bactéries environnementales et à des agents pathogènes humains[44].

L'utilisation d'antibiotiques en agriculture et l'application de fumier d'animaux traités avec des antibiotiques pour la fertilisation des sols augmentent la pression sélective des antibiotiques et la quantité de bactéries résistantes aux antibiotiques dans l'environnement[45,46]. Cette pression sélective peut stimuler le développement des résistances aux antibiotiques dans le sol et leur mobilisation et transfert vers des bactéries causant des infections chez l'humain. Par conséquent, les risques liés à l'utilisation d'antibiotiques en agriculture doivent être évalués afin de réglementer leur utilisation. De nombreuses études au cours des deux dernières décennies ont analysé les effets du compostage du fumier pollué par des antibiotiques sur le microbiome du sol. Ils ont montré que ce fumier peut augmenter l'abondance des bactéries résistantes aux antibiotiques, des gènes de résistance aux antibiotiques et des éléments génétiques mobiles[47–52]. Cependant, le sol est une matrice très complexe et diverse qui change avec le temps. Plusieurs facteurs, tels que les caractéristiques du sol (c'est-à-dire le pourcentage d'eau, la concentration d'oxygène ou la disponibilité des nutriments), la réduction de la biodisponibilité des antibiotiques en raison de leur adsorption dans les particules du sol et l'activité et la résilience des bactéries du sol peuvent modifier les effets des antibiotiques dans le sol. De plus, le sol est une matrice solide et le contact physique entre les bactéries est réduit par rapport aux environnements liquides. Ces différences peuvent empêcher l'évaluation précise du transfert horizontal des gènes de résistance aux antibiotiques des bactéries résistantes aux bactéries sensibles dans le sol. Les caractéristiques du sol peuvent fournir celui-ci avec une forte capacité de tamponner la pollution par des antibiotiques[53], en réduisant l'impact des antibiotiques sur le microbiome du sol et les conséquences potentielles de la pollution des sols par des antibiotiques sur la santé humaine. Ainsi, il est difficile d'obtenir une image globale des effets de la pollution par des antibiotiques sur le développement des résistances dans le sol, et après des décennies de recherche, l'amplitude de ce phénomène reste à élucider.

Cette étude compare les effets à court terme d'une gamme de concentrations de gentamicine sur le microbiome et le résistome des enrichissements bactériens et des microcosmes d'un sol agricole de La Côte de Saint André (France) en utilisant une approche métagénomique. La gentamicine est un aminoglycoside utilisé à la fois en thérapie humaine et en production alimentaire, et les gènes conférant une résistance à la gentamicine sont largement distribués dans l'environnement[54]. Nous avons émis l'hypothèse que la pollution par de la gentamicine a une plus grande influence sur les

bactéries du sol dans des milieux enrichis que dans les microcosmes, où l'adsorption de la gentamicine sur les particules du sol réduit sa biodisponibilité.

Les concentrations sous-inhibitrices et inhibitrices de gentamicine dans les bactéries du sol enrichies en milieu de culture ont été déterminées en évaluant la croissance bactérienne. Ensuite, des microcosmes du sol ont été pollués avec des concentrations inhibitrices de gentamicine ayant jusqu'à 1 mg par gramme de sol, alors que des cultures des bactéries du sol ont été polluées avec une concentration sous-inhibitrice et une concentration inhibitrice de gentamicine. La différence de réponse entre les communautés dans des milieux enrichis et des milieux environnementaux à plusieurs temps d'exposition a été évaluée en utilisant une combinaison d'approches dépendant de la culture et par métagénomique. L'ADN a été extrait des enrichissements bactériens et des microcosmes du sol, et sa biomasse a été approximativement estimée par le nombre de copies du gène 16S de l'ARNr. De plus, la région hypervariable V3-V4 du gène 16S de l'ARNr et les métagénomes de l'ADN extrait ont été séquencés, et l'abondance des gènes de résistance à la gentamicine a été évaluée par qPCR. Finalement, la fraction biodisponible de gentamicine dans le sol a été mesurée par HPLC-MS / MS.

Alors que dans les enrichissements bactériens du sol, une concentration inhibitrice de 12 µg de gentamicine par ml de milieu réduit considérablement la taille des communautés bactériennes et altère leur composition, dans les microcosmes du sol, où la gentamicine est fortement adsorbée sur les particules du sol et la diversité bactérienne est plus élevée, les concentrations inhibitrices jusqu'à 1 mg de gentamicine par gramme de sol n'a pas eu d'effets significatifs sur la croissance bactérienne, ni sur la composition du microbiome et du résistome du sol, après une exposition de 8 jours (chapitre III). Les analyses de HPLC-MS/MS ont montré que la gentamicine est fortement adsorbée dans les particules du sol, ce qui l'empêche d'avoir des effets sur les microcosmes du sol. Ceci expliquerait aussi partiellement les différences observées entre la réponse des bactéries du sol dans les microcosmes et les milieux enrichis, où la gentamicine n'est pas adsorbée sur les particules du sol et peut avoir un effet immédiat sur les bactéries du sol. Cette étude démontre les différences entre la réponse des communautés bactériennes à la pollution antibiotique dans les milieux enrichis et dans leur matrice environnementale et expose les limites des études basées sur la culture dans la surveillance de la résistance aux antibiotiques. En outre, il est nécessaire d'établir des liens entre les effets de la pollution antibiotique et les propriétés du sol.

Chapitre IV. Evaluation de la réponse du microbiome et résistome de l'eau à des concentrations sous-inhibitrices de gentamicine

Les concentrations d'antibiotiques trouvées dans les environnements pollués par des humains sont souvent sous-inhibitrices[55], c'est-à-dire trop faibles pour provoquer une inhibition visible de la croissance des bactéries sensibles en culture[56]. Cependant, des concentrations sous-inhibitrices d'antibiotiques peuvent provoquer une perturbation de la communauté microbienne. Ils peuvent souvent ralentir la croissance bactérienne[57] et la pression sélective qu'ils exercent peut être suffisante pour compenser le coût de la résistance, contribuant ainsi à la sélection des résistances. En fait, un nombre croissant d'études montre que des concentrations sous-inhibitrices d'antibiotiques peuvent sélectionner des bactéries résistantes aux antibiotiques[58] et induire un large éventail de réponses, y compris le quorum sensing, la persistance, la formation de biofilms et l'expression de gènes impliqués dans la résistance aux antibiotiques, la virulence et la réponse SOS[59]. En outre, des concentrations sous-inhibitrices d'antibiotiques peuvent sélectionner des mutants avec un coût de maintenance inférieur à ceux sélectionnés à des concentrations inhibitrices[57] et peuvent être suffisantes pour le maintien des plasmides de multirésistance[60]. Ces concentrations peuvent également stimuler l'intégration des gènes de résistance aux antibiotiques dans les éléments génétiques mobiles et la diffusion via le transfert horizontal des gènes[61–63]. Par conséquent, certains antibiotiques peuvent sélectionner des résistances indépendamment de leur dose et des concentrations inférieures à la concentration minimale inhibitrice des bactéries sensibles pourraient contribuer au développement des résistances autant que des concentrations inhibitrices.

La majorité des études publiées visant à élucider les effets des concentrations sous-inhibitrices d'antibiotiques sur les communautés bactériennes et leur résistome fondent leurs résultats sur des approches basées sur la culture bactérienne. Les bactéries doivent posséder des gènes impliqués dans la résistance aux antibiotiques qui sont traduits en protéines pour pousser dans des milieux pollués par des antibiotiques. Par conséquent, les approches basées sur la culture fournissent des preuves phénotypiques de la résistance aux antibiotiques et des mécanismes impliqués. Cependant, ces approches visent une partie réduite de la communauté bactérienne dans des conditions in vitro établies et ne tiennent pas nécessairement compte de la complexité des milieux naturels, où la croissance bactérienne et les interactions sont influencées par plusieurs facteurs et les effets des antibiotiques sur le microbiome peuvent varier. Les études métagénomiques évitent les biais des études basées sur des cultures et permettent d'accéder à une proportion beaucoup plus large du microbiome environnemental sans altérer les conditions de la matrice environnementale. Néanmoins, les gènes et taxons identifiés à l'aide d'approches métagénomiques ne répondent pas nécessairement activement à l'antibiotique. Ainsi, la combinaison d'études métagénomiques avec des analyses basées sur l'ARN devrait permettre de mieux comprendre les taxons et les gènes impliqués dans la réponse aux concentrations sous-inhibitrices d'antibiotiques.

Dans cette étude, nous avons cherché à déterminer si les concentrations sous-inhibitrices d'antibiotiques induisent une réponse au niveau taxonomique ou fonctionnel dans le milieu aquatique et à établir des relations entre la dose d'antibiotique et l'ampleur de la réponse, en utilisant la gentamicine comme modèle. La gentamicine est un aminoglycoside utilisé à la fois pour la thérapie humaine et vétérinaire, et des concentrations résiduelles de gentamicine excrétée ont été détectées dans les stations d'épuration des eaux usées[64]. Il a été démontré que la gentamicine à des concentrations sous-inhibitrices induisait le développement des résistances dans des cultures pures[65].

De plus, les gènes de résistance aux aminosides sont largement distribués dans les intégrons chromosomiques et plasmidiques de classe 1[66], qui sont considérés comme des facteurs majeurs de dissémination des résistances aux antibiotiques[67]. Ainsi, en utilisant une combinaison de métagénomique et d'analyses basées sur l'ARN, les bactéries et les gènes impliqués dans la réponse environnementale à la gentamicine à des concentrations sous-inhibitrices dans les microcosmes d'eau de rivière ont été identifiés et des cassettes d'intégrons de classe 1 ont été séquencées pour cribler les gènes de résistance aux aminosides. Nous avons émis l'hypothèse que les concentrations inhibitrices de gentamicine induiraient une réponse plus forte que les concentrations sous-inhibitrices, mais que les dernières provoqueraient des changements dans la composition des communautés microbiennes et augmenteraient l'abondance et la transcription des gènes liés à la résistance aux aminosides que l'on peut trouver en classe 1 intégrons. Des microcosmes d'eau du Rhône ont été pollués avec deux concentrations sous-inhibitrices et une concentration inhibitrice de gentamicine pour évaluer l'ampleur de la réponse du microbiome et du résistome de l'eau à cette pollution. Les effets de la gentamicine à des concentrations inhibitrices et sous-inhibitrices à différents temps d'exposition ont été évalués en analysant la croissance globale et les niveaux de transcription, la composition des communautés bactériennes totales et actives et l'abondance et la transcription de l'ARG sur les microcosmes de l'eau. Les intégrons de classe 1 ont été séquencés et criblés pour les gènes de résistance aux aminosides, et des relations dose-réponse ont été établies entre les concentrations de gentamicine mesurées dans l'eau par HPLC-MS / MS et l'abondance et l'expression des gènes de résistance aux aminoglycosides mesurées par qPCR / RT-qPCR.

Les résultats obtenus montrent qu'une concentration sous-inhibitrice de 50 ng de gentamicine par ml d'eau, qui n'a pas eu d'incidence sur la taille de la communauté bactérienne, a provoqué une augmentation significative des niveaux globaux de transcription après 2 jours d'exposition et induit des changements dans la composition des communautés bactériennes totales et actives, augmentant l'abondance des Limnohabitans après seulement 1 jour d'exposition. De plus, l'assemblage hybride de séquences métagénomiques courtes et longues séquencées à l'aide des technologies Illumina MiSeq et Oxford Nanopore, respectivement, a permis l'obtention d'un Limnohabitans MAG contenant une séquence partielle du gène *aac(6')-Ib8,* un gène codant pour une aminoglycoside acétyltransférase impliquée dans la résistance à la gentamicine. Finalement, la pollution par la gentamicine à 50 ng/ml a induit une augmentation de l'abondance et de la transcription des gènes *aac(6')* après 2 jours d'exposition, bien que l'ampleur de cette réponse soit inférieure à 800 ng/ml, une concentration inhibitrice de gentamicine qui provoque l'inhibition de la plupart des membres de la communauté et la sélection de clones résistants à la gentamicine. Ces gènes sont au moins en partie présents dans les intégrons de classe 1 et pourraient donc être mobilisés et disséminés dans l'environnement et aux microbiomes humain et animal.

Cette étude est la première à montrer que des concentrations sous-inhibitrices de gentamicine induisent une réponse dans le microbiome et le résistome environnementaux *in vivo* et établissent des liens entre la dose d'antibiotique et l'ampleur de la réponse. Ces résultats supportent la crainte que des concentrations sous-inhibitrices d'antibiotiques puissent sélectionner des résistances dans l'environnement[56] et méritent donc plus d'attention lors de l'évaluation des risques associés à la pollution environnementale par les antibiotiques.

Chapitre V. Evaluation des effets des facteurs anthropiques et environnementaux sur le microbiome et le résistome de la neige

Les montagnes, et en particulier leur couverture neigeuse, sont un indicateur sensible du changement climatique. Les montagnes abritent environ un tiers de toutes les espèces terrestres et fournissent de l'eau à près de la moitié de la population mondiale[68]. Les montagnes se réchauffent à un rythme similaire à celui de l'Arctique[69] , avec des impacts sur l'environnement (rétrécissement des glaciers, dégradation du pergélisol, inondations, changements dans la distribution des espèces et des écosystèmes) et sur les sociétés (tourisme d'été et d'hiver). Une source majeure d'incertitude est la mesure dans laquelle les activités humaines interagissent avec les facteurs climatiques pour modifier les processus biogéochimiques et la santé et le fonctionnement des écosystèmes. Parmi les activités connues pour avoir un impact sur les écosystèmes alpins, le tourisme est souvent cité. Le tourisme peut avoir une gamme d'impacts positifs et négatifs sur les écosystèmes, les communautés et les économies de montagne[70]. Les activités touristiques impliquent souvent le développement et l'utilisation intensive de pistes, de sentiers et de pistes sportives par des véhicules, des transports non motorisés et des piétons. La présence des visiteurs est également généralement concentrée dans de petites zones, ce qui contribue à l'augmentation du bruit et des déchets. Les effets environnementaux négatifs du tourisme peuvent inclure le défrichement de la végétation et l'érosion des sols, la modification des paysages et des débits d'eau critiques, la pollution de l'eau et de l'air, et le déplacement de la faune ou des changements de comportement[70]. L'introduction d'espèces et de maladies exotiques et envahissantes peut également avoir un impact négatif important sur l'environnement[71].

De plus, la neige est reliée à l'atmosphère, qui est considérée comme l'une des principales sources de micro-organismes dans la neige et permet le transport à grande vitesse de la poussière atmosphérique dans la neige[72]. Des gènes de résistance aux antibiotiques et des agents pathogènes polluant l'atmosphère peuvent être déposés sur Terre par les chutes de neige, ce qui pourrait augmenter leur dissémination[73,74]. De plus, les microorganismes présents dans la neige sont mieux adaptés au transport atmosphérique que d'autres microorganismes[75] et les bactéries résistantes aux antibiotiques des sites de neige pollués pourraient potentiellement disséminer vers des environnements éloignés. Des bactéries résistantes aux antibiotiques et des gènes de résistantes aux antibiotiques présents dans la neige pourraient également être disséminés dans l'eau douce par la fonte des neiges[76] , ce qui à son tour pourrait augmenter le risque de maladies d'origine hydrique.

Malgré ces impacts potentiels, l'effet du tourisme d'hiver sur la santé et le fonctionnement des écosystèmes de montagne reste à élucider. L'objectif de cette étude est de pallier ce manque d'information en étudiant l'impact du tourisme d'hiver sur les écosystèmes de montagne enneigés. Des échantillons de neige ont été prélevés sur deux bassins situés dans le parc national de Karkonosze, dans les Sudety Mountains en Pologne; un bassin Czarny Kociol Jagniatkowski non affecté et un bassin Kociol Malego Stawu soumis à des activités humaines bien documentées telles que le tourisme et le développement de chalets. De plus, des échantillons présentant différents niveaux de végétation ont été récupérés dans les deux bassins.

L'hypothèse principale de cette recherche était que les facteurs environnementaux et anthropiques ont un impact sur l'écologie de la neige et induisent des changements dans le microbiome de la neige et le résistome antibiotique en fournissant aux communautés bactériennes des niveaux plus élevés de

carbone organique et d'autres nutriments. Cela favoriserait la croissance d'une communauté bactérienne plus abondante, ce qui à son tour augmenterait l'abondance du résistome. Des recherches antérieures menées par notre groupe montrent que des niveaux accrus d'acides organiques induisent une compétition et augmentent l'abondance des gènes de résistance aux antibiotiques dans la neige[77], et qu'ils favorisent également l'établissement d'une communauté plus métaboliquement active[78]. Par conséquent, les sources anthropiques et environnementales de carbone organique pourraient stimuler la concurrence et la prolifération des gènes de résistance aux antibiotiques dans la neige. L'impact de la végétation et de l'activité anthropique sur le microbiome et le résistome de la neige a été évalué à l'aide d'une approche métagénomique (séquençage de la région V3-V4 du gène 16S de l'ARNr et séquençage métagénomique) et qPCR (amplification du gène 16S de l'ARNr). La taille, la composition, la richesse et la diversité des communautés bactériennes ont été les paramètres choisis pour déterminer les effets des facteurs anthropiques et environnementaux sur le microbiome de la neige, tandis que leurs effets sur le résistome des neiges ont été évalués par la détermination de la taille et de la composition du résistome, ainsi que la richesse et la diversité des gènes de résistance aux antibiotiques.

Les résultats de cette étude appuient l'hypothèse proposée. Premièrement, des niveaux plus élevés de végétation environnante ont augmenté la biomasse bactérienne, réduit la diversité des genres et façonné la composition des communautés bactériennes dans la neige. En fait, plusieurs genres qui étaient plus abondants dans les sites avec des niveaux plus élevés de végétation environnante utilisent le carbone comme source de nutriments, et les différences d'abondance des genres se sont accrues parallèlement aux différences dans la végétation environnante. De plus, la taille du résistome a été augmentée dans les sites avec plus de végétation environnante, et cette augmentation était liée à l'augmentation de la biomasse bactérienne. Enfin, le nombre de gènes de résistance aux antibiotiques conférant une résistance à toutes les classes d'antibiotiques identifiées dans la neige était plus élevé dans les sites avec des niveaux plus élevés de végétation environnante, tout comme l'abondance de plusieurs gènes de résistance aux antibiotiques.

Concernant l'activité anthropique, ces recherches suggèrent qu'elle induit des modifications du résistome des neiges tout en ayant un impact limité sur la composition des communautés bactériennes. Bien que des niveaux de biomasse bactérienne plus élevés aient été détectés dans les sites soumis à une activité anthropique, les facteurs anthropiques n'ont pas montré un fort impact sur la richesse et la diversité des genres ou la composition globale de la communauté, et les différences trouvées dans l'abondance des genres liées à l''activité anthropique étaient beaucoup plus faibles que celles liées à la végétation environnante. Étant donné que les sites soumis à une activité anthropique étaient entourés de végétation et que les bactéries du microbiome humain ont peu de chances de survivre dans des environnements extrêmes, les effets supplémentaires des déchets anthropiques peuvent ne pas nécessairement avoir un impact sur la communauté bactérienne principale. Cependant, l'augmentation de la taille de la communauté bactérienne dans les sites pollués par l'homme a été suivie d'une augmentation de la taille du résistome et du nombre de séquences de gènes de résistances aux antibiotiques associées à plusieurs classes d'antibiotiques. Enfin, neuf gènes impliqués dans la résistance à six classes d'antibiotiques ont été trouvés plus abondants dans les sites pollués anthropiques que dans la neige vierge. Ainsi, ces résultats soutiennent l'hypothèse que les activités humaines génèrent des quantités accrues de déchets qui fournissent au microbiome de la neige un supplément de carbone organique, stimulant la compétition et la prolifération des ARG de la même manière que la présence de végétation.

Chapitre VI. Conclusions et perspectives

La pollution anthropique de l'environnement avec des concentrations résiduelles d'antibiotiques et de polluants biologiques (bactéries résistantes aux antibiotiques et gènes de résistance aux antibiotiques) est une source de préoccupation majeure. Les antibiotiques peuvent imposer une pression sélective dans le microbiome environnemental stimulant le développement et la dissémination de résistances aux antibiotiques, potentiellement même à des concentrations sous-inhibitrices. De plus, la contamination biologique peut contribuer au réservoir du résistome environnemental, et les nutriments organiques présents dans les déchets générés par les activités humaines peuvent avoir un impact sur l'écologie microbienne environnementale et favoriser des changements dans le résistome environnemental. Dans les deux cas, les résistances aux antibiotiques présents dans l'environnement sont susceptibles de disséminer aux microbiomes humain et animal, aggravant encore la crise antibiotique à laquelle nous sommes actuellement confrontés. En revanche, les interactions entre les polluants chimiques et biologiques et le microbiome environnemental récepteur sont affectées par des facteurs environnementaux et les propriétés physico-chimiques de la matrice environnementale. Des recherches sont fortement nécessaires pour déterminer i) l'amplitude de la réponse du microbiome et du résistome environnementaux aux polluants chimiques et biologiques et la relation entre cette réponse et les facteurs environnementaux, ii) le potentiel de mobilisation et de dissémination de la résistance aux antibiotiques dans l'environnement, et iii) le risque de transfert des ARB et ARG environnementaux vers le microbiome humain et animal.

L'objectif principal de cette thèse était de fournir des éléments pour aider à élucider des questions majeures concernant le développement des résistances aux antibiotiques dans l'environnement sous pollution anthropique. Cette recherche a été divisée en trois axes principaux :

Dans un premier temps, une étude méthodologique visant à répondre à la question «Est-ce que la découverte de la richesse taxonomique et fonctionnelle/des gènes de résistance dans le sol est plus impactée par les méthodes d'extraction des acides nucléiques ou par la profondeur de séquençage?». Les résultats de cette étude montrent que la profondeur de séquençage a un impact plus important sur la découverte de la richesse taxonomique et fonctionnelle que le choix d'une méthode d'extraction, et que la raréfaction des séquences à une profondeur égale conduit à l'attribution de genres et de classes fonctionnelles uniques qui ont été réellement détectés par plus d'une méthode avant raréfaction. Par conséquent, le choix d'une méthode d'extraction d'ADN ou d'ARN semble moins critique que la nécessité d'optimiser les techniques de séquençage à haut débit pour étudier les communautés microbiennes du sol et leurs fonctions associées. En revanche, la découverte de la richesse des gènes de résistance aux antibiotiques est fortement impactée par la variation de la composition du résistome entre échantillons. Des variations entre méthodes ont également été détectées.

Les biais associés à la profondeur de séquençage Illumina MiSeq ont été considérés tout au long de cette thèse. Il a été observé que la diversité bactérienne plus élevée dans les microcosmes du sol que dans les enrichissements bactériens du sol était probablement la cause de la qualité inférieure de l'assemblage obtenu à partir des séquences métagénomiques du microcosme du sol, soulignant la nécessité de technologies de séquençage plus approfondies pour couvrir la diversité bactérienne du sol et analyser avec précision le résistome du sol en utilisant des approches métagénomiques **(chapitre III).** De plus, le séquençage à lecture longue utilisant la technologie Oxford Nanopore et l'assemblage hybride de lectures courtes et longues ont montré une amélioration de la complétion et de la définition

des génomes assemblés à partir de métagénomes de l'eau, prouvant son importance dans l'analyse métagénomique du résistome de communautés complexes **(chapitre IV)**. En outre, des analyses de qPCR ont été effectuées pour compléter le séquençage d'Illumina MiSeq lors de l'évaluation de l'abondance des gènes de résistance aux antibiotiques **(chapitres III et IV),** et le besoin d'une validation qPCR supplémentaire des différences d'abondance de gènes de résistance aux antibiotiques identifiées par du séquençage métagénomique a été signalé lorsque ces analyses n'ont pas pu être effectuées en raison de contraintes de temps **(chapitre V).**

Deuxièmement, la réponse à la pollution par la gentamicine aux concentrations inhibitrices et sous-inhibitrices a été évaluée dans deux environnements, le sol et l'eau douce, afin de répondre aux questions «Les concentrations sous-inhibitrices de gentamicine peuvent-elles induire une réponse dans le microbiome environnemental et le résistome?» , "Quel est le lien entre la dose de gentamicine et l'ampleur de cette réponse?" et «Cette réponse est-elle cohérente entre les environnements?». Ces deux études illustrent comment les effets du même antibiotique sur différents environnements sont fortement dépendants des facteurs environnementaux et des propriétés physico-chimiques. Cela surligne la nécessité d'une surveillance de la pollution par des antibiotiques dans un large éventail de contextes environnementaux et de la mesure systématique des concentrations d'antibiotiques et la caractérisation des propriétés physico-chimiques de l'environnement, ainsi que de séries temporelles pour tenir compte des changements saisonniers du milieu environnemental et de son microbiome associé.

Ces études soulignent également les limites des termes «sous-inhibiteur» et «inhibiteur» dans des environnements complexes. Puisque ces termes sont définis dans des cultures de souches isolées, ils peuvent ne pas être précis pour illustrer ce qui se passe chez des communautés environnementales complexes dans leur milieu naturel. Les recherches menées dans cette thèse montrent que des concentrations sous-inhibitrices peuvent stimuler le développement de résistances *in vivo* et que des concentrations inhibitrices peuvent ne pas impacter le résistome environnemental, en fonction des interactions entre l'antibiotique, la matrice environnementale et son microbiome. Ainsi, ces résultats suggèrent qu'il faut préconiser des études sur des microcosmes et de terrain lorsque cela est possible, car les études basées sur la culture ciblent une proportion réduite du microbiome environnemental et ne prennent pas en compte les facteurs environnementaux. Par conséquent, la réponse observée à l'aide de ces approches ne reflète pas nécessairement ce qui se passe dans l'environnement.

Néanmoins, ces études évaluent les effets à court terme d'une dose unique d'un seul antibiotique dans une seule matrice environnementale. Cette approche laisse plusieurs questions sans réponse. Premièrement, la biodisponibilité de la gentamicine pourrait augmenter à plusieurs doses ou à des temps d'exposition plus longs dans le sol et provoquer des perturbations à long terme sur les communautés microbiennes à un niveau superficiel, et l'étude sur l'eau douce ne fournit aucune information sur la réponse du microbiome environnemental à des temps d'exposition plus longs ni de la persistance de la résistance en l'absence de pression sélective imposée par l'antibiotique. Deuxièmement, différents sols ou environnements d'eau douce pourraient répondre différemment au même antibiotique, et des modèles devraient être établis pour relier la composition de la matrice environnementale et/ou les conditions environnementales et la réponse observée -ou l'absence de réponse-. Troisièmement, une combinaison d'antibiotiques à des concentrations sous-inhibitrices peut refléter plus précisément la situation se déroulant dans l'environnement. Il a été démontré que l'association de plusieurs antibiotiques à des niveaux sous-inhibiteurs peut modifier le résultat de la

sélection de résistance *in vitro*[60]. Par conséquent, des études analysant l'effet des antibiotiques combinés (et des agents co-sélectifs tels que les métaux et les biocides) à des concentrations sous-inhibitrices sur le microbiome environnemental et le résistome sont fortement nécessaires pour fournir des informations plus réalistes pour l'évaluation des risques. Enfin, ces expériences se concentrent sur les effets de la pollution par la gentamicine sur le développement des résistances aux antibiotiques dans l'environnement, et elles n'ont pas été conçues pour analyser la mobilisation et la diffusion des résistances aux antibiotiques dans l'environnement et la probabilité de transfert vers les microbiomes humains et animaux. Tous ces éléments sont essentiels pour acquérir une compréhension plus approfondie du risque posé par la pollution de l'environnement par des antibiotiques et devraient être envisagés dans un futur.

Finalement, une étude de terrain visait à éclairer la question «l'activité anthropique a-t-elle un impact sur la composition du microbiome et du résistome dans les environnements froids?». Cette recherche montre que certains organismes réagissent aux changements environnementaux provoqués par la végétation environnante et l'activité anthropique dans la neige. Il s'agit de la première étude à démontrer que le tourisme d'hiver induit des changements dans le résistome de la neige sans avoir un fort impact sur les communautés bactériennes. Cela suggère que les activités anthropiques pourraient polluer la neige par l'apport de matière organique présente dans les déchets et stimuler le développement de résistances aux antibiotiques dans le microbiome de la neige qui pourraient être disséminées par la suite dans l'atmosphère ou par la fonte des neiges. Cela met en évidence la nécessité de faire des études sur le développement de la résistance aux antibiotiques dans des environnements pollués par les activités humaines et la prise en compte des sources organiques de pollution en plus des polluants biologiques (gènes de résistance aux antibiotiques et de bactéries résistantes aux antibiotiques). Cependant, l'augmentation de l'abondance des gènes de résistance aux antibiotiques spécifiques dans les sites pollués par des activités humaines devrait être validée par qPCR, et des études complémentaires devraient évaluer le potentiel de dissémination de ces gènes.

Bibliographie

1. Levy, S. B. & Bonnie, M. Antibacterial resistance worldwide: Causes, challenges and responses. *Nat. Med.* **10,** S122–S129 (2004).

2. Sommer, M. O. A., Munck, C., Toft-Kehler, R. V. & Andersson, D. I. Prediction of antibiotic resistance: Time for a new preclinical paradigm? *Nat. Rev. Microbiol.* **15,** 689–696 (2017).

3. Basssett, EJ, Keith MS, Armelagos GJ, Martin DL, V. A. Tetracycline-Labeled Human Bone from Ancient. *Science (80-.).* **209,** 1532–1534 (1980).

4. Dcosta, V. M. *et al.* Antibiotic resistance is ancient. *Nature* **477,** 457–461 (2011).

5. Klein, E. Y. *et al.* Global increase and geographic convergence in antibiotic consumption between 2000 and 2015. *Proc. Natl. Acad. Sci.* 201717295 (2018). doi:10.1073/pnas.1717295115

6. Lamrabet, O., Martin, M., Lenski, R. E. & Schneider, D. Changes in Intrinsic Antibiotic Susceptibility during a Long- Term Evolution Experiment with Escherichia coli. *MBio* **10,** (2019).

7. Baquero, F., Alvarez-Ortega, C. & Martinez, J. L. Ecology and evolution of antibiotic resistance. *Environ. Microbiol. Rep.* **1,** 469–476 (2009).

8. Blair, J. M. A., Webber, M. A., Baylay, A. J., Ogbolu, D. O. & Piddock, L. J. V. Molecular mechanisms of antibiotic resistance. *Nat. Rev. Microbiol.* **13,** 42–51 (2015).

9. Munita, J. M., Arias, C. A., Unit, A. R. & Santiago, A. De. Mechanisms of Antibiotic Resistance. *Microbiol. Spectr.* **4,** (2016).

10. Leski, T. A. *et al.* Multidrug-resistant tet(X)-containing hospital isolates in Sierra Leone. *Int. J. Antimicrob. Agents* **42,** 83–86 (2013).

11. Gilmore, M. S., Lebreton, F. & Schaik, W. van. Genomic transition of Enterococci from gut commensals to leading causes of multidrug-resistant hospital infection in the antibiotic era. *Curr. Opin. Biotechnol.* **23,** 1–7 (2013).

12. Poirel, L., Héritier, C. & Nordmann, P. Chromosome-Encoded Ambler Class D B-Lactamase of Shewanella oneidensis as a Progenitor of Carbapenem-Hydrolyzing Oxacillinase. *Society* **48,** 348–351 (2004).

13. San Millan, A. Evolution of Plasmid-Mediated Antibiotic Resistance in the Clinical Context. *Trends Microbiol.* **26,** 978–985 (2018).

14. Ma, X. *et al.* rpoB gene mutations and molecular characterization of rifampin-resistant Mycobacterium tuberculosis isolates from Shandong Province, China. *J. Clin. Microbiol.* **44,** 3409–3412 (2006).

15. Varela, A. R., Macedo, G. N., Nunes, O. C. & Manaia, C. M. Genetic characterization of fluoroquinolone resistant Escherichia coli from urban streams and municipal and hospital effluents. *FEMS Microbiol. Ecol.* **91,** 1–12 (2015).

16. González, I. *et al.* Fluoroquinolone resistance mutations in the parC, parE, and gyrA genes of clinical isolates of viridans group streptococci. *Antimicrob. Agents Chemother.* **42,** 2792–2798 (1998).

17. Pruden, A., Arabi, M. & Storteboom, H. N. Correlation between upstream human activities and riverine antibiotic resistance genes. *Environ. Sci. Technol.* **46,** 11541–11549 (2012).

18. Zhu, Y. G. *et al.* Continental-scale pollution of estuaries with antibiotic resistance genes. *Nat. Microbiol.* **2,** (2017).

19. Xiang, Q. *et al.* Spatial and temporal distribution of antibiotic resistomes in a peri-urban area is associated significantly with anthropogenic activities. *Environ. Pollut.* **235,** 525–533 (2018).

20. Peng, M. *et al.* Alterations of Salmonella Typhimurium Antibiotic Resistance under Environmental

Pressure. *Appl. Environ. Microbiol.* **84,** (2018).

21. Suzuki, S., Kimura, M., Agusa, T. & Rahman, H. M. Vanadium accelerates horizontal transfer of tet(M) gene from marine Photobacterium to Escherichia coli. *FEMS Microbiol. Lett.* **336,** 52–56 (2012).

22. Flaherty, E. O., Solimini, A. G., Pantanella, F., Giusti, M. De & Cummins, E. Human exposure to antibiotic resistant- Escherichia coli through irrigated lettuce. *Environ. Int.* 1–11 (2018). doi:10.1016/j.envint.2018.11.022

23. Ma, L., Li, B. & Zhang, T. New insights into antibiotic resistome in drinking water and management perspectives: A metagenomic based study of small-sized microbes. *Water Res.* **152,** 191–201 (2019).

24. Roesch, L. F. W. *et al.* Pyrosequencing enumerates and contrasts soil microbial diversity. *ISME J.* **1,** 283–290 (2007).

25. Rappé, M. S. & Giovannoni, S. J. The Uncultured Microbial Majority. *Annu. Rev. Microbiol.* **57,** 369–394 (2003).

26. Cruaud, P. *et al.* Influence of DNA extraction method, 16S rRNA targeted hypervariable regions, and sample origin on microbial diversity detected by 454 pyrosequencing in marine chemosynthetic ecosystems. *Appl. Environ. Microbiol.* **80,** 4626–4639 (2014).

27. Lombard, N., Prestat, E., van Elsas, J. D. & Simonet, P. Soil-specific limitations for access and analysis of soil microbial communities by metagenomics. *FEMS Microbiol. Ecol.* **78,** 31–49 (2011).

28. Robe, P., Nalin, R., Capellano, C., Vogel, T. M. & Simonet, P. Extraction of DNA from soil. *Eur. J. Soil Biol.* **39,** 183–190 (2003).

29. Nacke, H. *et al.* Estimates of Soil Bacterial Ribosome Content and Diversity Are Employed. *Appl. Environ. Microbiol.* **82,** 2595–2607 (2016).

30. Zaheer, R. *et al.* Impact of sequencing depth on the characterization of the microbiome and resistome. *Sci. Rep.* **8,** 1–11 (2018).

31. Smith, D. P. & Peay, K. G. Sequence depth, not PCR replication, improves ecological inference from next generation DNA sequencing. *PLoS One* **9,** (2014).

32. Lundin, D. *et al.* Which sequencing depth is sufficient to describe patterns in bacterial α- and β-diversity? *Environ. Microbiol. Rep.* **4,** 367–372 (2012).

33. Griffiths, R. I., Whiteley, A. S., O'Donnell, A. G. & Bailey, M. J. 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 (2000).

34. Wang, Q., Garrity, G. M., Tiedje, J. M. & Cole, J. R. Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* **73,** 5261–5267 (2007).

35. Huson, D. H., Auch, A. F., Qi, J. & Schuster, S. C. MEGAN analysis of metagenomic data. *Genome Res.* **17,** 377–386 (2007).

36. Alcock, B. P. *et al.* CARD 2020: Antibiotic resistome surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids Res.* **48,** D517–D525 (2020).

37. Nesme, J. & Simonet, P. The soil resistome: A critical review on antibiotic resistance origins, ecology and dissemination potential in telluric bacteria. *Environ. Microbiol.* **17,** 913–930 (2015).

38. Monier, J. M. *et al.* Metagenomic exploration of antibiotic resistance in soil. *Curr. Opin. Microbiol.* **14,** 229–235 (2011).

39. Naidoo, Y., Valverde, A., Cason, E. D., Pierneef, R. E. & Cowan, D. A. A clinically important, plasmid-borne antibiotic resistance gene (β-lactamase TEM-116) present in desert soils. *Sci. Total Environ.* **719,** 137497 (2020).

40. Perron, G. G. *et al.* Functional characterization of bacteria isolated from ancient arctic soil exposes

diverse resistance mechanisms to modern antibiotics. *PLoS One* **10,** 1–19 (2015).

41. Riesenfeld, C. S., Goodman, R. M. & Handelsman, J. Uncultured soil bacteria are a reservoir of new antibiotic resistance genes. *Environ. Microbiol.* **6,** 981–989 (2004).

42. Willms, I. M. *et al.* Discovery of Novel Antibiotic Resistance Determinants in Forest and Grassland Soil Metagenomes. *Front. Microbiol.* **10,** 460 (2019).

43. Luo, W., Xu, Z., Riber, L., Hansen, L. H. & Sørensen, S. J. Diverse gene functions in a soil mobilome. *Soil Biol. Biochem.* **101,** 175–183 (2016).

44. Jiang, X. *et al.* Dissemination of antibiotic resistance genes from antibiotic producers to pathogens. *Nat. Commun.* **8,** 1–7 (2017).

45. Hoelzer, K. *et al.* Antimicrobial drug use in food-producing animals and associated human health risks: What, and how strong, is the evidence? *BMC Vet. Res.* **13,** 1–38 (2017).

46. Jechalke, S., Heuer, H., Siemens, J., Amelung, W. & Smalla, K. Fate and effects of veterinary antibiotics in soil. *Trends Microbiol.* **22,** 536–545 (2014).

47. Blau, K. *et al.* Manure and Doxycycline Affect the Bacterial Community and Its Resistome in Lettuce Rhizosphere and Bulk Soil. *Front. Microbiol.* **10,** (2019).

48. Wolters, B. *et al.* Bulk soil and maize rhizosphere resistance genes, mobile genetic elements and microbial communities are differently impacted by organic and inorganic fertilization. *FEMS Microbiol. Ecol.* **94,** (2018).

49. Kang, Y. *et al.* High diversity and abundance of cultivable tetracycline-resistant bacteria in soil following pig manure application. *Sci. Rep.* **8,** 1489 (2018).

50. Jechalke, S. *et al.* Increased abundance and transferability of resistance genes after field application of manure from sulfadiazine-treated pigs. *Appl. Environ. Microbiol.* **79,** 1704–1711 (2013).

51. Heuer, H. *et al.* IncP-1ε plasmids are important vectors of antibiotic resistance genes in agricultural systems: Diversification driven by class 1 integron gene cassettes. *Front. Microbiol.* **3,** 1–8 (2012).

52. Heuer, H. *et al.* Accumulation of sulfonamide resistance genes in arable soils due to repeated application of manure containing sulfadiazine. *Appl. Environ. Microbiol.* **77,** 2527–2530 (2011).

53. Durso, L. M. & Cook, K. L. Impacts of antibiotic use in agriculture : what are the benefits and risks ? *Curr. Opin. Microbiol.* **19,** 37–44 (2014).

54. Heuer, H. *et al.* Gentamicin resistance genes in environmental bacteria: Prevalence and transfer. *FEMS Microbiol. Ecol.* **42,** 289–302 (2002).

55. Almakki, A., Jumas-Bilak, E., Marchandin, H. & Licznar-Fajardo, P. Antibiotic resistance in urban runoff. *Sci. Total Environ.* **667,** 64–76 (2019).

56. Andersson, D. I. & Hughes, D. Microbiological effects of sublethal levels of antibiotics. *Nat. Rev. Microbiol.* **12,** 465–478 (2014).

57. Sandegren, L. Selection of antibiotic resistance at very low antibiotic concentrations. *Ups. J. Med. Sci.* **119,** 103–107 (2014).

58. Chow, L., Waldron, L. & Gillings, M. R. Potential impacts of aquatic pollutants: Sub-clinical antibiotic concentrations induce genome changes and promote antibiotic resistance. *Front. Microbiol.* **6,** 1–10 (2015).

59. Bruchmann, J., Kirchen, S. & Schwartz, T. Sub-inhibitory concentrations of antibiotics and wastewater influencing biofilm formation and gene expression of multi-resistant Pseudomonas aeruginosa wastewater isolates. *Environ. Sci. Pollut. Res.* **20,** 3539–3549 (2013).

60. Gullberg, E., Albrecht, L. M., Karlsson, C., Sandegren, L. & Andersson, D. I. Selection of a Multidrug Resistance Plasmid by Sublethal Levels of Antibiotics and Heavy Metals. **5,** 19–23 (2014).

61. Choung, S. *et al.* Transfer of antibiotic resistance plasmids in pure and activated sludge cultures in the presence of environmentally representative micro-contaminant concentrations. *Sci. Total Environ.* **468–469,** 813–820 (2014).

62. Shun-Mei, E. *et al.* Sub-inhibitory concentrations of fluoroquinolones increase conjugation frequency. *Microb. Pathog.* **114,** 57–62 (2018).

63. Jutkina, J., Rutgersson, C., Flach, C. F. & Joakim Larsson, D. G. An assay for determining minimal concentrations of antibiotics that drive horizontal transfer of resistance. *Sci. Total Environ.* **548–549,** 131–138 (2016).

64. Jia, S. *et al.* Fate of antibiotic resistance genes and their associations with bacterial community in livestock breeding wastewater and its receiving river water. *Water Res.* **124,** 259–268 (2017).

65. George, J. & Halami, P. M. Sub-inhibitory concentrations of gentamicin triggers the expression of aac(6')Ie-aph(2'')Ia, chaperons and biofilm related genes in Lactobacillus plantarum MCC 3011. *Res. Microbiol.* **168,** 722–731 (2017).

66. Zhang, A. N. *et al.* Conserved phylogenetic distribution and limited antibiotic resistance of class 1 integrons revealed by assessing the bacterial genome and plasmid collection. *Microbiome* **6,** 1–14 (2018).

67. Gillings, M. R. Integrons: Past, Present, and Future. *Microbiol. Mol. Biol. Rev.* **78,** 257–277 (2014).

68. Körner, C. & Paulsen, J. A world-wide study of high altitude treeline temperatures. *J. Biogeogr.* **31,** 713–732 (2004).

69. Mountain Research Initiative EDW Working Group. Elevation-dependent warming in mountain regions of the world. *Nat. Clim. Chang.* **5,** 424–430 (2015).

70. UNEP 2007 Annual Report. (2007).

71. Cowan, D. A. *et al.* Non-indigenous microorganisms in the Antarctic: Assessing the risks. *Trends Microbiol.* **19,** 540–548 (2011).

72. Maccario, L., Carpenter, S. D., Deming, J. W., Vogel, T. M. & Larose, C. Sources and selection of snow-specific microbial communities in a Greenlandic sea ice snow cover. *Sci. Rep.* **9,** 1–14 (2019).

73. Shen, F. & Yao, M. Are We Biologically Safe with Snow Precipitation? A Case Study in Beijing. *PLoS One* **8,** (2013).

74. Zhu, G. *et al.* Air pollution could drive global dissemination of antibiotic resistance genes. *ISME J.* (2020). doi:10.1038/s41396-020-00780-2

75. Harding, T., Jungblut, A. D., Lovejoy, C. & Vincent, W. F. Microbes in high arctic snow and implications for the cold biosphere. *Appl. Environ. Microbiol.* **77,** 3234–3243 (2011).

76. Rogers, S. O., Starmer, W. T. & Castello, J. D. Recycling of pathogenic microbes through survival in ice. *Med. Hypotheses* **63,** 773–777 (2004).

77. Bergk Pinto, B., Maccario, L., Dommergue, A., Vogel, T. M. & Larose, C. Do Organic Substrates Drive Microbial Community Interactions in Arctic Snow? *Front. Microbiol.* **10,** 1–13 (2019).

78. Zhu, C. *et al.* Snow microbiome functional analyses reveal novel aspects of microbial metabolism of complex organic compounds. *Microbiologyopen* **9,** 1–12 (2020).

Chapter I. A review on the effects of anthropogenic environmental pollution on antibiotic resistance development and dissemination

Abstract

Since the discovery of antibiotics during the 20th century, antibiotic therapy has enabled the treatment of infectious diseases and drastically reduced the mortality caused by bacterial pathogens. However, during the last few decades, new cases of antibiotic resistance in clinics and the emergence of Multidrug Resistant (MDR) clones worldwide have increased concerns and healthcare stress that may result in antibiotic therapy failure and the beginning of a post-antibiotic era. Antibiotics and antibiotic resistance predate the discovery of antibiotics by hundreds of thousands of years. However, the industrialization and extensive use of antibiotics in humans and animals has imposed a selective pressure without precedents on medical and agricultural bacterial communities and accelerated the development of antibiotic resistance in these communities at a global scale. This selective pressure can induce the acquisition of resistance traits through vertical evolution -mutations on a single strain that transfers the resistant trait to its progeny- and by horizontal gene transfer (HGT). Socioeconomic changes that have taken place since the Industrial Revolution also favor the worldwide dissemination of antibiotic resistant bacteria (ARB) and antibiotic resistance genes (ARGs) through travel and global trade. However, the interconnection between human, animal and environmental microbiomes and the anthropogenic impact on non-human microbiomes may also affect the outcome of antibiotic resistance in human pathogens.

Antibiotics are used worldwide to treat diseases caused by human and animal pathogens and for therapeutic and growth promotion in aquacultures and crops worldwide. This use of antibiotics can select for resistances in the animal microbiome and then can be disseminated to the human microbiome through direct contact and food consumption (zoonotic diseases). Furthermore, wild animals and pets can also act as vectors of antibiotic resistance to the human microbiome. On the other hand, antibiotic production and consequently antibiotic resistance have existed for as long as bacteria themselves in the environmental microbiome and today both pristine and urban environments are reservoirs of antibiotic resistance. However, as a consequence of anthropogenic activities, residual concentrations of antibiotics, ARGs and ARBs reach the environment mainly through wastewater treatment plants, manure fertilization, and the release from antibiotic producing facilities. This results in the development and selection of antibiotic resistance in environmental ecosystems and the dissemination of ARGs and ARBs from the environment to animal and human microbiomes. Nonetheless, the scope of this phenomenon remains unclear. This review will identify scientific advances and knowledge gaps concerning the effects of biological and chemical pollutants on environmental resistance development and dissemination.

A brief introduction of antibiotics and antibiotic resistance

The discovery of antibiotics during the 20th century is one of the major breakthroughs of modern science. Since Fleming's discovery of penicillin in 1928[1], the isolation of new compounds with antimicrobial activity and the development of antibiotic therapy has enhanced the treatment of bacterial infections that have provoked pandemics with dramatic consequences during the Medieval Era[2] and were the main cause of death at the beginning of the 20th century[3]. The main classes of antibiotics used in human therapy target a fairly reduced variety of processes[4,5]. Penicillins, cephalosporins, carbapenems, monobactams and glycopeptides inhibit cell wall synthesis, whereas sulphonamides and trimetophim interfere in the folic acid metabolic pathway (Figure 1). Other classes of antibiotics target the synthesis and expression of genetic material: fluoroquinolones interfere with DNA replication by inhibiting the DNA gyrase; rifampin inhibits the DNA-dependent RNA-polymerase and consequently RNA synthesis; and several antibiotic classes inhibit the synthesis of proteins by the ribosome. Tetracyclines and aminoglycosides target the 30S subunit of the ribosome, whereas macrolides, chloramphenicol and clindamycin target the 50S subunit. In addition, oxazoliodonones bind a different target in the 50S subunit of the ribosome and inhibit the formation of the 70S initiation complex.

Since their discovery, these antibiotics have been widely used to treat primary and opportunistic infections caused by bacteria. However, over the last decades, a continuous increase in antibiotic resistance in clinics[5,6] has led to the failure of antibiotic therapy and the re-emergence of several infections that were considered to be under control[5,7]. The mechanisms conferring resistance to antibiotics have been thoroughly reviewed elsewhere[8–11]. The main resistance mechanisms found in pathogens can be classified in three groups (Figure 2)[8]: those that reduce the concentration of antibiotics in the bacterial cell by a decreased penetration or an increased efflux, those that modify the target preventing the binding of the antibiotic and those mediated by enzymes that inactivate the target by hydrolysis (*i. e.* beta-lactamases) or modification (*i. e.* methyltransferases). Generally, the reduction of intracellular antibiotic concentration confers a lower level of resistance than target modification and antibiotic inactivation[8]. However, several pathogens overexpressing efflux pumps have been found in clinics[12,13]. Several different resistance mechanisms can affect the performance of a single antibiotic: some antibiotic classes are influenced by more than one – or even all – of these mechanisms[10]. For example, resistance to beta-lactams can be mediated by hydrolysis, by target alteration and by an increased efflux.

Antibiotic production is far more ancient than their discovery by the scientists in the 20th century. Antibiotics are naturally produced by bacteria as a mean of competition with other microorganisms that share the same environmental niche[14]. Therefore, antibiotic resistance is probably as ancient as antibiotic production, since antibiotic-producing bacteria would themselves need to be resistant and would impose a selective pressure on their neighbors that force them to adapt to this selective pressure and develop or obtain antibiotic resistance that would enable their survival. Several studies demonstrate that both antibiotics and antibiotic resistance have been here long before Fleming's scientific breakthrough. A study carried out in 1980 provided evidence of exposure to tetracycline in a bone that predated the antibiotic era by at least 1400 years[15], whereas a more recent metagenomic analysis of 30000 year-old permafrost sediments identified genes conferring resistance to several antibiotics, such as β-lactams and tetracyclines[16]. The presence of antibiotic resistance genes has also been observed in the human microbiome of people living in a village in South America with no known exposure to commercial antibiotics[17]. Thus, the human microbiome has potentially been exposed to antibiotics since long before their use in clinics. Even though antibiotic resistance predates the discovery of antibiotics, resistance development in a closed niche, such as the Amerindian village,

would not imply a global threat *per se*. However, the use, overuse and misuse of antibiotics to treat human infections over the last century has applied an increasingly strong selective pressure on the human microbiome and increased the development rate of antibiotic resistance[18] and the dissemination of resistance genes in the human microbiome. Resistance has now been identified for all antibiotics classes used in human therapy during the 20th century[5,14,19]. The mechanisms underlying resistance development and dissemination need to be understood in order to determine to what extent selective pressure in the environment affects the rise of resistance in human and animal microbiomes and pathogens.

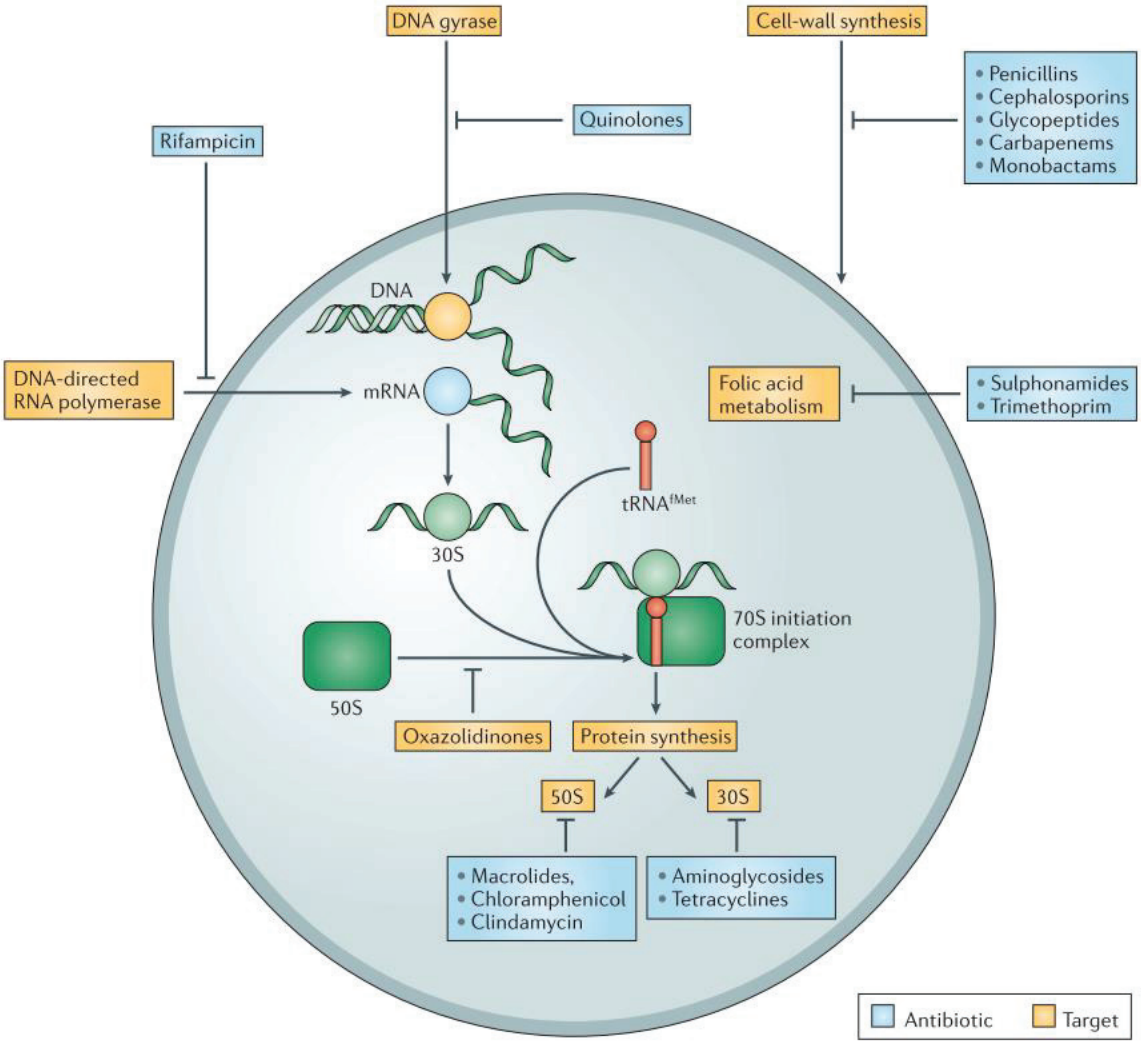


Figure 1. Targets of antibiotics. The main classes of antibiotics exploited by humans inhibit cell-wall synthesis, DNA synthesis, RNA synthesis, protein synthesis or folic acid metabolism. Figure from Lewis, 2013[4].

One of the main concerns regarding human health is the emergence of multidrug-resistant (MDR) bacterial clones[20,21]. These clones present highly-conserved genomes[22] that are consistent with their recent acquisition of their resistance mechanisms[19] and are often easily disseminated over large spatial distances even to other continents[23,24]. This ability to acquire and disseminate antibiotic resistance genes threatens antibiotic efficacy even for last-generation antibiotics. We are approaching the "post-antibiotic era" and unless alternative therapy options are developed or the spread of antibiotic resistances in human pathogens is reduced, antimicrobial therapy will no longer be effective. This would have dramatic consequences for human health. The inability to use antibiotics would affect the outcome of surgical procedures, transplantations or chemotherapy[25] in addition to the expected dramatic increase in mortality due to bacterial infections. Given the urgency of the situation, a global effort is needed in order to respond to the threat imposed by antibiotic resistance worldwide. Understanding the different factors that contribute to the rise of antibiotic resistance in human pathogens is critical.

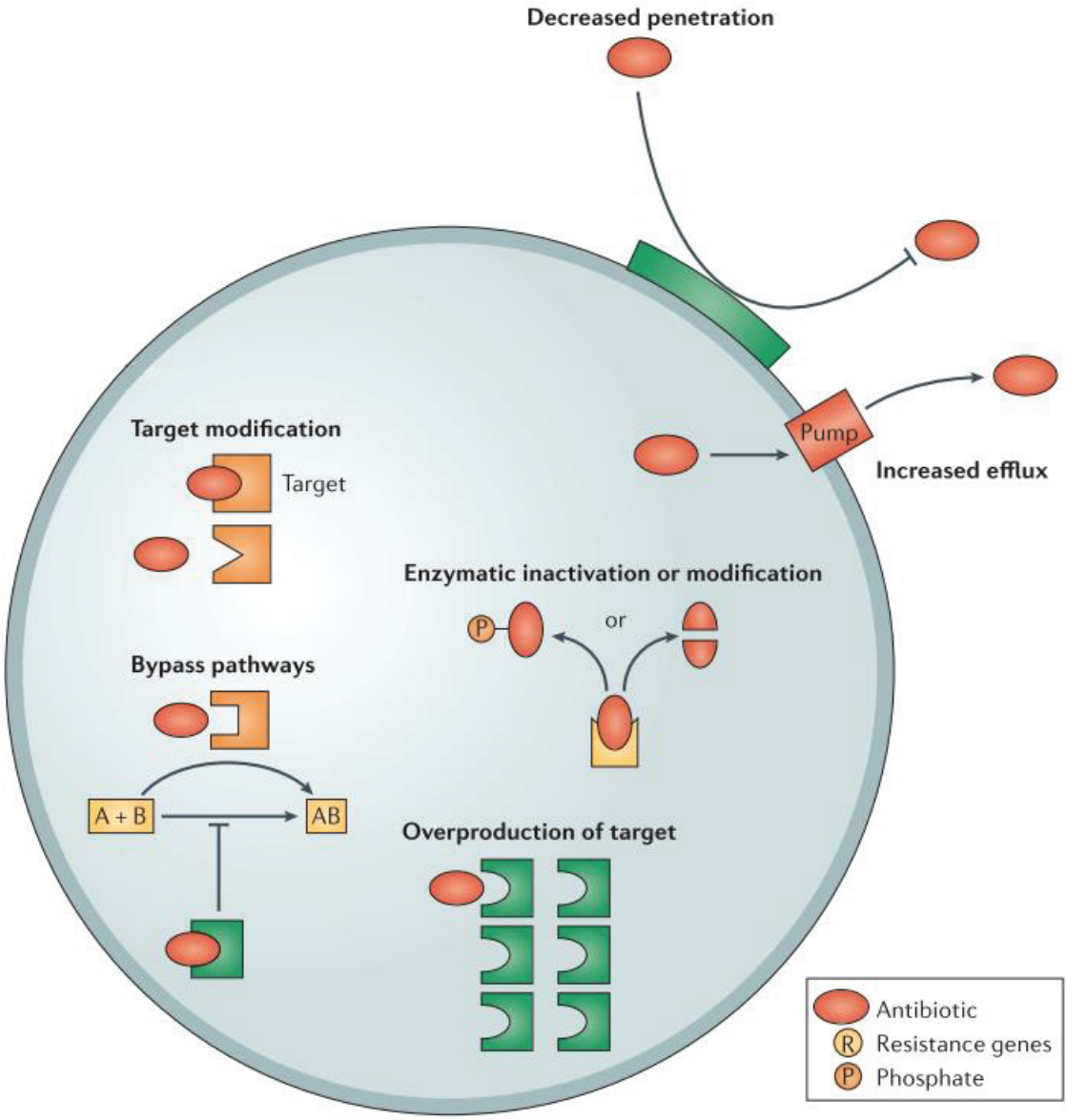


Figure 2. Mechanisms of antibiotic resistance. Resistance mechanisms include the reduction of intracellular antibiotic concentrations by a decreased penetration or an increased efflux, target modification and enzymatic inactivation of the the target by hydrolysis or modification. Figure from Lewis, 2013[4].

Antibiotic resistance dissemination: Horizontal Gene Transfer (HGT) and Mobile Genetic Elements (MGE)

Bacteria can be intrinsically resistant to an antibiotic[26] and/or acquire resistance through mutations of their genes and by horizontal gene transfer (Figure 3)[27]. Mutations in bacterial genes that affect antibiotic activity can happen during cell replication and be selected for and generate successful clones by vertical transfer[8,28]. The strain that carries this mutation can proliferate in the presence of antibiotics has a selective advantage. However, mutations can also generate a decreased fitness and are often lost in the absence of selective pressure[11]. On the other hand, antibiotic resistance can disseminate from bacteria containing a resistance gene to sensitive bacteria sharing the same niche by means of horizontal gene transfer (HGT)[6,8]. HGT events have shown to contribute to the development of the majority of antibiotic resistances currently found in pathogens[29–32]. However, the contribution of vertical evolution to the threat of antibiotic resistance must not be overlooked, since several mutant strains have resulted in pathogens of clinical relevance[33–36]. The three main mechanisms of HGT are conjugation, transduction and transformation (Figure 3).

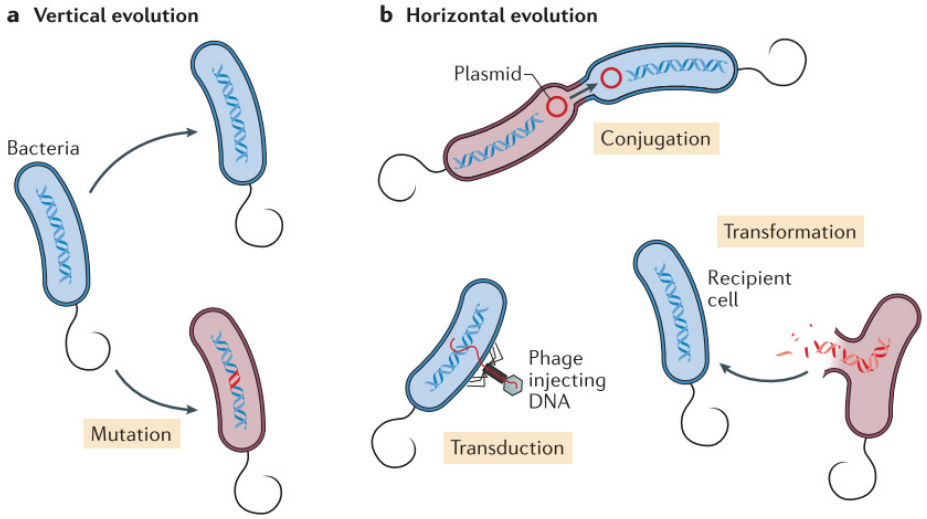


Figure 3. Transmission of antibiotic resistance. Blue cells: antibiotic sensitive bacteria. Red cells: antibiotic resistant bacteria. Resistance development can take place through two mechanisms: vertical evolution -the mutation of bacterial genes and their transfer to the progeny- or by horizontal gene transfer by conjugation, transduction or transformation. Figure from Sommer *et al.* 2017[6].

Conjugation involves direct contact between the donor bacteria and its recipient and the acquisition of new resistances through plasmids and integrative conjugative elements (ICE). Plasmids play a central role in the dissemination of ARGs through HGT[37]. Conjugative plasmids may display a narrow or broad host range, depending whether they are limited by a host phylogeny restriction[38]. Genes enabling plasmid replication and vertical transmission and genes encoding for conjugation proteins form the "backbone" core of housekeeping conjugative plasmid functions[39]. In addition, these plasmids contain genes that may benefit their host under specific environmental conditions. In the case of resistance plasmids, these genes include resistance genes as well as mobile genetic elements able to move within and between DNA molecules, such as integrons and transposons[40]. Transposons enable gene recombination, integration and excision from the bacterial host chromosome to the conjugative plasmid that mediates the horizontal transfer of the resistance gene to a recipient cell through conjugation[38].

Integrons mediate the capture and expression of exogenous genes. The intI integrase catalyzes the site-specific recombination of exogenous genes between the *attC* recombination site of the exogenous gene cassette and the *attI* recombination site of the integron (Figure 4). Several gene cassettes can be integrated into the same integron cassette array. This recombination is reversible. Gene cassettes can be excised as free circular DNA. The insertion of an exogenous gene at the *attI* site enables its expression by the integron promotor *Pc*[41]. Integrons are present in chromosomes, but they also are often associated to transposons and plasmids[42] to mediate their interchromosomal and intercellular mobility, respectively. Several integron classes have been defined based on the phylogeny of the *intI* gene. From these, class 1 integrons are often associated to ARGs, mobilized into transposons and plasmids and found in human pathogens[43]. They have also been proposed as markers for anthropogenic pollution[44].

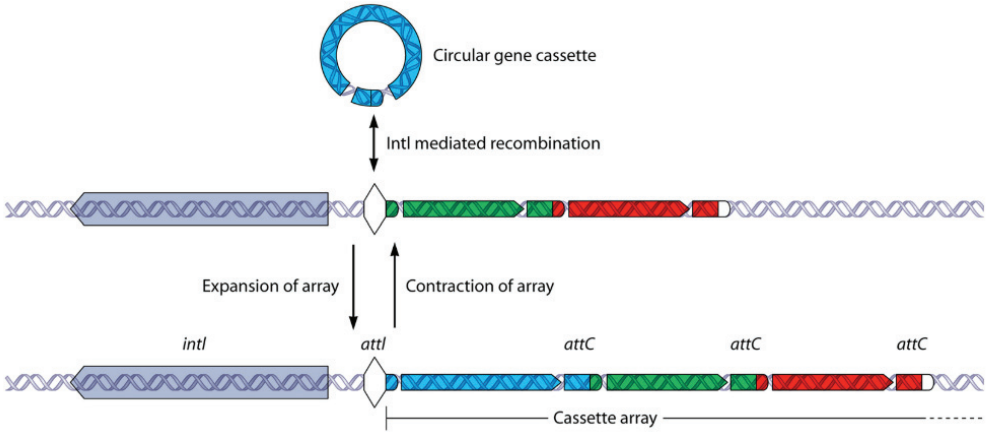


Figure 4. Recombination of exogenous gene cassettes. The intI integrase mediates the site-specific recombination of exogenous gene cassettes between *attI* and *attC*. Figure from Gillings, 2014[41].

Transposons are DNA segments that can move themselves and their associated resistance genes in a non-site specific manner intra or inter-molecularly within a bacterial cell[40]. In order to be transferred horizontally, these sequences need to be included in a plasmid. On the other hand, conjugative transposons, also known as integrative and conjugative elements (ICEs), encode a functional conjugation system that mediates their transfer between bacterial cells[45]. ICEs have been identified as the most abundant conjugative elements in prokaryotes[46]. They are normally integrated in the host chromosome as opposed to plasmids and are passively replicated within chromosomes. When the expression of the ICEs genes that mediate the excision and conjugation is induced, the ICE excises from the chromosome and can be transferred horizontally to a bacterial recipient[47]. Thus, ICEs can be disseminated both vertically and horizontally.

Conjugation is arguably the most efficient HGT mechanism[48] and poses a high risk for resistance dissemination, since multiple genes conferring resistance to antibiotics, metals or quaternary ammonium compounds can be located in the same gene cassette array as the gene selected under antibiotic pressure. Therefore, even in the absence of selective pressure for the rest of the genes in the plasmid or ICE, they can be co-selected with the gene that is primarily selected for[49–54]. Furthermore, although conjugation often seems to be limited to phylogenetically-close bacteria and plasmid-mediated conjugation often depends on the compatibility between the bacterial host and the plasmid[55], some conjugative elements can also be transferred among phylogenetically distant bacteria, and increase the scope of antibiotic resistance dissemination.

Transduction involves the action of bacteriophages, which are the most abundant organisms in the biosphere[56]. Bacteriophages infect bacteria by attaching to specific receptors on their surface and injecting their DNA (or RNA) in the bacterial genome. Virulent phages perform lytic cycles. They induce the formation of phage particles and lyse the bacterial cell to release the newly assembled phages instead of integrating their DNA into the host bacteria (Figure 4). On the other hand, the inserted DNA of temperate phages, also called prophages, can be integrated in the host chromosome and/or inserted as a free plasmid (Figure 5), and transmitted to the bacterial progeny (lysogenic cycle[57,58]). The injection of genetic material through bacteriophages has a broader spectrum than the lysis of bacterial cells[59]. Thus, ARGs in bacterial chromosomes and plasmids may be mobilized by phages during the infection cycle and disseminated between different bacterial species sharing the same niche. This gene mobilization can be carried out through generalized transduction, the mobilization of any fragment of the bacterial genome and through specialized transduction or mobilization of a few specific genes. In both cases, the survival of the inserted bacterial genes depends on the integration of these genes into the recipient bacterial genome: the presence of genes encoding for recombinases, integrases or transposases is critical for the acquisition of ARG in the recipient host[60]. Specialized transducing phages include an integrase gene that increases the probability of successful integration.

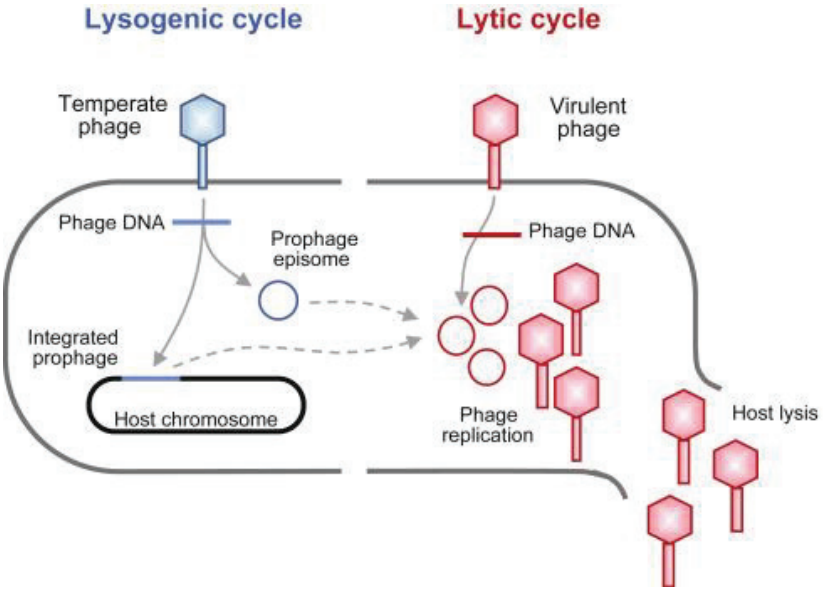


Figure 5. Infection and injection of DNA carried by phages into a bacterial host. After phage DNA insertion in the bacterial host, either the lysogenic cycle (blue) or the lytic cycle (red) can be induced. Figure from Fortier, 2017[61].

Finally, transformation takes place after the death and lysis of the donor bacterium, which releases its DNA for its uptake by a recipient cell. Then, the exogenous DNA segment internalized in the recipient cell is integrated into its genome by homologous (or heterologous) recombination. Thus, the newly acquired genes that provide the recipient bacterium with resistance to environmental stressors such as antibiotics can be transferred to their progeny[62]. Transformation is the only HGT mechanism that is controlled by the recipient cell, which contains all the genetic mechanisms involved[63]. However, this mechanism has some ecological bottlenecks that limit its occurrence. In order to undergo transformation, bacterial cells need to be competent[64]. Although competence can beinduced chemically or physically in certain environments, not all bacterial taxa are naturally competent. In addition, the proteins involved in competence of transformable bacteria are not expressed systematically as their expression relies on specific conditions such as growth alterations[63]. Thus, transformable bacteria are competent for a limited period of time in which transformation may take place.

Human, animal and environmental microbiomes and resistomes are interconnected

Along with the anthropogenic increase of the selective pressure imposed on bacteria, human society itself has experienced a drastic change over the last century. Now, international trade and tourism are critical to the global economy. This increased contact between individuals and environments from distant places also affects the resistance profile of the human microbiome and stimulates resistance dissemination around the planet[65–67]. In addition, the interconnection between the human, animal and environmental microbiomes and the anthropogenic impact on non-human microbiomes may also affect the outcome of antibiotic resistance in human pathogens.

Human and animal microbiomes are interconnected and infectious diseases caused by microorganisms present in the animal microbiome can disseminate to humans through direct contact or food consumption. The most recent and notorious example of this is the SARS-CoV-2 virus, which was originally hosted by bats and quickly disseminated to the human microbiome worldwide[68]. In addition, several food-born outbreaks caused by bacteria have been detected worldwide[69,70]. Approximately 60% of emerging diseases have a zoonotic origin[71]. The use of antibiotics on animals has been widely expanded over the last decades in order to prevent or treat infectious diseases as well as for growth promotion purposes[72]. Both uses were mainly aimed at increasing production in order to satisfy the demand for products of animal origin. These antibiotics can select for ARBs and ARGs and build a reservoir of antibiotic resistance in farm animal microbiomes[73]. Antibiotic resistance can then disseminate to the human microbiome either by food consumption and/or by direct contact, especially to high-risk groups such as farm workers[74]. Several studies link animal and the human microbiomes and show evidence of a shared resistome between humans and farm animals[75,76]. This kind of scenario also applies to crops. Produce that harbors a wide range of ARB and ARG[77,78] and is often eaten raw has been identified as the cause of several infectious outbreaks worldwide[79–81]. Due to the rising concern about the triggering effect of antibiotic use in agriculture and aquaculture on resistance development and dissemination, its use has been regulated in several countries[82,83]. However, resistances can disseminate not only through food consumption, but also through pets and wild animals, such as birds, insects, primates. Small mammals can act as vectors of ARB and ARG and contribute to their dissemination around the world[84–89]. Therefore, the animal resistome contributes to resistance development in human microbiomes and consequently in human pathogens.

The selective pressure imposed by antibiotics in nature has been selecting for resistance long before the human species started using them for their own purposes. Most of the antibiotics that were used in human therapy during the 20th century (some of which are still being used) have been isolated from soil bacteria[4]. Both clinically important and novel ARGs have been identified in soil subjected to relatively low anthropogenic pressure[90,91]. From a wider point of view, the environment can be considered as a reservoir of antibiotic resistant bacteria (ARB) and antibiotic resistance genes (ARGs)[92,93]. The ensemble of ARGs present in the environment is known as the environmental resistome[94]. ARGs have been identified even in pristine environments such as a cave isolated for over 4 million years[95] and lakes subjected to low anthropogenic impact[96]. Several urban settings, such as public restrooms[97], metro systems[98], surface waters[99], stormwater loadings[100] and air[101] have also been shown to harbor ARB and ARGs. Thus, both urban and pristine environments can be considered as reservoirs of ARB and ARGs and these resistance determinants can spread under specific conditions from the environmental resistome to commensal microorganisms and pathogens in the human microbiome[102,103]. Several resistance mechanisms found in clinics originated from environmental bacteria[104,105].

Despite that antibiotic production and antibiotic resistance occur naturally in the environment, the use of antibiotics in humans therapy and food production generates a flow of residual concentrations of antibiotics, as well as ARB and ARGs[106,107], to the environment (Figure 6). These contaminants may reach the environment through direct contact with humans, through hospital effluents, wastewater treatment plants, and through direct contact with contaminated crops and animal urine and feces or through manure applications for soil fertilization. Antibiotics from production plants may also be unintentionally released to the environment. The effects of both chemical and biological environmental pollutants on resistance development and dissemination are discussed below. Generally, biological pollutants (ARB and ARGs) may integrate into the environmental resistome from where they can disseminate back to the human microbiome. In addition, the selective pressure imposed by residual concentrations of antibiotics on human, animal and environmental microbiomes may enhance the development of antibiotic resistance[108,109], the horizontal transfer of ARGs to susceptible bacteria[110], and the dissemination of ARB between the three compartments. Environmental ARB and ARGs might disseminate to animal microbiomes by direct contact and to the human microbiome by direct contact, through drinking water[111,112], vegetables and animal vectors. Therefore, the anthropogenic input of antibiotics in the environment could result in the development of new resistances in clinics.

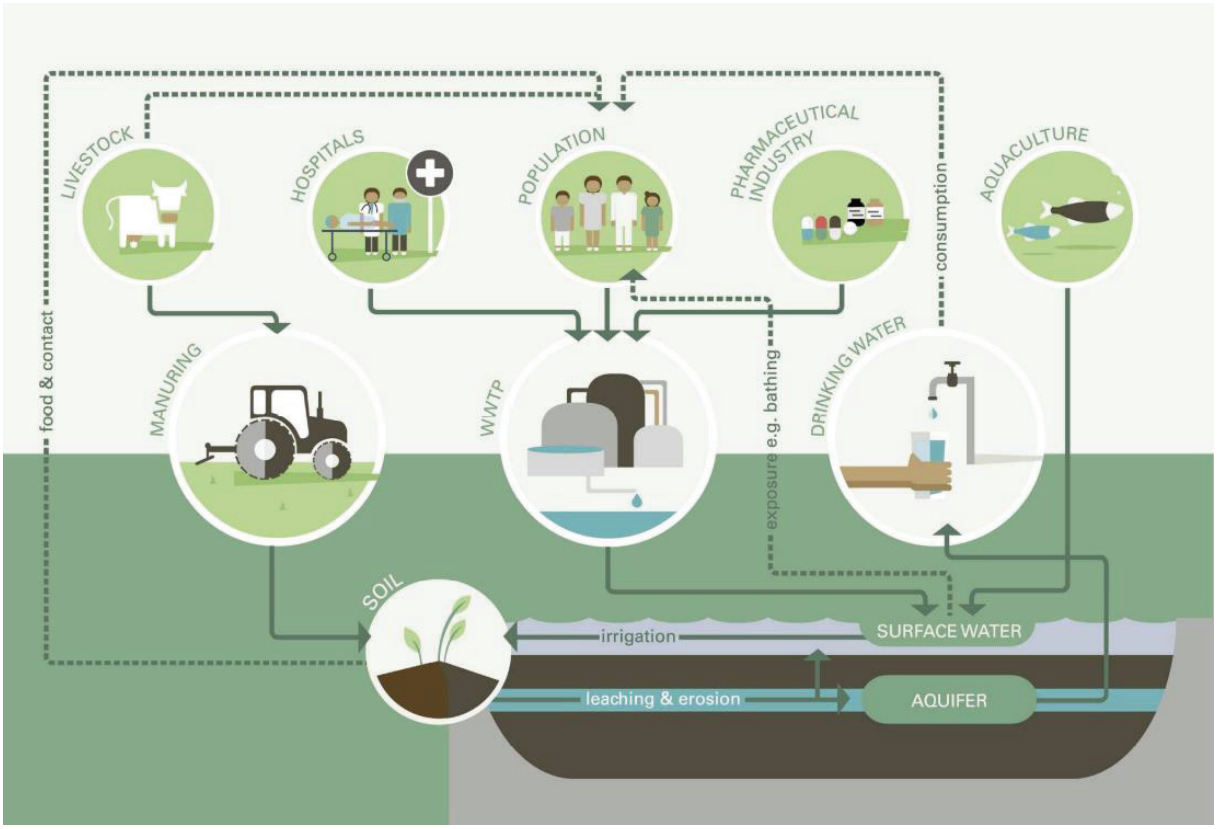


Figure 6. Routes of dissemination of antibiotics, ARB and ARG between humans, animals and the environment. Residual concentrations of antibiotics, antibiotic resistant bacteria (ARB) and antibiotic resistance genes (ARG) reach the environment mainly through hospital effluents, wastewater treatment plants, manure fertilization, aquacultures or release from antibiotic producing plants. This may result in the development and selection of antibiotic resistance in environmental settings and the dissemination of ARB and ARG between different environmental compartments and from the environment to animal and human microbiomes, either through direct contact with contaminated animals or environments or through food consumption or drinking water. Figure from Vikesland *et al.* 2017[113].

However, the proportion of ARGs found in pathogens is very low compared to those present in the environment[114,115]. This suggests that several bottlenecks regulate the dissemination of ARGs from environmental bacteria to human pathogens. These bottlenecks have been reviewed elsewhere[116]. Briefly, human pathogens need to share a niche with environmental bacteria in order to acquire resistance mechanisms from them by horizontal gene transfer, whose rate is in turn influenced by the phylogenic relation between the donor and the recipient[117]. In addition, the fitness cost of resistance acquisition makes it unlikely for this dissemination to take place without a selective pressure (*i.e*. environmental pollution with antibiotics, metals or biocides) and for the newly acquired ARGs to be maintained in the genome of the human pathogen in the absence of selective pressure[118]. Only those genes that present affordable fitness costs will successfully spread among bacteria. Thus, the scope of resistance selection in the environment and the dissemination of resistances to the human microbiome has not yet been elucidated[119].

Humans, livestock and their products, wild animals and pets, plants and the environment form a global resistome with shared elements between environments (Figure 7). As a consequence of the interconnection between human, animal and environmental microbiomes and the dissemination of antibiotic resistance between these environmental compartments[120–123], the "One Health" approach has emerged to tackle the antibiotic resistance black box with a cross-disciplinary approach[124]. This approach aims to incorporate resistance surveillance on agricultural and environmental settings, especially those affected by anthropogenic activities instead of focusing only on clinical scenarios.

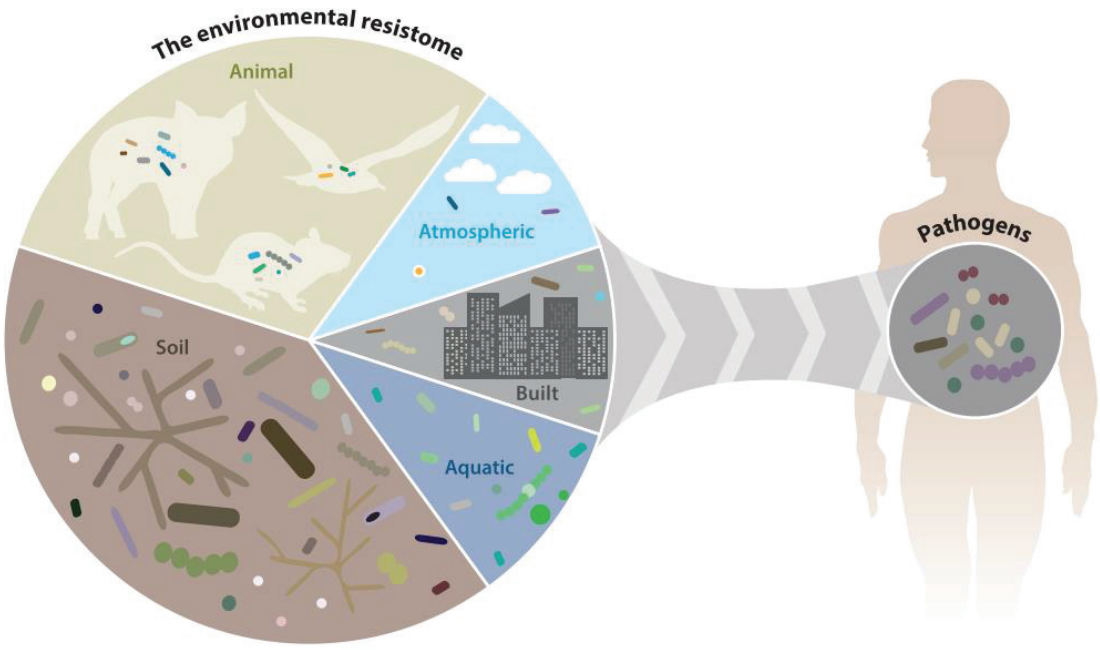


Figure 7. The environmental and animal microbiomes are connected to human pathogens. Antibiotic resistance in clinics mirrors and may originate from environmental compartments. Soil, aquatic, atmospheric, and built environments, as well as wild, farm and domestic animals, can act as reservoirs of antibiotic resistance. Furthermore, anthropogenic activities promote the selection and mobilization of resistance in these environmental compartments that can disseminate to human pathogens causing clinical infections. Figure from Surette *et al.* 2017[89].

Consequences of anthropogenic activity on resistance development and dissemination in the environment

Anthropogenic activities can pollute the surrounding environment with residual concentrations of antibiotics used in human therapy, food production, and with biological pollutants (ARB and ARGs). In the following section, the effects of both chemical and biological pollution on the environmental microbiome and their potential consequences on human health will be discussed.

Environmental pollution with biological pollutants

Waste from humans and animals that have been exposed to antibiotics can contain antibiotic resistant bacteria (ARB) and antibiotic resistance genes (ARGs). These may reach the environment through wastewater treatment plants[125–128] and manure[129,130] and they are considered "hotspots" of ARB and ARGs, which can survive in the environment even in the absence of selective pressure[131]. Effluents from hospital and urban wastewater treatment plants reach lakes, rivers and coastal waters and they are often reused for non-potable applications[132,133] such as soil irrigation. In addition, water sanitation in developing countries is often scarce and the environment is often polluted with untreated wastewater[134,135]. On the other hand, the application of contaminated manure and fecal pollution from aquacultures may affect the environmental microbiome and resistome[136,137]. The discharge of ARB and ARGs into the environment may increase environmental resistance[138–143]. These biological pollutants can disseminate back to the human and animal microbiomes[144,145] through food ingestion, drinking water and by direct contact while swimming in contaminated water[146] and contact with contaminated crops, and thus create a loop between the human, animal and environmental microbiomes that may compromise the efficacy of antibiotic therapy.

Yet, the scope of environmental contamination with biological pollutants remains unclear. For example, whereas several studies point to treated sewage as a source of antibiotic resistance genes in the receiving environment[147,148] even when trace levels of treated sewage are added to a receiving environment[149], another study suggests a negligible effect of treated fecal pollution on the resistome of the receiving environment[150]. Furthermore, geographical differences are reported between environmental resistomes and associated with different socio-economic factors. Developing countries with poorer sanitation systems presented the most divergent distribution of ARGs and were identified as hotspots for antibiotic resistance emergence in a recent study[151]. Further research is needed to assess the scope of pollution and the persistence of the biological pollutants in the receiving environments, to determine to what extent the exposure to environments polluted with fecal ARB and ARGs is associated to the dissemination of antibiotic resistance to the human microbiome, and to provide quantitative data concerning routes of resistance dissemination from the environment to the human microbiome[152].

Environmental pollution with chemical pollutants

Residual concentrations of antibiotics used in human therapy and animal production may reach the environment through hospital, urban and agricultural wastewater effluents[153], through untreated sewage[154], animal feces and manure application[155,156] and through intentional and accidental release from antibiotic production plants[157]. These residual concentrations of antibiotics may impose a selective pressure on environmental bacteria and result in the selection[158–160], development[91], mobilization[161] and dissemination of antibiotic resistance in the environment[162]. Furthermore, other pollutants such as metals[163] and organic compounds[164] have been shown to co-select for antibiotic resistance in the environment. ARB and ARGs selected in polluted environments may disseminate to the human microbiome in the case of direct exposure and indirectly through animal vectors[117].

However, the scope of this phenomenon and the consequences of environmental antibiotic pollution on human health are not clear. Further research is needed to elucidate dose-response relationships in the environment and to determine the mechanisms of resistance development and the rates of dissemination of resistance acquired under selective pressure in the environment[165].

The hypothesis of the selective window states that resistance selection only occurs at concentrations between the minimal inhibitory concentration (MIC) of susceptible and resistant bacteria[166]. The MIC is defined as the minimal concentration of antibiotics that inhibits visible growth in culture[167]. Thus, given that antibiotic concentrations found in natural and anthropogenically-polluted environments are often sub-inhibitory[168] (*i.e.* too low to cause a visible growth inhibition of susceptible bacteria in culture), their role on resistance development has been long overlooked. Nevertheless, sub-inhibitory concentrations may cause a perturbation of the microbial community without inhibiting overall growth. Sub-inhibitory concentrations are often sufficient to slow down bacterial growth[169] without causing inhibition (Figure 8). Therefore, although sub-inhibitory, non-lethal concentrations of antibiotics would not have an overall influence on bacterial community composition, the selective pressure they exert may be sufficient to offset the cost of resistance, thereby, contribute to resistance selection[170].

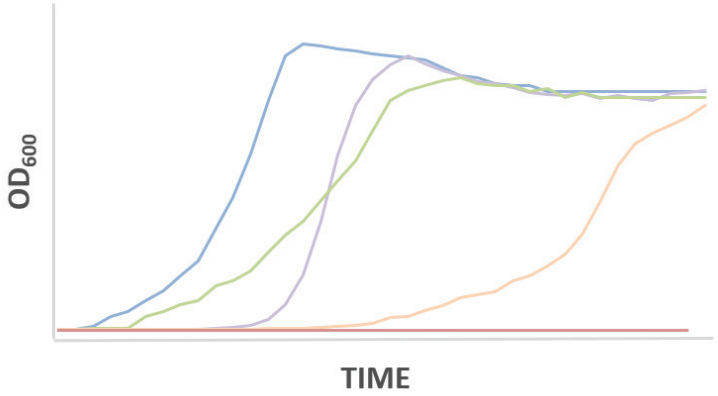


Figure 8. Effects of sub-inhibitory concentrations of antibiotics on bacterial growth in culture. Sub-inhibitory concentrations of antibiotics may slow down bacterial growth in culture without causing growth inhibition. These effects may be observed as a delay on the offset of the exponential phase (purple) compared to bacteria growing in media without antibiotics (blue), as a reduction of the slope of the curve (green) or both (orange). Nevertheless, all curves at sub-inhibitory levels of antibiotic pollution reach the same plateau as the one without antibiotics (*i.e.,* sub-inhibitory concentrations of antibiotics do not inhibit visible growth). On the other hand, inhibitory concentrations of antibiotics would visibly inhibit growth in culture (red).

Since environmental bacteria often produce antibiotics at sub-inhibitory concentrations, they may act as signaling molecules in natural environments[171] and induce a wide range of responses in bacterial communities. The responses induced by sub-inhibitory concentrations of antibiotics and their potential for resistance selection have been described elsewhere[167,172]. Briefly, an increasing body of evidence shows that sub-inhibitory concentrations of antibiotics may induce quorum sensing, persistence, biofilm formation and the expression of genes involved in antibiotic resistance, virulence and the SOS response[173]. The induction of the SOS response has been related to increased mutation rates, an activation of recombinases and the enhancement of horizontal gene transfer (HGT) by the induction of prophages, integrative conjugative elements and transposons[174,175] that in turn would increase the diversity of ARB and ARGs[176]. Furthermore, sub-inhibitory concentrations of antibiotics may also select for ARB[169] as illustrated by the strength of selection for $bla_{\text{CTX-M}}$ genes at sub-inhibitory concentrations appearing similar to that occurring at inhibitory concentrations[177]. Therefore, the idea of resistance selection being proportional to antibiotic concentration is in jeopardy and at least some antibiotics

select for resistance in a dose-independent manner and concentrations below the MIC of susceptible bacteria may contribute to resistance development as much as concentrations between in the selective window[170].

These studies have used culture-based approaches to characterize the effects of sub-inhibitory concentrations of antibiotics on individual strains and the cultivable fraction of complex communities. So while they provide critical information about the ecological roles of sub-inhibitory concentrations of antibiotics, these studies examined a small and specific portion of the bacterial community under established *in vitro* conditions and do not necessarily account for the complexity of natural environments[170], where bacterial growth and interactions are influenced by several factors. For instance, some PCR and qPCR based microcosm studies have observed a minor effect of spiked-in low concentrations of tetracycline, doxycycline and streptomycin applied with manure on ARG diversity and abundance in soil[178–180]. Nevertheless, given that antibiotics can be adsorbed by soil at different proportions depending on antibiotic structure and soil composition, each antibiotic-soil interaction will probably have different effects on the resistome and the probability for resistance selection in antibiotic-polluted soils cannot be underestimated. These studies targeted a limited set of genes and do not necessarily provide a complete view of the impact of pollution with antibiotics at sub-inhibitory concentrations on resistance development in soil. Therefore, studies combining metagenomics resistome screening and quantitative measurements on the abundance of a selected set of genes may provide a clearer view on the resistome response to antibiotic pollution at sub-inhibitory concentrations.

Given the complexity of environmental biotic and abiotic factors, determining the scope of resistance development and the risks associated to environmental antibiotic pollution at sub-inhibitory concentrations remains a challenging task. In addition, the effects of sub-inhibitory concentrations on complex communities may vary between the members of the community. While sub-inhibitory concentrations of antibiotics might not have an overall lethal effect, this does not necessarily imply that none of the members of a complex, environmental community will be sensitive to that antibiotic, and therefore, selected against at "sub-inhibitory" concentrations. Therefore, fundamental research is needed to determine whether sub-inhibitory antibiotic concentrations in natural settings can promote resistance development and selection[167,181] and to establish a quantitative relationship between antibiotic concentration and selective force in the environment[182] in order to identify threshold antibiotic concentrations that can impact the environmental microbiome and pose a potential risk for human health.

PhD objectives, approach and working hypothesis

Objective

The goal of this PhD was to evaluate the effects of anthropogenic environmental contamination on antibiotic resistance development and dissemination in the environment. Two types of contamination were analyzed separately. Firstly, the effects of chemical polluted with antibiotics were evaluated in soil and water microcosms. Secondly, both pristine and anthropogenically-impacted snow samples were analyzed to determine whether anthropogenic activities caused an increase in biological pollutants in snow.

Effects of chemical environmental pollution on resistance development and dissemination

Hypothesis

Gentamicin would be highly adsorbed onto soil particles and its bioavailability would be reduced shortly after administration. This added to the low activity of soil bacteria and the lower probability of contact in a solid matrix would result in a lack of visible impact of gentamicin even at high concentrations on the soil microbiome and resistome composition. If gentamicin had a small effect on any individual bacterial genome, it would be non-detectable due to the high diversity of the soil microbiome and the low abundance of most of individual members.

The increased contact between water bacteria (compared to soil bacteria) and the higher availability of gentamicin in water would result in a more propitious environmental setting to promote resistance development and dissemination due to sub-inhibitory concentrations of gentamicin. Thus, sub-inhibitory concentrations of gentamicin would induce shifts in the bacterial community and would increase the abundance and expression of ARGs without inhibiting overall community growth.

Approach

The effects of antibiotic environmental pollution on resistance development and dissemination in environmental settings were evaluated using a microcosm approach. Firstly, new techniques for DNA and RNA extraction from soil were compared to existent alternatives in order to improve access to the soil microbial diversity. Then, a model antibiotic, gentamicin, was chosen to evaluate its effects on microbial community, resistome composition and gene expression in the environment. Gentamicin is an aminoglycoside used in both human therapy and food production for which enzymatic inactivation is the most common resistance mechanism[183,184]. Gentamicin resistance genes are abundant and diverse in several environmental settings, such as manure, animal feces or seawater[185], and sub-inhibitory concentrations of gentamicin have been shown to stimulate resistance selection and biofilm formation *in vitro*[186] and resistance dissemination[187]. Here, the concentrations that were inhibitory and sub-inhibitory in soil and water bacteria enriched in culture media were determined by evaluating bacterial growth. Agricultural soil microcosms were polluted with a range of inhibitory concentrations of gentamicin from 1 µg/g to 1 mg/g of soil and culture-dependent and metagenomic approaches were combined to analyze the difference in the response between the communities in enriched media and environmental settings at several exposure times. The bioavailable fraction of gentamicin in soil was measured by HPLC-MS/MS. In addition, river water microcosms were polluted with both inhibitory and sub-inhibitory concentrations of gentamicin, and the response to soil microbial communities to gentamicin pollution was evaluated, in order to determine whether sub-inhibitory concentrations of gentamicin affected the river water microbiome and resistome. Class 1 integrons were sequenced and screened for aminoglycoside resistance genes, and dose-response relationships were established between gentamicin concentrations measured in water by HPLC-MS/MS and the abundance and expression of aminoglycoside resistance genes measured by qPCR/RT-qPCR.

Effects of anthropogenic activities on the snow microbiome and resistome

Hypothesis

Both environmental and anthropogenic factors influence snow ecology and induce changes in the snow microbiome and antibiotic resistome by providing bacterial communities with higher levels of organic carbon and other nutrients. This would support the growth of a more abundant bacterial community that would increase the abundance of the antibiotic resistome. Previous research performed by our group has shown that increased levels of organic acids induced competition and increased the abundance of ARGs in snow[188]. These increased concentrations also supported the establishment of a more metabolically active community[189]. Therefore, both anthropogenic and environmental sources of organic carbon could stimulate competition and ARG proliferation in snow.

Approach

Snow samples were obtained from two watersheds located in the Karkonosze National Park in the Sudety Mountains in Poland. One of them was subjected to human activities such as tourism and cottage development and the other one was hypothetically unaffected by local anthropogenic activity. In addition, snow with different levels of surrounding vegetation were sampled from both catchments. DNA was extracted from snow samples and the impact of vegetation and anthropogenic activity on the snow microbiome and resistome was evaluated using a metagenomics and qPCR approach. Bacterial community size, composition, genus richness and diversity were the parameters chosen to determine the effects of anthropogenic and environmental factors on the snow microbiome, whereas their effects on the snow resistome were assessed by the determination of the size and composition of the antibiotic resistome, as well as ARG richness and diversity.

References

1. Gould, K. Antibiotics: From prehistory to the present day. *J. Antimicrob. Chemother.* **71,** 572–575 (2016).

2. Smith, P. W., Watkins, K. & Hewlett, A. Infection control through the ages. *Am. J. Infect. Control* **40,** 35–42 (2012).

3. da Cunha, B. R., Fonseca, L. P. & Calado, C. R. C. Antibiotic discovery: Where have we come from, where do we go? *Antibiotics* **8,** (2019).

4. Lewis, K. Platforms for antibiotic discovery. *Nat. Rev. Drug Discov.* **12,** 371–387 (2013).

5. Levy, S. B. & Bonnie, M. Antibacterial resistance worldwide: Causes, challenges and responses. *Nat. Med.* **10,** S122–S129 (2004).

6. Sommer, M. O. A., Munck, C., Toft-Kehler, R. V. & Andersson, D. I. Prediction of antibiotic resistance: Time for a new preclinical paradigm? *Nat. Rev. Microbiol.* **15,** 689–696 (2017).

7. Porter, J. D. H. & Mcadam, K. P. W. J. The re-emergence of tuberculosis. *Annu. Rev. Public Health* **15,** 303–323 (1994).

8. Blair, J. M. A., Webber, M. A., Baylay, A. J., Ogbolu, D. O. & Piddock, L. J. V. Molecular mechanisms of antibiotic resistance. *Nat. Rev. Microbiol.* **13,** 42–51 (2015).

9. Du, D. *et al.* Multidrug efflux pumps: structure, function and regulation. *Nat. Rev. Microbiol.* 1–17 (2018). doi:10.1038/s41579-018-0048-6

10. Morar, M. & Wright, G. D. The Genomic Enzymology of Antibiotic Resistance. *Annu. Rev. Genet.* **44,** 25–51 (2010).

11. Munita, J. M., Arias, C. A., Unit, A. R. & Santiago, A. De. Mechanisms of Antibiotic Resistance. *Microbiol. Spectr.* **4,** (2016).

12. Pumbwe, L. & Piddock, L. J. V. Two Efflux Systems Expressed Simultaneously in Multidrug-Resistant Pseudomonas aeruginosa. *Antimicrob. Agents Chemother.* **44,** 2861–2864 (2000).

13. Kosmidis, C. *et al.* Expression of multidrug resistance efflux pump genes in clinical and environmental isolates of Staphylococcus aureus. *Int. J. Antimicrob. Agents* **40,** 204–209 (2012).

14. Waglechner, N. & Wright, G. D. Antibiotic resistance: It's bad, but why isn't it worse? *BMC Biol.* **15,** 1–8 (2017).

15. Basssett, EJ, Keith MS, Armelagos GJ, Martin DL, V. A. Tetracycline-Labeled Human Bone from Ancient. *Science (80-.).* **209,** 1532–1534 (1980).

16. Dcosta, V. M. *et al.* Antibiotic resistance is ancient. *Nature* **477,** 457–461 (2011).

17. Clemente, J. C. *et al.* The microbiome of uncontacted Amerindians. *Sci. Adv.* (2015).

18. Klein, E. Y. *et al.* Global increase and geographic convergence in antibiotic consumption between 2000 and 2015. *Proc. Natl. Acad. Sci.* 201717295 (2018). doi:10.1073/pnas.1717295115

19. Klemm, E. J., Wong, V. K. & Dougan, G. Emergence of dominant multidrug-resistant bacterial clades : Lessons from history and whole-genome sequencing. *PNAS* **115,** 12872–12877 (2018).

20. Molina-Santiago, C., de Vicente, A. & Romero, D. The race for antimicrobials in the multidrug resistance era. *Microb. Biotechnol.* 2016–2018 (2017). doi:10.1111/1751-7915.12884

21. Peleg, A. Y., Seifert, H. & Paterson, D. L. Acinetobacter baumannii: Emergence of a successful pathogen. *Clin. Microbiol. Rev.* **21,** 538–582 (2008).

22. Kingsley, R. A. *et al.* Epidemic multiple drug resistant Salmonella Typhimurium causing invasive disease in sub-Saharan Africa have a distinct genotype. *Genome Res.* **19,** 2279–2287 (2009).

23. Nicolas-Chanoine, M. H. *et al.* Intercontinental emergence of Escherichia coli clone O25:H4-ST131 producing CTX-M-15. *J. Antimicrob. Chemother.* **61,** 273–281 (2008).

24. Wong, V. K. *et al.* Phylogeographical analysis of the dominant multidrug-resistant H58 clade of Salmonella Typhi identifies inter- and intracontinental transmission events. *Nat. Genet.* **47,** 632–639 (2015).

25. Laxminarayan, R. *et al.* Antibiotic resistance-the need for global solutions. *Lancet Infect. Dis.* **13,** 1057–1098 (2013).

26. Lamrabet, O., Martin, M., Lenski, R. E. & Schneider, D. Changes in Intrinsic Antibiotic Susceptibility during a Long- Term Evolution Experiment with Escherichia coli. *MBio* **10,** (2019).

27. Baquero, F., Alvarez-Ortega, C. & Martinez, J. L. Ecology and evolution of antibiotic resistance. *Environ. Microbiol. Rep.* **1,** 469–476 (2009).

28. McLean, R. C. & San Millan, A. The Evolution of Antibiotic Resistance. *Science (80-.).* **365,** (2019).

29. Leski, T. A. *et al.* Multidrug-resistant tet(X)-containing hospital isolates in Sierra Leone. *Int. J. Antimicrob. Agents* **42,** 83–86 (2013).

30. Gilmore, M. S., Lebreton, F. & Schaik, W. van. Genomic transition of Enterococci from gut commensals to leading causes of multidrug-resistant hospital infection in the antibiotic era. *Curr. Opin. Biotechnol.* **23,** 1–7 (2013).

31. Poirel, L., Héritier, C. & Nordmann, P. Chromosome-Encoded Ambler Class D B-Lactamase of Shewanella oneidensis as a Progenitor of Carbapenem-Hydrolyzing Oxacillinase. *Society* **48,** 348–351 (2004).

32. San Millan, A. Evolution of Plasmid-Mediated Antibiotic Resistance in the Clinical Context. *Trends Microbiol.* **26,** 978–985 (2018).

33. Ma, X. *et al.* rpoB gene mutations and molecular characterization of rifampin-resistant Mycobacterium tuberculosis isolates from Shandong Province, China. *J. Clin. Microbiol.* **44,** 3409–3412 (2006).

34. Wang, H., Dzink-Fox, J. L., Chen, M. & Levy, S. B. Genetic characterization of highly fluoroquinolone-resistant clinical Escherichia coli strains from China: Role of acrR mutations. *Antimicrob. Agents Chemother.* **45,** 1515–1521 (2001).

35. González, I. *et al.* Fluoroquinolone resistance mutations in the parC, parE, and gyrA genes of clinical isolates of viridans group streptococci. *Antimicrob. Agents Chemother.* **42,** 2792–2798 (1998).

36. Conrad, S. *et al.* gyrA Mutations in high-level fluoroquinolone-resistant clinical isolates of Escherichia coli. *J. Antimicrob. Chemother.* **38,** 443–455 (1996).

37. Stokes, H. W. & Gillings, M. R. Gene flow, mobile genetic elements and the recruitment of antibiotic resistance genes into Gram-negative pathogens. *FEMS Microbiol. Rev.* **35,** 790–819

(2011).

38. Sultan, I. *et al.* Antibiotics, Resistome and Resistance Mechanisms: A Bacterial Perspective. *Front. Microbiol.* **9,** (2018).

39. Thomas, C. M. Paradigms of plasmid organization. *Mol. Microbiol.* **37,** 485–491 (2000).

40. Partridge, S. R., Kwong, S. M., Firth, N. & Jensen, S. O. Mobile genetic elements associated with antimicrobial resistance. *Clin. Microbiol. Rev.* **31,** 1–61 (2018).

41. Gillings, M. R. Integrons: Past, Present, and Future. *Microbiol. Mol. Biol. Rev.* **78,** 257–277 (2014).

42. Gillings, M. R. Class 1 integrons as invasive species. *Curr. Opin. Microbiol.* **38,** 10–15 (2017).

43. Gillings, M. *et al.* The evolution of class 1 integrons and the rise of antibiotic resistance. *J. Bacteriol.* **190,** 5095–5100 (2008).

44. Gillings, M. R. *et al.* Using the class 1 integron-integrase gene as a proxy for anthropogenic pollution. *ISME J.* **9,** 1269–1279 (2015).

45. Delavat, F., Miyazaki, R., Carraro, N., Pradervand, N. & van der Meer, J. R. The hidden life of integrative and conjugative elements. *FEMS Microbiol. Rev.* **41,** 512–537 (2017).

46. Guglielmini, J., Quintais, L., Garcillán-Barcia, M. P., de la Cruz, F. & Rocha, E. P. C. The repertoire of ice in prokaryotes underscores the unity, diversity, and ubiquity of conjugation. *PLoS Genet.* **7,** (2011).

47. Johnson, C. M. & Grossman, A. D. Integrative and Conjugative Elements (ICEs): What They Do and How They Work. *Annu. Rev. Genet.* **49,** 577–601 (2015).

48. Lood, R., Ertürk, G. & Mattiasson, B. Revisiting antibiotic resistance spreading in wastewater treatment plants - Bacteriophages as a much neglected potential transmission vehicle. *Front. Microbiol.* **8,** 1–7 (2017).

49. Hughes, D. & Andersson, D. I. Evolutionary Trajectories to Antibiotic Resistance. *Annu. Rev. Microbiol.* **71,** 579–596 (2017).

50. Baker-Austin, C., Wright, M. S., Stepanauskas, R. & McArthur, J. V. Co-selection of antibiotic and metal resistance. *Trends Microbiol.* **14,** 176–182 (2006).

51. Pal, C., Bengtsson-Palme, J., Kristiansson, E. & Larsson, D. G. J. Co-occurrence of resistance genes to antibiotics, biocides and metals reveals novel insights into their co-selection potential. *BMC Genomics* **16,** (2015).

52. Poole, K. At the Nexus of Antibiotics and Metals: The Impact of Cu and Zn on Antibiotic Activity and Resistance. *Trends Microbiol.* **25,** 820–832 (2017).

53. Bottery, M. J., Wood, A. J. & Brockhurst, M. A. Temporal dynamics of bacteria-plasmid coevolution under antibiotic selection. *ISME J.* 559–562 (2018). doi:10.1038/s41396-018-0276-9

54. Li, L. G., Xia, Y. & Zhang, T. Co-occurrence of antibiotic and metal resistance genes revealed in complete genome collection. *ISME J.* **11,** 651–662 (2017).

55. Bañuelos-Vazquez, L. A., Torres Tejerizo, G. & Brom, S. Regulation of conjugative transfer of plasmids and integrative conjugative elements. *Plasmid* **91,** 82–89 (2017).

56. Clokie, M. R. J., Millard, A. D., Letarov, A. V. & Heaphy, S. Phages in nature. *Bacteriophage* **1,**

31–45 (2011).

57. Torres-Barceló, C. The disparate effects of bacteriophages on antibiotic-resistant bacteria. *Emerg. Microbes Infect.* **7,** (2018).

58. Colavecchio, A., Cadieux, B., Lo, A. & Goodridge, L. D. Bacteriophages contribute to the spread of antibiotic resistance genes among foodborne pathogens of the Enterobacteriaceae family - A review. *Front. Microbiol.* **8,** 1–13 (2017).

59. Kenzaka, T., Tani, K. & Nasu, M. High-frequency phage-mediated gene transfer in freshwater environments determined at single-cell level. *ISME J.* **4,** 648–659 (2010).

60. Brown-Jaque, M., Calero-Cáceres, W. & Muniesa, M. Transfer of antibiotic-resistance genes via phage-related mobile elements. *Plasmid* **79,** 1–7 (2015).

61. Fortier, L.-C. *Chapter Five - The Contribution of Bacteriophages to the Biology and Virulence of Pathogenic Clostridia*. *Advances in Applied Microbiology* (2017).

62. DI., A. & D., H. Selection and Transmission of Antibiotic-Resistant Bacteria. *Microbiol. Spectr.* **5,** 1–17 (2017).

63. Johnston, C., Martin, B., Fichant, G., Polard, P. & Claverys, J. P. Bacterial transformation: Distribution, shared mechanisms and divergent control. *Nat. Rev. Microbiol.* **12,** 181–196 (2014).

64. Chen, I. & Dubnau, D. DNA uptake during bacterial transformation. *Nat. Rev. Microbiol.* **2,** 241–249 (2004).

65. Mutreja, A. Bacterial frequent flyers. *Nat. Rev. Microbiol.* **10,** 734 (2012).

66. Bengtsson-Palme, J. *et al.* The human gut microbiome as a transporter of antibiotic resistance genes between continents. *Antimicrob. Agents Chemother.* **59,** 6551–6560 (2015).

67. Memish, Z. A., Venkatesh, S. & Shibl, A. M. Impact of travel on international spread of antimicrobial resistance. *Int. J. Antimicrob. Agents* **21,** 135–142 (2003).

68. Wu, D., Wu, T., Liu, Q. & Yang, Z. The SARS-CoV-2 outbreak: what we know. *Int. J. Infect. Dis.* **94,** 44–48 (2020).

69. Tchatchouang, C. D. K. *et al.* Listeriosis outbreak in south africa: A comparative analysis with previously reported cases worldwide. *Microorganisms* **8,** (2020).

70. Cheng, R. A., Eade, C. R. & Wiedmann, M. Embracing diversity: Differences in virulence mechanisms, disease severity, and host adaptations contribute to the success of nontyphoidal salmonellaas a foodborne pathogen. *Front. Microbiol.* **10,** 1–20 (2019).

71. Jones, K. E. *et al.* Global trends in emerging infectious diseases. *Nature* **451,** 990–993 (2008).

72. Durso, L. M. & Cook, K. L. Impacts of antibiotic use in agriculture : what are the benefits and risks ? *Curr. Opin. Microbiol.* **19,** 37–44 (2014).

73. Alexander, T. W. *et al.* Farm-to-fork characterization of Escherichia coli associated with feedlot cattle with a known history of antimicrobial use. *Int. J. Food Microbiol.* **137,** 40–48 (2010).

74. Price, L. B. *et al.* Elevated risk of carrying gentamicin-resistant escherichia coli among U.S. poultry workers. *Environ. Health Perspect.* **115,** 1738–1742 (2007).

75. Ma, L. *et al.* Metagenomic Assembly Reveals Hosts of Antibiotic Resistance Genes and the

Shared Resistome in Pig, Chicken, and Human Feces. *Environ. Sci. Technol.* **50,** (2016).

76. Hu Y, Yang X, Li J, Lv N, Liu F, Wu J, Lin IY, Wu N, Weimer BC, Gao GF, Liu Y, Z. B. The transfer network of bacterial mobile resistome connecting animal and human microbiome. *Am. J. Vet. Res.* **82,** 6672–6681 (2016).

77. Zhu, B., Chen, Q., Chen, S. & Zhu, Y.-G. Does organically produced lettuce harbor higher abundance of antibiotic resistance genes than conventionally produced? *Environ. Heal. Perspect.* (2016). doi:10.1016/j.envint.2016.11.001

78. Blau, K. *et al.* The Transferable Resistome of Produce. *MBio* **9,** 1–15 (2018).

79. Carstens, C. K., Salazar, J. K. & Darkoh, C. Multistate Outbreaks of Foodborne Illness in the United States Associated With Fresh Produce From 2010 to 2017. *Front. Microbiol.* **10,** 1–15 (2019).

80. Wendel, A. M. *et al.* Multistate Outbreak of Escherichia coli O157:H7 Infection Associated with Consumption of Packaged Spinach, August–September 2006: The Wisconsin Investigation . *Clin. Infect. Dis.* **48,** 1079–1086 (2009).

81. Buchholz, U., Bernard, H., Werber, D., M.Böhmer, M. & Remschmidt, C. German Outbreak of Escherichia coli O104:H4 Associated with Sprouts. *N. Engl. J. Med.* **365,** 1763–1770 (2011).

82. Hoelzer, K. *et al.* Antimicrobial drug use in food-producing animals and associated human health risks: What, and how strong, is the evidence? *BMC Vet. Res.* **13,** 1–38 (2017).

83. Tang, K. L. *et al.* Restricting the use of antibiotics in food-producing animals and its associations with antibiotic resistance in food-producing animals and human beings: a systematic review and meta-analysis. *Lancet Planet. Heal.* **1,** e316–e327 (2017).

84. Vredenburg, J. *et al.* Quinolone-resistant Escherichia coli isolated from birds of prey in Portugal are genetically distinct from those isolated from water environments and gulls in Portugal, Spain and Sweden. *Environ. Microbiol.* **16,** 995–1004 (2014).

85. Wu, J., Huang, Y., Rao, D., Zhang, Y. & Yang, K. Evidence for Environmental Dissemination of Antibiotic Resistance Mediated by Wild Birds. *Front. Microbiol.* **9,** 1–12 (2018).

86. Furness, L. E., Campbell, A., Zhang, L., Gaze, W. H. & McDonald, R. A. Wild small mammals as sentinels for the environmental transmission of antimicrobial resistance. *Environ. Res.* **154,** 28–34 (2017).

87. Williams, S. H. *et al.* New York City House Mice (Mus musculus) as Potential Reservoirs for Pathogenic Bacteria and Antimicrobial Resistance Determinants. *MBio* **9,** e00624-18 (2018).

88. Weiss, D. *et al.* Antibiotic-Resistant Escherichia coli and Class 1 Integrons in Humans, Domestic Animals, and Wild Primates in Rural Uganda. *Appl. Environ. Microbiol.* **84,** 1–10 (2018).

89. Surette, M. & Wright, G. D. Lessons from the Environmental Antibiotic Resistome. *Annu. Rev. Microbiol.* **71,** 309–329 (2017).

90. Naidoo, Y., Valverde, A., Cason, E. D., Pierneef, R. E. & Cowan, D. A. A clinically important, plasmid-borne antibiotic resistance gene (β-lactamase TEM-116) present in desert soils. *Sci. Total Environ.* **719,** 137497 (2020).

91. Lau, C. H. F., van Engelen, K., Gordon, S., Renaud, J. & Topp, E. Novel antibiotic resistance determinants from agricultural soil exposed to antibiotics widely used in human medicine and animal farming. *Appl. Environ. Microbiol.* **83,** (2017).

92. Allen, H. K. *et al.* Call of the wild: Antibiotic resistance genes in natural environments. *Nat.*

Rev. Microbiol. **8,** 251–259 (2010).

93. Nesme, J. *et al.* Large-scale metagenomic-based study of antibiotic resistance in the environment. *Curr. Biol.* **24,** 1096–1100 (2014).

94. Perry, J. A., Westman, E. L. & Wright, G. D. The antibiotic resistome: What's new? *Curr. Opin. Microbiol.* **21,** 45–20 (2014).

95. Bhullar, K. *et al.* Antibiotic resistance is prevalent in an isolated cave microbiome. *PLoS One* **7,** 1–11 (2012).

96. Czekalski, N., Sigdel, R., Birtel, J., Matthews, B. & Bürgmann, H. Does human activity impact the natural antibiotic resistance background? Abundance of antibiotic resistance genes in 21 Swiss lakes. *Environ. Int.* **81,** 45–55 (2015).

97. Mkrtchyan, H. V., Russell, C. A., Wang, N. & Cutler, R. R. Could Public Restrooms Be an Environment for Bacterial Resistomes? *PLoS One* **8,** 1–6 (2013).

98. Kang, K. *et al.* The Environmental Exposures and Inner- and Intercity Traffic Flows of the Metro System May Contribute to the Skin Microbiome and Resistome. *Cell Rep.* **24,** 1190–1202.e5 (2018).

99. Narciso-da-Rocha, C. & Manaia, C. M. Multidrug resistance phenotypes are widespread over different bacterial taxonomic groups thriving in surface water. *Sci. Total Environ.* **563–564,** 1–9 (2016).

100. Garner, E. *et al.* Stormwater loadings of antibiotic resistance genes in an urban stream. *Water Res.* **123,** 144–152 (2017).

101. Wang, Y., Wang, C. & Song, L. Distribution of antibiotic resistance genes and bacteria from six atmospheric environments: Exposure risk to human. *Sci. Total Environ.* **694,** 133750 (2019).

102. Chamosa, L. S. *et al.* Lateral Antimicrobial Resistance Genetic Transfer is active in the open environment. *Sci. Rep.* **7,** 1–12 (2017).

103. Ju, F. *et al.* Antibiotic resistance genes and human bacterial pathogens: Co-occurrence, removal, and enrichment in municipal sewage sludge digesters. *Water Res.* **91,** 1–10 (2016).

104. Yoon, E. *et al.* Origin in Acinetobacter guillouiae and Dissemination of the Aminoglycoside-Modifying Enzyme Aph(3=)-VI. *MBio* **5,** (2014).

105. Poirel, L., Rodriguez-Martinez, J.-M., Mammeri, H., Liard, A. & Nordmann, P. Origin of Plasmid-Mediated Quinolone Resistance Determinant QnrA. *Antimicrob. Agents Chemother.* **49,** 3523–3525 (2005).

106. Pruden, A., Arabi, M. & Storteboom, H. N. Correlation between upstream human activities and riverine antibiotic resistance genes. *Environ. Sci. Technol.* **46,** 11541–11549 (2012).

107. Zhu, Y. G. *et al.* Continental-scale pollution of estuaries with antibiotic resistance genes. *Nat. Microbiol.* **2,** (2017).

108. Xiang, Q. *et al.* Spatial and temporal distribution of antibiotic resistomes in a peri-urban area is associated significantly with anthropogenic activities. *Environ. Pollut.* **235,** 525–533 (2018).

109. Peng, M. *et al.* Alterations of Salmonella Typhimurium Antibiotic Resistance under Environmental Pressure. *Appl. Environ. Microbiol.* **84,** (2018).

110. Suzuki, S., Kimura, M., Agusa, T. & Rahman, H. M. Vanadium accelerates horizontal transfer of tet(M) gene from marine Photobacterium to Escherichia coli. *FEMS Microbiol. Lett.* **336,** 52–

56 (2012).

111. O'Flaherty, E., Borrego, C. M., Balcázar, J. L. & Cummins, E. Human exposure assessment to antibiotic-resistant Escherichia coli through drinking water. *Sci. Total Environ.* **616–617,** 1356–1364 (2017).

112. Ma, L., Li, B. & Zhang, T. New insights into antibiotic resistome in drinking water and management perspectives: A metagenomic based study of small-sized microbes. *Water Res.* **152,** 191–201 (2019).

113. Vikesland, P. J. *et al.* Toward a Comprehensive Strategy to Mitigate Dissemination of Environmental Sources of Antibiotic Resistance. *Environ. Sci. Technol.* **51,** 13061–13069 (2017).

114. Gillings, M. R., Paulsen, I. T. & Tetu, S. G. Genomics and the evolution of antibiotic resistance. *Ann. N. Y. Acad. Sci.* **1388,** 92–107 (2017).

115. Pal, C., Bengtsson-Palme, J., Kristiansson, E. & Larsson, D. G. J. The structure and diversity of human, animal and environmental resistomes. *Microbiome* **4,** 54 (2016).

116. Martínez, J. L. Bottlenecks in the transferability of antibiotic resistance from natural ecosystems to human bacterial pathogens. *Front. Microbiol.* **2,** 1–6 (2012).

117. Bengtsson-Palme, J., Kristiansson, E. & Larsson, D. G. J. Environmental factors influencing the development and spread of antibiotic resistance. *FEMS Microbiol. Rev.* 68–80 (2017). doi:10.1093/femsre/fux053

118. Martínez, J. L. Ecology and Evolution of Chromosomal Gene Transfer between Environmental Microorganisms and Pathogens. *Microbiol. Spectr.* **6,** 1–16 (2018).

119. Smalla, K., Cook, K., Djordjevic, S. P., Klümper, U. & Gillings, M. Environmental dimensions of antibiotic resistance: Assessment of basic science gaps. *FEMS Microbiol. Ecol.* 1–6 (2018). doi:10.1093/femsec/fiy195

120. Varela, A. R. *et al.* Molecular evidence of the close relatedness of clinical, gull and wastewater isolates of quinolone-resistant Escherichia coli. *J. Glob. Antimicrob. Resist.* **3,** 286–289 (2015).

121. Zhou, Z. C. *et al.* Prevalence and transmission of antibiotic resistance and microbiota between humans and water environments. *Environ. Int.* **121,** 1155–1161 (2018).

122. Cantas, L. *et al.* A brief multi-disciplinary review on antimicrobial resistance in medicine and its linkage to the global environmental microbiota. *Front. Microbiol.* **4,** 1–14 (2013).

123. Forsberg, K. J., Reyes, A., Wang, B., Selleck, E. M. & Morten, O. A. The shared antibiotic resistome of soil bacteria and human pathogens. *Science* **337,** 1107–1111 (2012).

124. TIEDJE, J. M. *et al.* Antibiotic Resistance Genes in the Human-Impacted Environment: A One Health Perspective. *Pedosphere* **29,** 273–282 (2019).

125. Gouliouris, T. *et al.* Detection of vancomycin-resistant Enterococcus faecium hospital-adapted lineages in municipal wastewater treatment plants indicates widespread distribution and release into the environment. *Genome Res.* **29,** 626–634 (2019).

126. Varela, A. R., Macedo, G. N., Nunes, O. C. & Manaia, C. M. Genetic characterization of fluoroquinolone resistant Escherichia coli from urban streams and municipal and hospital effluents. *FEMS Microbiol. Ecol.* **91,** 1–12 (2015).

127. Ju, F. *et al.* Wastewater treatment plant resistomes are shaped by bacterial composition, genetic exchange, and upregulated expression in the effluent microbiomes. *ISME J.* (2018).

doi:10.1038/s41396-018-0277-8

128. Ng, C. *et al.* Characterization of metagenomes in urban aquatic compartments reveals high prevalence of clinically relevant antibiotic resistance genes in wastewaters. *Front. Microbiol.* **8,** 1–12 (2017).

129. Heuer, H., Schmitt, H. & Smalla, K. Antibiotic resistance gene spread due to manure application on agricultural fields. *Curr. Opin. Microbiol.* **14,** 236–243 (2011).

130. Mcallister, T. A. & Topp, E. Role of livestock in microbiological contamination of water : Commonly the blame , but not always the source. *Anim. Front.* **2,** 17–27 (2012).

131. Tamminen, M. *et al.* Tetracycline resistance genes persist at aquaculture farms in the absence of selection pressure. *Environ. Sci. Technol.* **45,** 386–391 (2011).

132. Garner, E. *et al.* Metagenomic Characterization of Antibiotic Resistance Genes in Full-Scale Reclaimed Water Distribution Systems and Corresponding Potable Systems. *Environ. Sci. Technol.* **52,** 6113–6125 (2018).

133. Fahrenfeld, N., Ma, Y., O'Brien, M. & Pruden, A. Reclaimed water as a reservoir of antibiotic resistance genes: Distribution system and irrigation implications. *Front. Microbiol.* **4,** 1–10 (2013).

134. Bürgmann, H. *et al.* Water and sanitation: An essential battlefront in the war on antimicrobial resistance. *FEMS Microbiol. Ecol.* **94,** (2018).

135. Obayiuwana, A., Ogunjobi, A., Yang, M. & Ibekwe, M. Characterization of Bacterial Communities and Their Antibiotic Resistance Profiles in Wastewaters Obtained from Pharmaceutical Facilities in Lagos and. *Int. J. Environ. Res. Public Health* **15,** (2018).

136. Muziasari, W. I. *et al.* The resistome of farmed fish feces contributes to the enrichment of antibiotic resistance genes in sediments below baltic sea fish farms. *Front. Microbiol.* **7,** 1–10 (2017).

137. Xie, W. Y., Shen, Q. & Zhao, F. J. Antibiotics and antibiotic resistance from animal manures to soil: A review. *Eur. J. Soil Sci.* 181–195 (2017). doi:10.1111/ejss.12494

138. Chen, B. *et al.* Class 1 integrons, selected virulence genes, and antibiotic resistance in Escherichia coli isolates from the Minjiang River, Fujian Province, China. *Appl. Environ. Microbiol.* **77,** 148–155 (2011).

139. He, L. Y. *et al.* Discharge of swine wastes risks water quality and food safety: Antibiotics and antibiotic resistance genes from swine sources to the receiving environments. *Environ. Int.* **92–93,** 210–219 (2016).

140. Power, M. L. *et al.* Escherichia coli out in the cold: Dissemination of human-derived bacteria into the Antarctic microbiome. *Environ. Pollut.* **215,** 58–65 (2016).

141. Jia, S. *et al.* Fate of antibiotic resistance genes and their associations with bacterial community in livestock breeding wastewater and its receiving river water. *Water Res.* **124,** 259–268 (2017).

142. Jechalke, S. *et al.* Effects of 100 years wastewater irrigation on resistance genes, class 1 integrons and IncP-1 plasmids in Mexican soil. *Front. Microbiol.* **6,** 1–10 (2015).

143. Muziasari, W. I. *et al.* Sulphonamide and trimethoprim resistance genes persist in sediments at Baltic Sea aquaculture farms but are not detected in the surrounding environment. *PLoS One* **9,** (2014).

144. Vaz-Moreira, I., Nunes, O. C. & Manaia, C. M. Bacterial diversity and antibiotic resistance in water habitats: Searching the links with the human microbiome. *FEMS Microbiology Reviews* **38,** (2014).

145. Adegoke, A. A., Amoah, I. D., Stenström, T. A., Verbyla, M. E. & Mihelcic, J. R. Epidemiological Evidence and Health Risks Associated With Agricultural Reuse of Partially Treated and Untreated Wastewater: A Review. *Front. Public Heal.* **6,** 1–20 (2018).

146. Leonard, A. F. C. *et al.* Exposure to and colonisation by antibiotic-resistant E. coli in UK coastal water users: Environmental surveillance, exposure assessment, and epidemiological study (Beach Bum Survey). *Environ. Int.* **114,** 326–333 (2018).

147. Karkman, A., Pärnänen, K. & Larsson, D. G. J. Fecal pollution explains antibiotic resistance gene abundances in anthropogenically impacted environments. *Nat. Commun.* **10,** (2019).

148. Czekalski, N., Gascón Díez, E. & Bürgmann, H. Wastewater as a point source of antibiotic-resistance genes in the sediment of a freshwater lake. *ISME J.* **8,** 1381–1390 (2014).

149. Lehmann, K. *et al.* Trace levels of sewage effluent are sufficient to increase class 1 integron prevalence in freshwater biofilms without changing the core community. *Water Res.* **106,** 163–170 (2016).

150. Bondarczuk, K. & Piotrowska-Seget, Z. Microbial diversity and antibiotic resistance in a final effluent-receiving lake. *Sci. Total Environ.* **650,** 2951–2961 (2019).

151. Hendriksen, R. S. *et al.* Global monitoring of antimicrobial resistance based on human sewage. *Nat. Commun.* **10,** (2019).

152. Larsson, D. G. J. *et al.* Critical knowledge gaps and research needs related to the environmental dimensions of antibiotic resistance. *Environ. Int.* **117,** 132–138 (2018).

153. Zhou, L. J. *et al.* Occurrence and fate of eleven classes of antibiotics in two typical wastewater treatment plants in South China. *Sci. Total Environ.* **452–453,** 365–376 (2013).

154. Chen, H. *et al.* Antibiotics in the coastal environment of the Hailing Bay region, South China Sea: Spatial distribution, source analysis and ecological risks. *Mar. Pollut. Bull.* **95,** 365–373 (2015).

155. Zhou, L. J. *et al.* Excretion masses and environmental occurrence of antibiotics in typical swine and dairy cattle farms in China. *Sci. Total Environ.* **444,** 183–195 (2013).

156. Zhang, M. *et al.* Fate of veterinary antibiotics during animal manure composting. *Sci. Total Environ.* **650,** 1363–1370 (2019).

157. Larsson, D. G. J. Pollution from drug manufacturing: review and perspectives. *Philos. Trans. R. Soc.* **369,** (2014).

158. Hoa, P. T. P. *et al.* Antibiotic contamination and occurrence of antibiotic-resistant bacteria in aquatic environments of northern Vietnam. *Sci. Total Environ.* **409,** 2894–2901 (2011).

159. Varela, A. R., André, S., Nunes, O. C. & Manaia, C. M. Insights into the relationship between antimicrobial residues and bacterial populations ina hospital-urban wastewater treatment plant system. *Water Res.* **54,** 327–336 (2014).

160. Rutgersson, C. *et al.* Fluoroquinolones and qnr genes in sediment, water, soil, and human fecal flora in an environment polluted by manufacturing discharges. *Environ. Sci. Technol.* **48,** 7825–7832 (2014).

161. Marathe, N. P. *et al.* Functional metagenomics reveals a novel carbapenem-hydrolyzing

mobile beta-lactamase from Indian river sediments contaminated with antibiotic production waste. *Environ. Int.* **112,** 279–286 (2018).

162. Larsson, D. G. J. Antibiotics in the environment. *Ups. J. Med. Sci.* **119,** 108–112 (2014).

163. Zhao, Y. *et al.* Evidence for co-selection of antibiotic resistance genes and mobile genetic elements in metal polluted urban soils. *Sci. Total Environ.* **656,** 512–520 (2019).

164. Jiao, Y. N., Chen, H., Gao, R. X., Zhu, Y. G. & Rensing, C. Organic compounds stimulate horizontal transfer of antibiotic resistance genes in mixed wastewater treatment systems*.Chemosphere* (2017).

165. Ashbolt, N. J. Human Health Risk Assessment (HHRA) for Environmental Development and Transfer of Antibiotic Resistance. *Environ. Health Perspect.* **121,** 993–1002 (2013).

166. Sandegren, L. Selection of antibiotic resistance at very low antibiotic concentrations. *Ups. J. Med. Sci.* **119,** 103–107 (2014).

167. Andersson, D. I. & Hughes, D. Microbiological effects of sublethal levels of antibiotics. *Nat. Rev. Microbiol.* **12,** 465–478 (2014).

168. Almakki, A., Jumas-Bilak, E., Marchandin, H. & Licznar-Fajardo, P. Antibiotic resistance in urban runoff. *Sci. Total Environ.* **667,** 64–76 (2019).

169. Gullberg, E. *et al.* Selection of resistant bacteria at very low antibiotic concentrations. *PLoS Pathog.* **7,** 1–9 (2011).

170. Murray, A. K. *et al.* Novel Insights into Selection for Antibiotic Resistance in Complex Microbial Communities. *MBio* **9,** 1–12 (2018).

171. Berglund, B. Environmental dissemination of antibiotic resistance genes and correlation to anthropogenic contamination with antibiotics. *Infect. Ecol. Epidemiol.* **5,** 28564 (2015).

172. Bernier, S. P. & Surette, M. G. Concentration-dependent activity of antibiotics in natural environments. *Front. Microbiol.* **4,** (2013).

173. Bruchmann, J., Kirchen, S. & Schwartz, T. Sub-inhibitory concentrations of antibiotics and wastewater influencing biofilm formation and gene expression of multi-resistant Pseudomonas aeruginosa wastewater isolates. *Environ. Sci. Pollut. Res.* **20,** 3539–3549 (2013).

174. Gillings, M. R. & Stokes, H. W. Are humans increasing bacterial evolvability? *Trends Ecol. Evol.* **27,** 346–352 (2012).

175. Zhang, Y. *et al.* Sub-inhibitory concentrations of heavy metals facilitate the horizontal transfer of plasmid-mediated antibiotic resistance genes in water environment. *Environ. Pollut.* **237,** 74–82 (2018).

176. Chow, L., Waldron, L. & Gillings, M. R. Potential impacts of aquatic pollutants: Sub-clinical antibiotic concentrations induce genome changes and promote antibiotic resistance. *Front. Microbiol.* **6,** 1–10 (2015).

177. Murray, A. K. *et al.* Novel insights into selection for antibiotic resistance in complex microbial communities. *MBio* **9,** 1–12 (2018).

178. Schmitt, H., Stoob, K., Hamscher, G., Smit, E. & Seinen, W. Tetracyclines and tetracycline resistance in agricultural soils: Microcosm and field studies. *Microb. Ecol.* **51,** 267–276 (2006).

179. Blau, K. *et al.* Soil texture-depending effects of doxycycline and streptomycin applied with manure on the bacterial community composition and resistome. *FEMS Microbiol. Ecol.* **94,** 1–

11 (2017).

180. Kyselková, M. *et al.* Cow excrements enhance the occurrence of tetracycline resistance genes in soil regardless of their oxytetracycline content. *Chemosphere* **93,** 2413–2418 (2013).

181. Choung, S. *et al.* Transfer of antibiotic resistance plasmids in pure and activated sludge cultures in the presence of environmentally representative micro-contaminant concentrations. *Sci. Total Environ.* **468–469,** 813–820 (2014).

182. Berendonk, T. U. *et al.* Tackling antibiotic resistance: The environmental framework. *Nat. Rev. Microbiol.* **13,** 310–317 (2015).

183. Ramirez, M. S., Nikolaidis, N. & Tolmasky, M. E. Rise and dissemination of aminoglycoside resistance: The aac(6′)-Ib paradigm. *Front. Microbiol.* **4,** 1–13 (2013).

184. Ramirez .S, M. & Tolmasky .E, M. Aminoglycoside Modifing Enzymes. *Drug Resist. Updat.* **13,** 151–171 (2011).

185. Heuer, H. *et al.* Gentamicin resistance genes in environmental bacteria: Prevalence and transfer. *FEMS Microbiol. Ecol.* **42,** 289–302 (2002).

186. George, J. & Halami, P. M. Sub-inhibitory concentrations of gentamicin triggers the expression of aac(6′)Ie-aph(2″)Ia, chaperons and biofilm related genes in Lactobacillus plantarum MCC 3011. *Res. Microbiol.* **168,** 722–731 (2017).

187. Jutkina, J., Marathe, N. P., Flach, C. F. & Larsson, D. G. J. Antibiotics and common antibacterial biocides stimulate horizontal transfer of resistance at low concentrations. *Sci. Total Environ.* **616–617,** 172–178 (2018).

188. Bergk Pinto, B., Maccario, L., Dommergue, A., Vogel, T. M. & Larose, C. Do Organic Substrates Drive Microbial Community Interactions in Arctic Snow? *Front. Microbiol.* **10,** 1–13 (2019).

189. Zhu, C. *et al.* Snow microbiome functional analyses reveal novel aspects of microbial metabolism of complex organic compounds. *Microbiologyopen* **9,** 1–12 (2020).

Chapter II. Optimization of nucleic acid extraction from soil. Evaluation of the effect of nucleic acid extraction and sequencing depth on taxonomic, functional and ARG discovery.

Abstract

Soil ecosystems harbour the highest biodiversity on our planet. Although next generation sequencing (NGS) techniques have increased our access to the soil microbiome, each step of soil metagenomics presents inherent biases that prevent the accurate definition of the soil microbiome and its ecosystem function. Biases related to nucleic acid extraction have been particularly well documented with no one method being able to avoid them. In this study, we compared the phylogenetic and functional richness detected by five DNA extraction methods and the phylogenetic richness measured by four RNA extraction methods applied to two different soils. The V3-V4 hypervariable region of the 16S rRNA gene and cDNA was sequenced to determine the taxonomical richness measured by each method at the genus level. Functional and ARG richness were evaluated by metagenomics sequencing. Despite variable DNA extraction methods, sequencing depth had a greater influence on bacterial richness discovery at both taxonomical and functional levels than these extraction methods, whereas DNA extraction methods and resistome variability between triplicates played a more important role than sequencing depth on ARG richness discovery. Furthermore, at an equal sequencing depth, the differences observed between DNA extraction methods in both soils were more likely a product of random subsampling than that of the DNA extraction itself. Therefore, an optimization of each step of soil metagenomics workflow is needed in order to sequence samples at an equal depth and to be able to perform more accurate metagenomics comparisons.

Introduction

The soil ecosystem arguably harbours the highest diversity of microorganisms of any ecosystem[1]. Unravelling the composition and function of the soil microbiome is critical to better understanding the role of these microbial communities in soil function and ecosystem services. The use of next generation sequencing (NGS) techniques has increased our access to the microbial communities present in soil, especially to the large proportion of uncultured microorganisms[2,3]. Nevertheless, each methodological step from soil sampling to sequence annotation presents inherent biases that limit the depth and reliability of soil microbiome analyses[4–7]. Of all these biases, the ones associated with DNA extraction have been particularly emphasized for their effects[8,9]. DNA can be adsorbed by soil compounds such as clay[10,11], which when combined with the presence of lysis-recalcitrant bacteria[12], reduces DNA extraction efficiency. Moreover, organic matter and humic acids that are known to potentially inhibit enzymatic reactions[13] are often co-extracted with DNA. These biases also apply to RNA extraction[14], which is increasingly used to identify potentially active communities and genes. Many studies have compared DNA and RNA extraction methods and documented the biases imposed by lysis procedures[15–18]. This has provoked proposals of different methods for nucleic acid extraction and purification from soil over the past few decades[15,19,20] in an attempt to obtain an unbiased picture of soil microbiome biodiversity[9]. Nevertheless, no method has been shown to overcome all the biases described above and the debate on the choice of nucleic extraction methods is still ongoing.

The criteria used to define the performance of a nucleic acid extraction method vary between studies and range from nucleic acid yield to phylogenetic diversity. However, higher yields, purity, and integrity do not always imply an improvement in bacterial diversity discovery[4]. Analyses of the relative abundance of taxonomic groups have a limited potential for selecting nucleic acid extraction methods, since the biases

associated with soil metagenomics prevent us from determining the actual distribution of soil microbial populations within a community. Furthermore, nucleic acid extraction methods may modify the relative abundance of detected communities without affecting bacterial richness discovery. The goal of this paper was not to resolve the debate concerning the use of relative abundance versus richness. We believe that bacterial richness measurements provide a more objective comparison of the performance of nucleic acid extraction methods, since biodiversity calculations with a limited set of sequences are strongly biased by evenness, which depends on the number of sequences. In addition, the actual relative abundances of the soil microbiome remain unknown.

On the other hand, sequencing depth has also shown an important effect on bacterial discovery. Several studies have shown that low sequencing depths may bias the evaluation of microbial communities, both in terms of richness and diversity[21–23]. However, to the best of our knowledge, the relative contribution of sequencing depth and DNA or RNA extraction to soil bacterial richness discovery has not been evaluated yet. In this study, we compared the phylogenetic and functional richness detected by two novel semi-automated methods for DNA extraction and purification to that measured by two commercial kits and the phenol/chloroform method as described by Griffiths *et al* [24]. Since antibiotic resistance functions are rare in the environmental microbiome, ARG richness discovery may be more unequal depending on sequencing depth and, therefore, follow a different pattern than other functional richness discovery. Thus, the ARG richness detected by these five methods was also assessed. Furthermore, we evaluated the effect of sequencing depth on the taxonomic richness measured by two semi-automated methods for RNA extraction and purification to that measured by a commercial kit and the phenol/chloroform method. Nucleic acids were extracted from two soil samples and the V3-V4 hypervariable region of the 16S rRNA gene and cDNA was sequenced to determine the Operational Taxonomic Unit (OTU). DNA and cDNA richness measured by each method was defined at the genus level for each sequence using the Ribosome Database Project (RDP) database and the RDP Bayesian classifier[25]. Metagenomics sequences obtained from DNA samples were annotated using MEGAN6[26] and the SEED hierarchical subsystems. The third level of the SEED classification was selected for functional richness discovery analysis. Finally, metagenomic reads were blasted against the CARD antibiotic gene database[27] using Diamond[28] to identify ARGs.

Materials and methods

Soil sampling

Two soils were selected for this study (Table 1). Soils were sampled at an experimental farm (Scottish Agricultural College, Craibstone, Scotland, Grid reference NJ872104) and at a field planted with mature corn at La Côte Saint Andre, France. All samples were kept at 4°C before DNA extraction. Aliquots of 250 mg of sample were taken and frozen to -20°C prior RNA extraction. Details about the Scottish Agricultural College soil composition were provided by Kemp *et al.*[29] Physical characterization of the soil from La Côte de Saint André was performed by the CESAR (Centre Scientifique Agricole Régional) using standard methods (NFX 31-107, NFX 31-117, ISO 10694, ISO 13878).

	Scottish Agricultural College	**La Côte de Saint André**
Sand	73.85%	42.9%
Silt	20.04%	43.6%
Clay	6.11%	13.5%
pH	4.5	7.24
Organic matter	5.97%	2.92 %
Organic C	3.79%	1.7 %
Total N	0.45%	0.17 %

Table 1. Physical characterization of samples selected for this study.

DNA extraction and purification

DNA was extracted from 250 mg of sample using the DNeasy PowerSoil Kit (QIAGEN), the ZymoBIOMICS DNA Mini Kit (Zymo Research) and the Phenol/Chloroform extraction method described by Griffiths *et al* [24], as well as a new semi-automated protocol in which the Maxwell® RSC Instrument (Promega) and a prototype version of the Maxwell® Fecal Microbiome Kit (Promega) are used for DNA purification. Two modifications of this protocol were tested, referred as Maxwell 1 and Maxwell 2 methods. All DNA extractions were performed in triplicate. In the Maxwell 1 method, 250 mg of sample were diluted in 1 ml of Lysis Buffer (Promega) and heated for 5 minutes at 95 °C. Samples underwent bead-beating twice at 5.5 m/s for 30 seconds in Lysis Matrix E tubes (MP Biomedicals) and centrifuged at 10600 x g for 5 minutes. Then, 300 µl of supernatant were added to 300 µl of Binding Buffer (Promega) and loaded into a Maxwell RSC cartridge containing magnetic beads for DNA purification on the Maxwell RSC Instrument, according to the Technical Manual TM473. A second purification using the ProNex Size-Selective Purification System (Promega) was carried out to reduce humic acids carryover. In the Maxwell 2 method, two variants were introduced in the previously described protocol: 500 µl of Lysis Buffer were mixed with 500 µl of 0.5 M Sodium Phosphate Buffer (pH 7.0) and added to 250 mg of sample, and cells were lysed without bead-beating.

DNA quantification and quantitative PCR (qPCR) assays

DNA concentrations were measured using the Qubit Fluorometer and the Qubit dsDNA HS Assay Kit (ThermoFisher). Then, total bacterial abundance was estimated by quantifying the V3 region of the 16S rRNA gene by qPCR using the "universal" primers 341F (5'-CCT ACG GGA GGC AGC AG- 3') and 534R (5'-ATT ACC GCG GCT GCT GGC A-3')[30,31]. qPCR assays were carried out using the Corbett Rotor-Gene 6000 (QIAGEN) in a 20 µl reaction volume containing GoTaq® qPCR Master Mix (Promega), 0.75 µM of each primer and 2 µl of DNA at ≤2.5 ng/µl. Two non-template controls were also included in all the assays. Standard curves for all the assays were obtained using 10-fold serial dilutions of a linearized plasmid pGEM®-T Easy Vector (10^2 to 10^7 copies) containing the 16S rRNA gene of *Pseudomonas aeruginosa* PAO1. Cycling conditions for qPCR amplification were 95 °C for 2 minutes followed by 30 cycles of 95 °C for 15 seconds, 60 °C for 30 seconds and 72 °C for 30 seconds. Melting curves were generated after amplification by increasing the temperature from 60 °C to 95 °C.

RNA extraction and purification

RNA was extracted from 250 mg of sample using the RNeasy PowerSoil Total RNA Kit (QIAGEN), two semi-automated protocols using the Maxwell® RSC Instrument (Promega) and either the Maxwell® RSC miRNA Tissue Kit (Promega) or the Maxwell® RSC simplyRNA Tissue Kit (Promega), and the Phenol/Chloroform extraction method described by Griffiths *et al* [24] with an extra DNase treatment using the RQ1 RNase-Free DNase (Promega). In semi-automated protocols, 250 mg of sample were diluted in 1 ml of CTAB Buffer (Promega) with 2% 1-thioglycerol and underwent bead-beating at 5.5 m/s for 30 seconds in Lysis Matrix E tubes (MP Biomedicals). Then, samples were centrifuged at 10600 x g for 5 minutes. Finally, 300 µl of supernatant were added to 300 µl of Lysis Buffer (Promega), and 400 µl of this mix were loaded into a Maxwell RSC cartridge containing magnetic beads for RNA purification on the Maxwell RSC Instrument using the Maxwell miRNA Tissue Kit (according to the Technical Manual TM441) or the Maxwell simplyRNA Kit (according to the Technical Manual TM416). Both protocols using Promega RNA extraction kits include DNase treatment.

16S rRNA gene and cDNA V3-V4 amplicon sequencing and analysis

All DNA samples were diluted to 2.5 ng/µl before 16S rRNA gene PCR amplification and sequencing, except for some samples extracted using the ZymoBIOMICS DNA Kit, which were already at lower concentrations. Regarding RNA, samples were diluted to 1 ng/µl and retrotranscribed using the GoScript™ Reverse Transcriptase (Promega) and random primers (Promega). Then, the V3-V4 hypervariable regions of bacterial 16S rRNA gene and cDNA were amplified using the Titanium Taq DNA Polymerase (Takara

Clontech) and the Platinum Taq Polymerase (Invitrogen), respectively, and forward 341F with Illumina overhang (5'-TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG CCT ACG GGN GGC WGC AG-3') and reverse 785F with Illumina overhang (5'-GTC TCG TGG GCT CGG AGA TGT GTA TAA GAG ACA GGT CTC GTG GGC TCG GAG ATG TGT ATA AGA GAC AGG ACT ACH VGG GTA TCT AAT CC-3') primers[5] to determine the Operational Taxonomic Unit (OTU). Cycling conditions for PCR amplification were 95 °C for 3 minutes followed by 25 cycles of 95 °C for 30 seconds, 55 °C for 30 seconds and 72 °C for 30 seconds and a final extension step at 72°C for 5 minutes. DNA libraries were prepared based on Illumina's "16S Metagenomics Library Prep Guide" (15044223 Rev. B) using the Platinum Taq DNA Polymerase (Invitrogen) and the Nextera XT Index Kit V2 (Illumina). DNA and cDNA sequencing with a 15% PhiX spike-in was performed using the MiSeq System and the MiSeq Reagent Kit v3 (Illumina). Reads were trimmed to meet a quality score of Q20. Pair-ended reads were assembled using PANDAseq[32] at a sequence length between 410 and 500 bp and an overlap length between 20 and 100 bp, using the rdp_mle algorithm. Then, each of the DNA and cDNA sequences was annotated using the Ribosome Database Project (RDP) database and the RDP Bayesian classifier using an assignment confidence cut-off of 0.6[25] and singletons were removed. In order to evaluate the effect of sequencing depth in taxonomic richness assessment, the genus richness detected in each sample was determined using the vegan package (version 2.5-6[33]) in R (version 3.5.1) and plotted as a function of sequencing depth. Then, triplicates from each DNA extraction method/soil couple were pooled and a subsample of each method pool was randomly rarefied at the lowest size of the cohort using the vegan package in R. Venn diagrams were obtained using the VennDiagram package (version 1.6.20[34]) in R and the number of sequences annotated as unique genera (*i.e.*, genera detected by a single method) were inferred from the rarefied sample. Finally, the distribution of each unique genera before rarefaction and the number of sequences associated to those that were detected by a single method before rarefaction were determined.

Metagenomics sequencing and analysis

Metagenomics libraries were prepared from <1 ng of DNA using the Nextera XT Library Prep Kit and Indexes (Illumina), as detailed in Illumina's "Nextera XT DNA Library Prep Kit" reference guide (15031942 v03). DNA sequencing with a 1% PhiX spike-in was performed using the MiSeq System and the MiSeq Reagent Kit v2 (Illumina). Reads were trimmed and filtered using USEARCH[35] where all nucleotides had a quality score of Q20 and above were selected. In addition, sequences that had one nucleotide below Q20 were also accepted, but the nucleotide below Q20 was replaced with an "N". A minimum length of 120 bp was required before blasting against the nr database using Diamond default parameters and a coverage of 60%. Sequences were functionally annotated using MEGAN6 and the SEED hierarchical subsystems. The third level of hierarchical functional subsystems classification was selected for richness discovery analysis and singletons were removed. Samples extracted from both soils using the ZymoBIOMICS DNA Mini Kit were excluded from the analysis, since the ensemble of annotated reads from extraction triplicates did not add up to 10000. To evaluate the effect of sequencing depth in functional richness assessment, the functional class richness detected in each sample was determined using the vegan package in R and plotted as a function of sequencing depth. Then, after pooling triplicates from each method/soil couple, a subsample of each method pool was randomly rarefied at the lowest size of the cohort using the vegan package in R. Venn diagrams were obtained using the VennDiagram R package, and the number of sequences annotated as unique functional classes (*i.e.*, functional classes detected by a single method) were inferred from the rarefied sample. Then, the distribution of each unique functional class before rarefaction and the number of sequences associated to those that were detected by a single method before rarefaction were determined. Finally, metagenomic reads were blasted against the CARD antibiotic gene database using Diamond to identify ARGs. The obtained results were filtered at a minimum identity of 60%, a minimum length of 50 amino acids and an e-value of $10e^{-10}$. The best hit was chosen.

ARGs that had less than 10 copies total from in all samples were removed and the ARG richness measured in each sample was determined and plotted as a function of sequencing depth.

Results

DNA quantification and quantitative PCR (qPCR) assays

DNA samples were quantified and amplified before library preparation to assess the concentrations and number of copies of the 16S rRNA gene obtained by each method (Table 2). Both DNA concentrations and 16S rRNA gene copies obtained using the ZymoBIOMICS DNA Kit were lower than those measured by the other methods for both soils. The highest DNA concentrations were obtained from both soils using the Phenol/Chloroform method and the DNeasy PowerSoil Kit. However, when performing DNA amplification, similar numbers of 16S rRNA gene copies were obtained after DNA extraction using the Maxwell 1 and 2 methods, which are modifications of the prototype Maxwell Fecal Microbiome Kit (Promega). Therefore, PCR inhibition was greater in samples with higher DNA concentrations (*i.e*. extracted using the Phenol/Chloroform method and the DNeasy PowerSoil Kit). In order to reduce amplification biases and normalize the amount of starting material for library preparation, all samples were diluted to 2.5 ng/µl before 16S rRNA gene PCR amplification and sequencing.

Soil	Method	DNA concentration (ng/µl)	16S rRNA gene copies/µl
Scottish Agricultural College soil	Maxwell 1	9.05 ± 1.35	153,955 ± 52,412
	Maxwell 2	6.01 ± 1.50	64,279 ± 9,133
	Phenol/Chloroform method	23.40 ± 1.90	189,213 ± 15,036
	DNeasy PowerSoil Kit	17 ± 3.12	165,311 ± 65,886
	ZymoBIOMICS DNA Mini Kit	2.79 ± 0.64	20,056 ± 27,979
La Côte de Saint André soil	Maxwell 1	6.43 ± 0.97	160,608 ± 14,701
	Maxwell 2	5.06 ± 0.54	170,391 ± 16,323
	Phenol/Chloroform method	15.20 ± 2.76	211,949 ± 45,407
	DNeasy PowerSoil Kit	22.90 ± 1.15	88,904 ± 25,638
	ZymoBIOMICS DNA Mini Kit	0.12 ± 0.06	2,557 ± 1,376

Table 2. DNA concentrations and 16S rRNA gene copy numbers from the Scottish Agricultural College (Craibstone, Scotland) and La Côte de Saint André (France) soils (250 mg) using different DNA extraction and purification strategies. Maxwell 1 and 2: modifications of the prototype Maxwell Fecal Microbiome Kit (Promega). All data (averages and standard deviations) are based on three separate soil samples for each method.

Sequencing depth effect on taxonomic and functional richness discovery

In order to assess the relative contribution of sequencing depth and DNA and RNA extraction methods on taxonomic richness discovery, the richness measured in each sample was determined and plotted as a function of sequencing depth (Figure 1). The functional richness detected in each DNA sample was also plotted as a function of sequencing depth (Figure 2). Less than 10000 reads were obtained from ZymoBIOMICS DNA Mini Kit (Zymo Research) triplicate pools after metagenomics sequencing and annotation. Therefore, they were excluded from functional richness analysis. Sequencing depth showed a higher influence on bacterial richness discovery at both taxonomical and functional levels than DNA extraction methods: measured richness increased proportionally with sequencing depth regardless of the method used for DNA extraction. Furthermore, DNA extraction triplicates that had been sequenced at different depths had access to a different proportion of bacterial DNA richness (Figures 1 and 2). The same tendency was observed when extracting RNA (Figure 1): sequencing depth had a higher impact on genus richness discovery than the choice of the RNA extraction method. In addition, a PCoA analysis showed four main clusters representing the total and active communities extracted from both soil samples (Figure

3). Globally, samples from the same soil extracted using different DNA or RNA extraction methods showed a similar composition. However, samples sequenced at a lower depth were more distant to the rest of the cluster, regardless of the method used for DNA or RNA extraction. Given these results, DNA samples were normalized by size (sequence or read numbers) before comparing the taxonomic and functional richness measured by different DNA extraction methods in order to limit the effect of sequencing depth. After removing singletons and pooling method triplicates, subsamples of each method were randomly rarefied at 58645 sequences annotated as genera for the Scottish Agricultural College soil, 36311 sequences annotated as genera for La Côte de Saint André soil, 32693 sequences annotated as functions for the Scottish Agricultural College soil and 159748 sequences annotated as functions for La Côte de Saint André soil, as this represented values that all samples could meet, even the samples with the lowest number of reads.

Figure 1. Effect of sequencing depth on taxonomic richness discovery (top: DNA extraction; bottom:RNA extraction). Maxwell 1 and 2: modifications of the prototype Maxwell Fecal Microbiome Kit (Promega). PowerSoil: DNeasy PowerSoil Kit (QIAGEN). Zymo: ZymoBIOMICS DNA Mini Kit (Zymo Research). miRNA: Maxwell RSC miRNA Tissue Kit (Promega). simplyRNA: Maxwell RSC simplyRNA Tissue Kit (Promega). Triangles: DNA from the Scottish Agricultural College soil. Circles: DNA from La Côte de Saint André soil. Diamonds: cDNA from the Scottish Agricultural College soil. Squares: cDNA from La Côte de Saint André soil. Yellow and green: Promega's Maxwell 1 and 2 methods, respectively. Orange: Phenol/Chloroform DNA and RNA methods. Grey: DNeasy and RNeasy PowerSoil Kit (QIAGEN). Blue: ZymoBIOMICS DNA Mini Kit (Zymo Research). Dark blue Maxwell RSC miRNA Tissue Kit (Promega). Pink: Maxwell RSC simplyRNA Tissue Kit (Promega). The taxonomical richness measured in each sample was determined and plotted as a function of sequencing depth. For each method/soil pair, triplicates were performed and plotted in the graph.

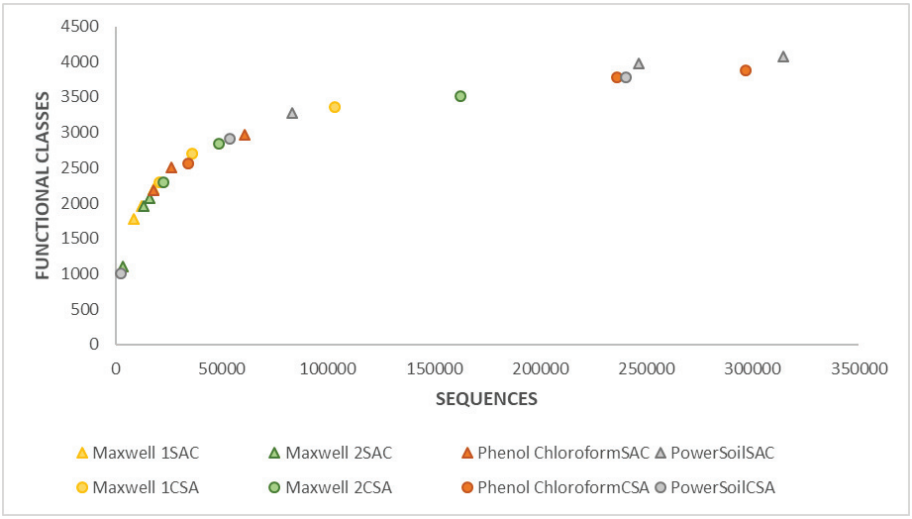


Figure 2. Effect of sequencing depth on functional richness discovery. Maxwell 1 and 2: modifications of the prototype Maxwell Fecal Microbiome Kit (Promega). PowerSoil: DNeasy PowerSoil Kit (QIAGEN). Triangles: DNA from the Scottish Agricultural College soil. Circles: DNA from La Côte de Saint André soil. Yellow and green: Promega's Maxwell 1 and 2 methods, respectively. Orange: Phenol/Chloroform method. Grey: DNeasy PowerSoil Kit (QIAGEN). The functional richness measured in each sample was determined and plotted as a function of sequencing depth. For each method/soil pair, triplicates were performed and plotted in the graph. Less than 10000 reads were obtained from ZymoBIOMICS DNA Mini Kit (Zymo Research) triplicate pools after metagenomics sequencing and annotation. Therefore, they were excluded from functional richness analysis.

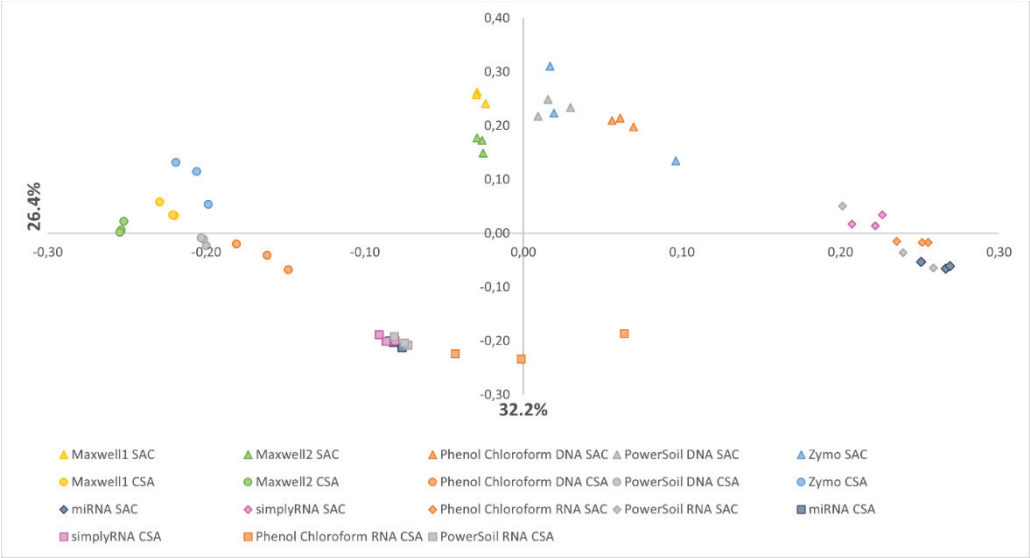


Figure 3. PCoA showing the total and active communities extracted from the Scottish Agricultural College soil and La Côte de Saint André soil using different DNA and RNA extraction methods. Maxwell 1 and 2: variants of the prototype Maxwell® Fecal Microbiome Kit (Promega). PowerSoil: DNeasy PowerSoil Kit (QIAGEN). Zymo: ZymoBIOMICS DNA Mini Kit (Zymo Research). miRNA: Maxwell® RSC miRNA Tissue Kit (Promega). simplyRNA: Maxwell® RSC simplyRNA Tissue Kit (Promega). Triangles: DNA from the Scottish Agricultural College soil. Circles: DNA from La Côte de Saint André soil. Diamonds: cDNA from the Scottish Agricultural College soil. Squares: cDNA from La Côte de Saint André soil. Yellow and green: Promega's Maxwell 1 and 2 methods, respectively. Orange: Phenol/Chloroform DNA and RNA methods. Grey: DNeasy and RNeasy PowerSoil Kit (QIAGEN). Blue: ZymoBIOMICS DNA Mini Kit (Zymo Research). Dark blue Maxwell RSC miRNA Tissue Kit (Promega). Pink: Maxwell RSC simplyRNA Tissue Kit (Promega). n=3.

DNA extraction methods comparison at an equal sequencing depth

At an equal sequencing depth, genera and functional classes shared by all methods as well as those detected by only one method for each of the two soils were represented using Venn Diagrams (Figure 4). The proportion of genera detected by a single method represented 8.29% of the Scottish Agricultural College soil pool and 12.84% of the pool from La Côte de Saint André soil. For functional classes, the ones measured by a single method represented 20.22% of the Scottish Agricultural College pool and 12.11% of La Côte de Saint André pool. The average relative abundance of the 25 most abundant genera and functional classes and their relative abundances in each method subsample were inferred from the rarefied pool (Tables S1, S2, S3 and S4 in Supplementary Material). For taxonomic genera, most values fit into the average plus/minus one standard deviation for the Maxwell 1 and phenol/chloroform methods and the DNeasy PowerSoil Kit. The Maxwell 2 method globally showed higher relative abundances than the average plus one standard deviation in both soils, whereas the relative abundances of the most abundant genera were often lower than the average minus one standard deviation in the ZymoBIOMICS DNA Mini Kit rarefied pools. The most abundant functional classes were more likely to be between the average plus/minus one standard deviation than the genus distribution. Most of the function values outside one standard deviation from the mean belonged to the Maxwell 2 method for both soils.

Complementary analysis had shown that only a small proportion of annotated genera and functional classes are relatively abundant in the analysed ecosystems. Most of the annotated genera and functional classes were present at low abundance, with less than 30 associated sequences each (Figures S1 and S2 in Supplementary Material). Therefore, random subsampling of low abundant genera and functional classes could generate a loss of information, detecting artificially unique genera and functional classes. To determine whether the genera and functional classes classified as unique were a product of random rarefaction, their distribution between the different methods before rarefaction was determined. A genus or functional class was validated as unique when it was detected by only one method both before and after rarefaction. Otherwise, it was classified as a false assignation due to random rarefaction (Tables 3 and 4). The percentage of false assignation of unique genera or functional classes and its correlation to the number of sequences before rarefaction were assessed. Regarding genera, the percentage of false assignment ranged between 33% and 87% and it showed no correlation to the sample size before rarefaction with R^2 values of 0.31 and 0.003 in the Scottish Agricultural College soil and La Côte de Saint André soil, respectively (Tables 3 and 4). On the other hand, there was a negative correlation between the percentage of false assignment of functional classes and the sample size before rarefaction, with R^2 values of 0.99 and 0.96 in the Scottish Agricultural College soil and La Côte de Saint André soil, respectively (Tables 3 and 4).

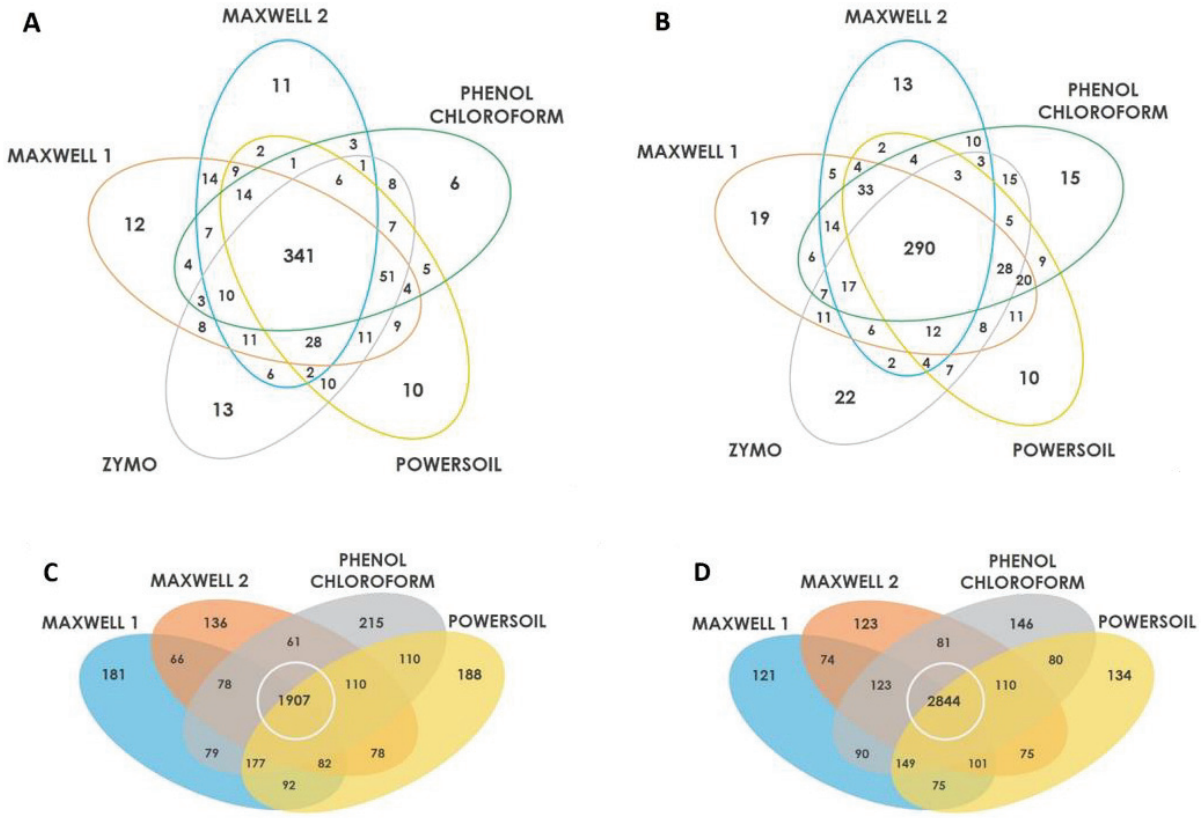


Figure 4. Venn Diagrams representing shared and unique (A) genera from DNA extracted from the Scottish Agricultural College soil; (B) genera from DNA extracted from La Côte de Saint André soil; (C) functional classes from DNA extracted from the Scottish Agricultural College soil; (D) functional classes from DNA extracted from La Côte de Saint André soil. Maxwell 1 and 2: modifications of the prototype Maxwell Fecal Microbiome Kit (Promega). POWERSOIL: DNeasy PowerSoil Kit (QIAGEN); ZYMO: ZymoBIOMICS DNA Mini Kit (Zymo Research). Less than 10000 reads were obtained from ZymoBIOMICS DNA Mini Kit (Zymo Research) triplicate pools after metagenomics sequencing and annotation. Therefore, they were excluded from functional richness analysis. Sequences from triplicates for each method/soil couple were pooled and singletons removed. Then, a subsample of each method pool was randomly rarefied using the vegan package in R at the lowest size of the cohort (58645 sequences annotated as genera for the Scottish Agricultural College soil; 36311 sequences annotated as genera for La Côte de Saint André soil; 32693 sequences annotated as functions for the Scottish Agricultural College soil; 159748 sequences annotated as functions for La Côte de Saint André soil). Venn Diagrams were obtained using the R package VennDiagram.

Soil	Method	Sequences before rarefaction	Detected unique genera	Validated unique genera	% False assignation of unique genera
SCOTTISH AGRICULTURAL COLLEGE	Maxwell 1	80628	12	8	33
	Maxwell 2	58645	11	4	64
	Phenol Chloroform	64805	6	3	50
	PowerSoil	76270	10	3	70
	Zymo	82187	13	8	38
	TOTAL	**362535**	**52**	**26**	**50**
	R^2 between sequences before rarefaction and % false assignation of unique genera = 0.31				
LA COTE DE SAINT ANDRE	Maxwell 1	61656	19	3	84
	Maxwell 2	70676	13	3	77
	Phenol Chloroform	77866	15	2	87
	PowerSoil	108447	10	3	70
	Zymo	36311	22	7	68
	TOTAL	**354956**	**79**	**18**	**77**
	R^2 between sequences before rarefaction and % false assignation of unique genera = 0.003				

Table 3. Percentage of false assignation of unique genera due to sample size normalization and correlation to sample size before rarefaction. Maxwell 1 and 2: modifications of the prototype Maxwell Fecal Microbiome Kit (Promega). PowerSoil: DNeasy PowerSoil Kit (QIAGEN); Zymo: ZymoBIOMICS DNA Mini Kit (Zymo Research). Detected unique genera: genera identified in a single method after sample rarefaction at the lowest size of the cohort (58645 sequences for the Scottish Agricultural College soil and 36311 sequences for La Côte de Saint André soil). Validated unique genera: genera identified in a single method both before and after rarefaction. False assignation of unique genera: genera detected as unique after rarefaction but detected by more than one method before rarefaction. All data (averages and standard deviations) are based on three separate soil samples for each method.

Finally, the number of sequences associated with each validated unique genus and functional class before sample size normalization was inferred. Most genera detected by only one method had less than six annotated sequences before rarefaction (Tables S5 and S6 in Supplementary Material). The only exception was *Staphylococcus,* which was detected by the ZymoBIOMICS DNA Mini Kit and had 95 and 130 annotated sequences before rarefaction in the Scottish Agricultural College soil and La Côte de Saint André soil, respectively. Regarding functions, the same tendency was generally observed with less than six sequences annotated per unique function before size normalization. However, in the Scottish Agricultural College soil, more than six sequences were associated to half of the functional classes detected only by the DNeasy PowerSoil Kit, the largest pool of the cohort before rarefaction (Table S7 in Supplementary Material). Thus, size normalization did not appear to account for the sequencing depth effect on bacterial richness discovery.

Soil	Method	Sequences before rarefaction	Detected unique functional classes	Validated unique functional classes	% False assignation of unique functional classes
SCOTTISH AGRICULTURAL COLLEGE	Maxwell 1	39204	181	11	94
	Maxwell 2	32693	136	12	91
	Phenol Chloroform	104693	215	21	90
	PowerSoil	644248	188	118	37
	TOTAL	**820838**	**720**	**162**	**78**
	R^2 between sequences before rarefaction and % false assignation of unique functions = 0.99				
LA COTE DE SAINT ANDRE	Maxwell 1	159748	121	39	68
	Maxwell 2	234007	123	54	56
	Phenol Chloroform	566982	146	95	35
	PowerSoil	296570	134	57	57
	TOTAL	**1257307**	**524**	**245**	**53**
	R^2 between sequences before rarefaction and % false assignation of unique functions = 0.96				

Table 4. Percentage of false assignation of unique functions due to sample size normalization and correlation to sample size before rarefaction. Maxwell 1 and 2: modifications of the prototype Maxwell Fecal Microbiome Kit (Promega). PowerSoil: DNeasy PowerSoil Kit (QIAGEN). Less than 10000 reads were obtained from ZymoBIOMICS DNA Mini Kit (Zymo Research) triplicate pools after metagenomics sequencing and annotation. Therefore, they were excluded from functional richness analysis. Detected unique functional classes: functional classes identified in a single method after sample rarefaction at the lowest size of the cohort (32693 sequences for the Scottish Agricultural College soil and 159748 sequences for La Côte de Saint André soil). Validated unique functional classes: functional classes identified in a single method both before and after rarefaction. False assignation of unique functional classes: functional classes detected as unique after rarefaction but detected by more than a method before rarefaction. All data (averages and standard deviations) are based on three separate soil samples for each method.

Sequencing depth effect on ARG richness discovery

In order to evaluate the relative contribution of DNA extraction methods and sequencing depth on ARG abundance and ARG richness discovery, the ARG richness and the number of ARG reads detected in each DNA sample were plotted as a function of sequencing depth (Figure 5). Less than 10000 reads were obtained from ZymoBIOMICS DNA Mini Kit (Zymo Research) triplicate pools after metagenomics sequencing and annotation. Therefore, they were excluded from ARG richness analysis. The maximum richness measured from soil metagenomic DNA represented roughly 10% of the references found in the CARD database (320 ARG detected in soil versus 3057 reference sequences in the CARD database) and a plateau was observed between 300 and 320 ARG regardless of the extraction method or soil type (Figure 5, top left). The ARG richness detected by the different extraction methods showed some important variability between samples, except for the triplicates sequenced from DNA extracted from La Côte de Saint André soil using the method Maxwell 1, which detected between 250 and 280 ARGs. In fact, similar or higher ARG richness was measured from these samples than from samples sequenced at a higher sequencing depth. Regarding ARG reads annotation, more reads were detected in some samples than in others and this was not necessarily due to a higher sequencing depth (Figure 5, top right). In addition, although the number of ARG reads and the ARG richness obtained from each sample was not obviously related to sequencing depth, measured ARG richness was clearly dependent on the number of annotated ARG reads (Figure 5, bottom). Thus, differences in measured ARG richness seemed to be related to variability between sample resistome content, DNA extraction methods and, to a lesser extent,sequencing depth.

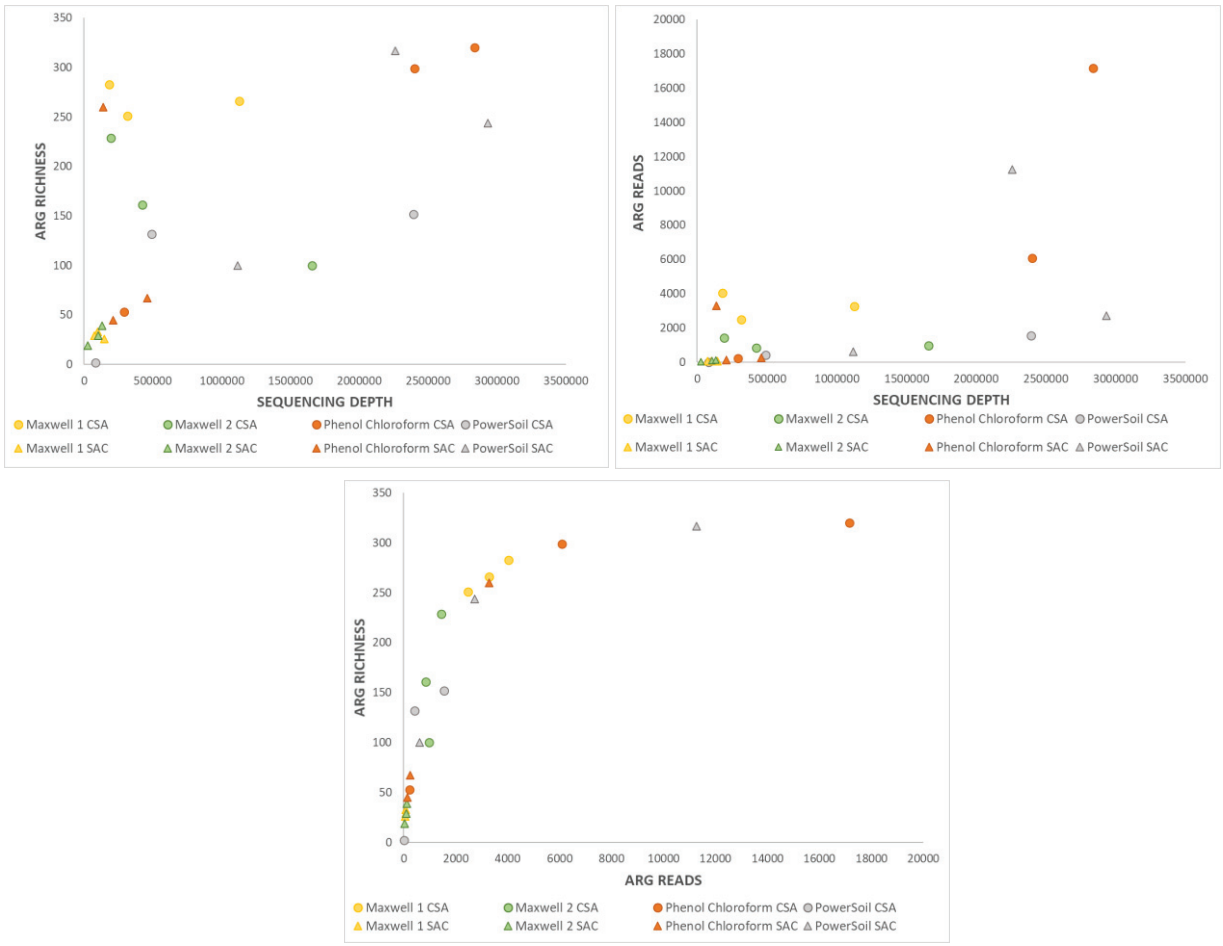


Figure 5. Effect of sequencing depth on ARG richness discovery (top left) and ARG reads detection (top right); effect of ARG read detection on ARG richness discovery (bottom). Maxwell 1 and 2: modifications of the prototype Maxwell Fecal Microbiome Kit (Promega). PowerSoil: DNeasy PowerSoil Kit (QIAGEN). Triangles: DNA from the Scottish Agricultural College soil. Circles: DNA from La Côte de Saint André soil. Yellow and green: Promega's Maxwell 1 and 2 methods, respectively. Orange: Phenol/Chloroform method. Grey: DNeasy PowerSoil Kit (QIAGEN). For each method/soil pair, triplicates were performed and plotted in the graph.

Discussion

In this study, we compared the bacterial richness detected by different DNA and RNA extraction methods. Although some methods detected a higher ARG richness than others sequenced at a higher depth (Figure 5), greater sequencing depth increased taxonomic and functional richness discovery regardless of the method used for DNA or RNA extraction (Figures 1 and 2). In other words, at a cursory level, all methods detected a similar taxonomical and functional richness at equal sequencing depths. In addition, higher differences in bacterial community composition were found between soil samples than between DNA or RNA extraction methods (Figure 3). Thus, sequencing depth has a larger influence on taxonomic and functional richness discovery and on the observed composition of bacterial communities than DNA or RNA extraction methods, whereas some methods arguably detected a higher ARG richness than others did. However, variability between triplicates in terms of ARG content also plays an important role on ARG richness discovery and a plateau of 320 ARGs was reached regardless of soil type or DNA extraction method, suggesting that ARG richness in these two soil types is similar and represents around 10% of the ARG richness identified in the CARD database. Previous studies have shown that sequencing depth has an impact on bacterial richness and diversity discovery [21–23] but, unfortunately, they were not coupled to

extraction method variants since these also have been considered as critical for bacterial richness and diversity [15–18]. Thus, this is the first study (to the best of our knowledge) evaluating both effects simultaneously and demonstrating that the impact of sequencing depth on soil taxonomic and functional richness discovery is greater than that of DNA and RNA extraction methods, while this tendency is less obvious when it comes to ARG richness discovery.

The main implication of the observed unequal sampling depths and their effect on richness discovery was the need for normalizing sample sizes in order to be able to perform comparisons between DNA extraction methods. One of the most popular approaches for sample normalization is the use of relative abundance [6] that have often been selected as a criteria to compare the performance of DNA extraction methods. However, since soil DNA complete diversity has not been extracted and sequenced yet[36], the actual relative abundances of soil microbiome cannot be known and the obtained profiles cannot be validated. In addition, relative abundances are potentially influenced by sequencing depth, since the probability of detecting rare OTUs or functions increases with the number of sequences and, therefore, affects the relative proportions of high (and not unique) abundant versus low abundant OTUs or functions. Finally, differences in sample lysis and DNA purification may affect the relative abundances of the detected communities without modifying measured richness. In other words, the use of different DNA extraction methods might detect different proportions of the same communities rather than different OTUs or functions. Altogether, the informative potential of relative abundance measurements appears to be considerably limited and the performance of DNA extraction methods should be compared using absolute values.

Thus, after pooling the DNA from the three samples, we normalized sample size by randomly rarefying every method pool at the lowest measured sequencing depth. Only a few of the genera and functional classes present in the non-rarefied pools were abundant (see Figures S1 and S2 in Supplementary Material) as both genera and function pools were mainly composed by low-abundant elements. This pattern of taxonomical distribution has already been observed in a study comparing soils from 237 different locations, where only 2% of the ensemble of bacterial phylotypes were found to be dominant [37]. For functions, this distribution is consistent with a few functions being shared between different taxa and implicated in common bacterial ecology processes, while a large pool of low abundant genes confers functions specific to single taxonomical groups. Given this distribution of genera and functions, random sub-sampling of pools mainly composed by low-abundant sequences could result in the selection of false positives (falsely considered unique to the soil or the extraction method). Indeed, our results show that random rarefaction results in the false identification as unique of genera and functional classes that were detected by several methods before rarefaction (Tables 2 and 3).

Furthermore, there was a negative correlation between sample size before rarefaction and the percentage of false identification of unique functional classes, suggesting that size normalization of unequal pools is not enough to avoid the effect of sequencing depth and perform accurate comparisons between DNA extraction methods. The number of sequences associated to the OTUs and functional classes that were detected by a single method both before and after rarefaction was relatively low, which raises questions about their real “uniqueness” in extracted DNA samples. Only *Staphylococcus* discovery by the ZymoBIOMICS DNA Mini Kit from both soils had enough annotated sequences before rarefaction to provide a certain confidence on the uniqueness of the genera. Nevertheless, the possibility that *Staphylococcus* was present in DNA eluates extracted using other methods and was overlooked during any of the steps following DNA extraction cannot be ruled out. Given that sequencing itself is a subsampling of extracted DNA, the differences detected between DNA extraction methods could be due to any other step of the soil metagenomics workflow, from sampling to sequence annotation, rather than DNA extraction. The higher numbers of sequences associated to unique functional classes are only observed in the largest pool of the cohort and, thus, support this hypothesis.

Although the effects of sequencing depth on bacterial discovery have been observed before, to the best of our knowledge, this is the first study comparing both nucleic acid extraction and sequencing depth on bacterial richness discovery. This comparison can help determine which investments are more urgently needed to improve the metagenomics workflow. Our results show that sequencing depth has a more determinant effect on the bacterial richness measured from the same microbial community than the method chosen for DNA extraction. Furthermore, the different DNA extraction "triplicates" that had been sequenced at different depths had access to a different proportion of bacterial DNA richness. Similar results have been reported in a study analysing soil fungi, where higher levels of dissimilarity between replicates were found at low sequencing depths [22]. Therefore, differences found between DNA extraction methods are probably due to DNA subsampling for sequencing instead of to the choice of a DNA extraction method, and we are currently unable to determine which of these differences account for DNA extraction technique variability. For RNA extraction, the same tendency was observed for taxonomic richness discovery, supporting the hypothesis that sequencing depth has a higher impact than nucleic acid extraction on bacterial richness discovery. However, the relative contribution of RNA extraction methods and sequencing depth on functional richness discovery were not evaluated in this study. This kind of study would be of interest, since functional richness discovery in metatranscriptomic studies is critical to analyse the potential activity of soil microbial communities.

Since samples were sequenced at unequal depths, comparisons between DNA extraction methods needed a size normalization that inevitably caused a loss of information and the selection of false positives due to random subsampling. On the other hand, the effect of sequencing depth is still observed in functional analysis after rarefaction, thus, size normalization is not sufficient to compensate for the effect of sequencing depth on bacterial richness discovery. Whereas DNA extraction methods may affect sequencing depth, other factors, such as DNA amplification, library preparation, sequencing techniques and sequence annotation may also contribute to this inequality and lead to inaccurate comparisons between methods. Therefore, efforts should be made to optimize each of these steps in order to sequence representative samples of extracted DNA at a sufficient and equal depth. This would not only facilitate the accurate comparison between DNA extraction methods but would also help define standard methods for soil metagenomics that would improve metagenomic comparison and eventually lead to accurate profiles of soil microbiomes.

Finally, the main goal of this methodological study was to determine which method was the best choice for resistome analysis in soil. Given that sequencing depth showed a stronger effect on taxonomic and the functional richness discovery than ARG discovery, this decision was made based on ARG richness discovery. Although some samples extracted with other methods showed a higher ARG richness than triplicates from La Côte de Saint André extracted using the Maxwell 1 method, the Maxwell 1 showed a more consistent ARG richness discovery than other methods. Therefore, this method could arguably be more adequate to resistome richness discovery in soil and was selected for the next study.

References

1. Roesch, L. F. W. *et al.* Pyrosequencing enumerates and contrasts soil microbial diversity. *ISME J.* **1,** 283–290 (2007).

2. Rappé, M. S. & Giovannoni, S. J. The Uncultured Microbial Majority. *Annu. Rev. Microbiol.* **57,** 369–394 (2003).

3. Hugenholtz, P., Goebel, B. M. & Pace, N. R. Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity. *J. Bacteriol.* **180,** 4765–4774 (1998).

4. Cruaud, P. *et al.* Influence of DNA extraction method, 16S rRNA targeted hypervariable regions, and sample origin on microbial diversity detected by 454 pyrosequencing in marine chemosynthetic ecosystems. *Appl. Environ. Microbiol.* **80,** 4626–4639 (2014).

5. Klindworth, A. *et al.* Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res.* **41,** 1–11 (2013).

6. Hugerth, L. W. & Andersson, A. F. Analysing microbial community composition through amplicon sequencing: From sampling to hypothesis testing. *Front. Microbiol.* **8,** 1–22 (2017).

7. Tremblay, J. *et al.* Primer and platform effects on 16S rRNA tag sequencing. *Front. Microbiol.* **6,** 1–15 (2015).

8. Lombard, N., Prestat, E., van Elsas, J. D. & Simonet, P. Soil-specific limitations for access and analysis of soil microbial communities by metagenomics. *FEMS Microbiol. Ecol.* **78,** 31–49 (2011).

9. Robe, P., Nalin, R., Capellano, C., Vogel, T. M. & Simonet, P. Extraction of DNA from soil. *Eur. J. Soil Biol.* **39,** 183–190 (2003).

10. Levy-Booth, D. J. *et al.* Cycling of extracellular DNA in the soil environment. *Soil Biol. Biochem.* **39,** 2977–2991 (2007).

11. Paget, E., Monrozier, L. J. & Simonet, P. Adsorption of DNA on Clay-Minerals - Protection against DNaseI and Influence on Gene-Transfer. *FEMS Microbiol. Lett.* (1992). doi:DOI 10.1111/j.1574-6968.1992.tb05435.x

12. Frostegård, Å. *et al.* Quantification of bias related to the extraction of DNA directly from soils. *Appl. Environ. Microbiol.* **65,** 5409–5420 (1999).

13. Tebbe, C. C. & Vahjen, W. Interference of humic acids and DNA extracted directly from soil in detection and transformation of recombinant DNA from bacteria and a yeast. *Appl. Environ. Microbiol.* **59,** 2657–2665 (1993).

14. Nacke, H. *et al.* Estimates of Soil Bacterial Ribosome Content and Diversity Are Employed. *Appl. Environ. Microbiol.* **82,** 2595–2607 (2016).

15. Kauffmann, I. M., Schmitt, J. & Schmid, R. D. DNA isolation from soil samples for cloning in different hosts. *Appl. Microbiol. Biotechnol.* **64,** 665–670 (2004).

16. Miller, D. N., Bryant, J. E., Madsen, E. L. & Ghiorse, W. C. Evaluation and optimization of DNA extraction and purification procedures for soil and sediment samples. *Appl. Environ. Microbiol.* **65,** 4715–4724 (1999).

17. Töwe, S. *et al.* Improved protocol for the simultaneous extraction and column-based separation of DNA and RNA from different soils. *J. Microbiol. Methods* **84,** 406–412 (2011).

18. Wang, Y., Morimoto, S., Ogawa, N., Oomori, T. & Fujii, T. An improved method to extract RNA from soil with efficient removal of humic acids. *J. Appl. Microbiol.* **107,** 1168–1177 (2009).

19. Roose-Amsaleg, C. L., Garnier-Sillam, E. & Harry, M. Extraction and Purification of Microbial DNA from Soil and Sediment Samples. *Appl. Soil Ecol.* **18,** 47–60 (2001).

20. Marketa Sagova-Mareckova, Ladislav Cermak, Jitka Novotna, Kamila Plhackova, Jana Forstova, J. K.

Innovative Methods for Soil DNA Purification Tested in Soils with Widely Differing Characteristics. *Appl. Environ. Microbiol.* **74,** 2902–2907 (2008).

21. Zaheer, R. *et al.* Impact of sequencing depth on the characterization of the microbiome and resistome. *Sci. Rep.* **8,** 1–11 (2018).

22. Smith, D. P. & Peay, K. G. Sequence depth, not PCR replication, improves ecological inference from next generation DNA sequencing. *PLoS One* **9,** (2014).

23. Lundin, D. *et al.* Which sequencing depth is sufficient to describe patterns in bacterial α- and β-diversity? *Environ. Microbiol. Rep.* **4,** 367–372 (2012).

24. Griffiths, R. I., Whiteley, A. S., O'Donnell, A. G. & Bailey, M. J. 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 (2000).

25. Wang, Q., Garrity, G. M., Tiedje, J. M. & Cole, J. R. Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* **73,** 5261–5267 (2007).

26. Huson, D., Mitra, S., Ruscheweyh, H., Weber, N. & Schuster, S. C. Integrative analysis of environmental sequences using MEGAN4. *Genome Res.* **21,** 1552–1560 (2011).

27. Alcock, B. P. *et al.* CARD 2020: Antibiotic resistome surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids Res.* **48,** D517–D525 (2020).

28. Benjamin Buchfink, Xie, C. & Huson, D. H. Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* **12,** 59–60 (2015).

29. Kemp, J. S., Paterson, E., Gammack, S. M., Cresser, M. S. & Killham, K. Leaching of genetically modified Pseudomonas fluorescens through organic soils: Influence of temperature, soil pH, and roots. *Biol. Fertil. Soils* **13,** 218–224 (1992).

30. Muyzer, G., Hottentrager, S., Teske, A. & Wawer, C. Denaturing gradient gel electrophoresis of PCR-amplified 16S rDNA—a new molecular approach to analyse the genetic diversity of mixed microbial communities. In *Molecular microbial ecology manual;* (eds. Akkermans, A., van Elsas, J. & de Bruijn, F.) 1–23 (Kluwer Academic Publishers, 1995).

31. Watanabe, K., Kodama, Y. & Harayama, S. Design and evaluation of PCR primers to amplify bacterial 16S ribosomal DNA fragments used for community fingerprinting. *J. Microbiol. Methods* **44,** 253–262 (2001).

32. Masella, A. P., Bartram, A. K., Truszkowski, J. M., Brown, D. G. & Neufeld, J. D. PANDAseq: Paired-end assembler for illumina sequences. *BMC Bioinformatics* **13,** 31 (2012).

33. Oksanen, J. *et al.* vegan: Community Ecology Package. R package version 2.4-2. *Community Ecol. Packag.* **2.5-6,** 1–296 (2019).

34. Chenn, H. Generate High-Resolution Venn and Euler Plots. VennDiagram package. *R Packag.* 33 (2018).

35. Edgar, R. C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26,** 2460–2461 (2010).

36. Delmont, T. O., Simonet, P. & Vogel, T. M. Describing microbial communities and performing global comparisons in the omic era. *ISME J.* **6,** 1625–1628 (2012).

37. Delgado-Baquerizo, M. *et al.* A global atlas of the dominant bacteria found in soil. *Science (80-.).* **359,** 320–325 (2018).

Chapter III. Gentamicin adsorption onto soil particles prevents it from having overall short-term effects on the soil microbiome and resistome

Abstract

Antibiotics used in agriculture may reach the surrounding environment, impose a selective pressure on soil bacterial communities and stimulate the development and dissemination of antibiotic resistance in the soil microbiome. On the other hand, the adsorption of antibiotics onto soil particles can buffer the effects of antibiotic pollution on soil microbial communities. The scope of antibiotic pollution on resistance development and ARG dissemination in soil and the link to soil properties needs to be elucidated. This study compared the short-term effects of a range of gentamicin concentrations on the microbiome and resistome of bacterial enrichments and microcosms of an agricultural soil using a metagenomics approach. We hypothesized that gentamicin pollution has a greater influence on soil bacteria in enriched media than in microcosms, where gentamicin adsorption onto soil particles reduces its bioavailability. Gentamicin impact on bacterial biomass was roughly estimated by the number of 16S rRNA gene copies. In addition, soil microbiome and resistome response to gentamicin pollution was evaluated by 16S rRNA gene and metagenomic sequencing, respectively. Finally, gentamicin bioavailability in soil was determined by HPLC-MS/MS. While gentamicin pollution at the scale of µg/g strongly influenced the bacterial communities in soil enrichments, concentrations up to 1 mg/g were strongly adsorbed onto soil particles and did not cause significant changes in the microbiome and resistome of soil microcosms. This study demonstrates the differences between the response of bacterial communities to antibiotic pollution in enriched media and in their environmental matrix and exposes the limitations of culture-based studies in antibiotic resistance surveillance. Furthermore, establishing links between the effects of antibiotic pollution and soil properties is needed.

Introduction

Among all the ecosystems present on Earth, soil harbors the highest microbial diversity[1] and is likely the biggest reservoir of antibiotics. Most of the antibiotics currently used in human therapy and food production have been isolated from soil bacteria and fungi[2]. A natural consequence of this production of antibiotics by soil bacteria is the development of antibiotic resistance in soil. Soil is considered as one of the main environmental reservoirs of antibiotic resistance genes (ARGs)[3]. Both clinically relevant and novel ARGs have been identified even in low-anthropogenically impacted soils[4–7], showing that antibiotic resistance occurs in soil even in the absence of a strong anthropogenic selective pressure. Furthermore, antibiotic resistance genes present in soil are often associated with mobile genetic elements (MGEs)[8] and, therefore, can be transferred to both other environmental bacteria and human pathogens[9].

The use of antibiotics in agriculture and the application of manure from antibiotic-treated animals for soil fertilization increase antibiotic selective pressure and ARG-containing microorganisms in the environment[10,11]. This selective pressure may enhance the development of antibiotic resistance in soil and its mobilization and transfer to clinically relevant bacteria. Therefore, the risks associated with the use of antibiotics in agriculture must be evaluated in order to regulate their use. Multiple studies over the last two decades have analyzed the effects of antibiotic-polluted manure composting on the soil microbiome. They have shown that this manure can increase the abundance of antibiotic resistant bacteria (ARB), ARGs and MGEs[12–17]. However, soil is a highly complex and diverse matrix that changes over time[18]. Several factors, such as soil characteristics (*i.e.* water content, oxygen concentration or nutrient availability)[19,20], the percentage of reduced bioavailability of antibiotics due to their

adsorption onto soil[21,22] and the activity and resilience of soil bacteria[23,24] may alter the effects of antibiotics in soil. In addition, soil is a solid matrix and the physical contact between bacteria and antibiotic compound is reduced in comparison to liquid environments. These differences may inhibit accurate evaluation of the horizontal transfer of ARGs between resistant and susceptible bacteria in soil. Soil characteristics can provide a strong buffering capacity for antibiotic pollution[25] by reducing the impact of antibiotics on the soil microbiome and the potential consequences of soil pollution with antibiotics on human health. Thus, a global picture of the effects of antibiotic pollution on resistance development in soil is difficult to determine and after decades of research the scope of this phenomenon still remains unclear[26].

This study compared the response of soil bacteria to antibiotic pollution in soil microcosms and enrichments in selective media. Gentamicin is an aminoglycoside used in both human therapy and food production and genes conferring resistance to gentamicin are widely distributed in the environment[27]. A range of concentrations of gentamicin from 1 µg/g of soil up to 1 mg/g of soil were added to an agricultural soil with no previous known exposure to gentamicin to evaluate their effects on the soil microbiome and resistome. Bacterial enrichments were contaminated with gentamicin up to 12 µg/ml. The main hypothesis was that gentamicin would be strongly adsorbed onto soil particles and that gentamicin pollution has a higher effect on soil bacteria in enriched media than in soil microcosms, since it would be more bioavailable and bacteria in enriched media are more responsive to gentamicin given their higher activity, lower diversity, and the higher availability of nutrients. This work demonstrates the differences in the microbiome and resistome response to gentamicin pollution between soil microcoms and bacterial enrichments and exposes the limitation of culture-based studies in antibiotic resistance surveillance in terrestrial ecosystems.

Materials and methods

Soil sampling

Soil was sampled from a plowed corn field at La Côte de Saint André, France (45,38N - 5,26E) on January 2018, on which manure from farm animals treated with cefalexin, neomycin, cefalonium, tetracycline, oxytetracycline, tylosin and sulfamidine was applied. This soil had no previous exposure to gentamicin. Soil characteristics are described in Table 1. Ninety-six sampling points were randomly selected within the field and soil was shoveled at 20 cm depth. Five kg of sample were mixed together and kept at 4°C until the start of the experiments.

	La Côte de Saint André
Sand	42.9%
Silt	43.6%
Clay	13.5%
pH	7.24
Organic matter	2.92 %
Organic C	1.7 %
Total N	0.17 %

Table 1. Physical characterization of La Côte de Saint André soil (France).

Determination of gentamicin effect on soil enrichments

First, in order to determine whether gentamicin had an effect on the enriched fraction of soil bacteria in liquid enrichments and which gentamicin concentrations were sub-inhibitory at a cursory level, soil bacteria were extracted in 0.9% NaCl solution, and 150 µl of 1:10 Tryptic Soy Broth (TSB) medium containing soil bacteria without antibiotics or with gentamicin at 0, 0.01, 0.05, 0.1, 0.5 or 11 µg/ml were transferred to a 96-well culture plate and incubated for 24 hours at 29°C under continuous

shaking in the MultiSkan GO Plate Reader (Thermo Scientific). Optical Density at 600 nm (OD_{600}) was measured every hour over the 24-hour incubation. Then, ANOVA tests and t-student tests between the OD_{600} measured at each gentamicin concentration after 24-hour incubation and the average between all groups were performed using the ggpubr package in R.

Enrichment of soil bacteria in 1:10 TSB medium polluted with gentamicin

Soil bacteria were extracted in 0.9% NaCl solution. 0.5 ml of extracted bacteria were added to 4.5 ml of 1:10 TSB medium without antibiotics or polluted with gentamicin at 0.1 and 12 µg/ml. Enrichments were incubated at 29°C for 24 hours. Then, 1 ml of soil bacteria enrichments was centrifuged at 2000 x g for 5 minutes. The pellet was resuspended in 100 µl of 10 mg/ml lysozyme and 400 µl of TE buffer. After heating at 37°C for 30 minutes shaking at 800 rpm, the lysate was purified using the Maxwell RSC Instrument and the Maxwell RSC Whole Blood DNA Kit (Promega).

Microcosm experiment

Soil from La Côte de Saint André stored at 4°C was sifted at 4 mm and homogenized, and 100 gram microcosms without vegetation were prepared in polypropylene containers, since gentamicin has shown to be highly adsorbed onto glass[28]. Soil water retention capacity was of 24.2% ± 0.64%. Microcosms were left at room temperature overnight before pollution with 1 µg/g, 100 µg/g or 1 mg/g of gentamicin (Duchera Biochemie). Serial dilutions of gentamicin were made in water, and 1 ml of solution was applied to soil four times intermittently, mixing with a metal bar in between applications. Triplicates were made for each concentration, as well as for non-polluted samples. DNA was extracted after 0, 2, and 8-day incubation at ambient temperature and light without moisture maintenance treatment.

DNA extraction from soil microcosms

DNA was extracted using the Maxwell RSC Instrument (Promega) and a prototype version of the Maxwell Fecal Microbiome Kit (Promega). 250 mg of sample were diluted in 1 ml of Lysis Buffer (Promega) and heated for 5 minutes at 95 °C. Samples underwent bead-beating twice at 5.5 m/s for 30 seconds in Lysis Matrix E tubes (MP Biomedicals) and centrifuged at 10600 x g for 5 minutes. Then, 300 µl of supernatant were added to 300 µl of Binding Buffer (Promega) and loaded into a Maxwell RSC cartridge containing magnetic beads for DNA purification on the Maxwell RSC Instrument, according to the Technical Manual TM473.

DNA quantification and quantitative PCR (qPCR) assays in soil microcosms

DNA concentrations extracted from soil microcosms and bacterial enrichments were assessed using the Qubit Fluorometer and the Qubit dsDNA HS Assay Kit (ThermoFisher). The size of the total bacterial community was estimated by quantifying the V3 region of the 16S rRNA gene by qPCR using the "universal" primers 341F (5'-CCT ACG GGA GGC AGC AG- 3') and 534R (5'-ATT ACC GCG GCT GCT GGC A-3')[29,30]. One of the primer pairs (F: 5'-CATGACCTTGCGATGCTCTATG-3'; R: 5'-TCCAAGAGCAACGTACGACTG-3') designed to in Chapter IV was used to target a 201-bp conserved region of 15 genes belonging to the AAC(6') acetyltransferase family: *aac(6')-Ib, aac(6')-Ib', aac(6')-Ib-cr, aac(6')-Ib3, aac(6')-Ib4, aac(6')-Ib7, aac(6')-Ib8, aac(6')-Ib9, aac(6')-Ib10, aac(6')-Ib11, aac(6')-30/aac(6')-Ib', aac(6')-Ib-Hangzhou, aac(6')-Ib-Suzhou, aac(3)-Ib/aac(6')-IIb", ant(3")-II/aac(6')-IId.* qPCR assays were carried out using the Corbett Rotor-Gene 6000 (QIAGEN) in a 20 µl reaction volume containing GoTaq qPCR Master Mix (Promega), 0.75 µM of each primer and 2 µl of DNA at ≤2.5 ng/µl. Two non-template controls were also included in all the assays. Standard curves for 16S rRNA gene assays were obtained using 10-fold serial dilutions of a linearized plasmid pGEM-T Easy Vector (10^7 to 10^2 copies) containing the 16S rRNA gene of *Pseudomonas aeruginosa* PAO1. Cycling conditions for 16S rRNA gene qPCR amplification were 95 °C for 2 minutes followed by 30 cycles of 95 °C for 15

seconds, 60 °C for 30 seconds and 72 °C for 30 seconds. Standard curves for *aac(6')* gene qPCR assays were obtained from river water DNA, cloned and transformed using the TOPO TA cloning Kit (Invitrogen), linearized and diluted (10^7 – 10^2 copies/µl). Cycling conditions for qPCR amplification were 95 °C for 2 minutes followed by 35 cycles of 95 °C for 30 seconds, 55 °C for 30 seconds and 72 °C for 30 seconds. Melting curves were generated after amplification by increasing the temperature from 60 °C to 95 °C. The number of copies per µl of reaction obtained from the amplification of *aac(6')* genes was normalized by the copies of the 16S rRNA gene per µl of reaction to assess their relative abundance in soil microcosms and bacterial enrichments. Then, ANOVA tests and t-student tests between the DNA concentration, copies of 16S rRNA gene per µl of reaction and relative abundance of *aac(6')* genes measured from each group (each gentamicin concentration after 24-hour exposure for bacterial enrichments and each gentamicin concentration and each exposure time for microcosms) and the average between all groups were carried out using the ggpubr package in R.

16S rRNA gene and cDNA sequencing and analysis

The V3-V4 hypervariable region of bacterial 16S rRNA gene weas amplified from DNA obtained from both enriched bacteria and soil microcosms except for enriched media polluted at 12 µg/ml due to insufficient DNA concentration. DNA was amplified using the Titanium Taq DNA Polymerase (Takara Clontech) forward 341F with Illumina overhang (5'-TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG CCT ACG GGN GGC WGC AG-3') and reverse 785F with Illumina overhang (5'-GTC TCG TGG GCT CGG AGA TGT GTA TAA GAG ACA GGT CTC GTG GGC TCG GAG ATG TGT ATA AGA GAC AGG ACT ACH VGG GTA TCT AAT CC-3') primers[31]. Amplification conditions were as follows: 95 °C for 3 minutes followed by 25 cycles of 95 °C for 30 seconds, 55 °C for 30 seconds and 72 °C for 30 seconds and a final extension step at 72°C for 5 minutes. DNA libraries were prepared from amplified products based on Illumina's "16S Metagenomics Library Prep Guide" (15044223 Rev. B) using the Platinum Taq DNA Polymerase (Invitrogen) and the Nextera XT Index Kit V2 (Illumina). DNA sequencing with a 15% PhiX spike-in was performed using the MiSeq System and the MiSeq Reagent Kit v2 (Illumina). Reads were trimmed to meet a quality score of Q20. Then, pair-ended reads were assembled using PANDAseq[32] at a sequence length between 410 and 500 bp and an overlap length between 20 and 100 bp, using the rdp_mle algorithm. Finally, each of the DNA sequences was annotated to the genus level using the Ribosome Data Project (RDP) database and the RDP Bayesian classifier using an assignment confidence cut-off of 0.6[33]. Three microcosm samples were excluded from further analyses due to insufficient sequencing depth: a triplicate polluted at 1 µg/g of gentamicin after 0-day exposure, a triplicate polluted at 1 µg/g of gentamicin after 2-day exposure and a non-polluted triplicate after 8-day exposure. The genera that had less than 10 associated sequences in the ensemble of sequences from enriched media and soil microcosms were removed. Then, the relative abundances of the remaining genera were calculated and plotted individually. PCoA analyses were performed using R to compare the community composition of soil microcosms and bacterial enrichments. In addition, the genus richness measured in each sample was determined using the vegan package in R[34]. Then, ANOVA tests and t-student tests between the genus richness detected in each group (each gentamicin concentration after 24-hour exposure for bacterial enrichments and each gentamicin concentration and each exposure time for microcosms) and the average between all groups were carried out using the ggpubr package in R.

Metagenomics sequencing and analysis

Metagenomics libraries were prepared from <1 ng of DNA obtained from both enriched bacteria and soil microcosms using the Nextera XT Library Pep Kit and Indexes (Illumina), as detailed in Illumina's "Nextera XT DNA Library Prep Kit" reference guide (15031942 v03). DNA sequencing with a 1% PhiX spike-in was performed using the MiSeq System and the MiSeq Reagent Kit v2 (Illumina). Two approaches were used to evaluate the antibiotic resistome of non-polluted and gentamicin-polluted

soil bacterial enrichments and microcosms using metagenomics sequences. First, sequences from soil bacterial enrichments and soil microcosms were co-assembled in order to generate metagenome assembled genomes (MAGs) and associate possible resistome elements to concrete taxa. Metagenomic reads obtained from soil microcosms or bacterial enrichments were filtered according to the criteria described by Minoche *et al*[35]. Then, reads were co-assembled using MEGAHIT[36] to generate contigs, and reads were mapped onto the contigs using Bowtie 2[37] to generate BAM files. Profiles were created for each individual sample and merged using the anvi'o metagenomic workflow[38]. Samples were binned based on differential coverage and sequence composition. A bin was considered as a MAG when it showed a completion higher than 50% and a redundancy lower than 10%. In order to determine whether the assembled contigs contained ARGs, the merged profile was blasted against the CARD database[39] using Diamond. The obtained results were filtered at a minimum identity of 60% and a minimum length of 33 amino acids and the best-hit was selected. A summary of these assemblies is found in Table S1 in Supplementary Information. Since the assembly represented a very low proportion of the total of sequenced reads from soil microcosms, information regarding soil resistome composition could be lost using this approach. Therefore, the resistome screening was carried out on non-assembled reads from bacterial enrichments and soil microcosms in order to evaluate all the sequences. Reads were trimmed using the Fastq Quality Trimmer tool of the FASTX-Toolkit. Nucleotides that did not meet a minimum quality score of Q20 were trimmed from the sequences, and sequences shorter than 100 nucleotides after trimming were removed. Then, reads from R1 and R2 were concatenated and blasted against the CARD database using Diamond. The obtained results were filtered at a minimum identity of 60% and a minimum length of 33 amino acids, and the best hit was chosen. Genes that had less than 10 copies in the ensemble of samples were removed and the relative abundance of each gene was calculated. The genes that increased their relative abundance over time in non-polluted samples were removed*, i.e.* the genes for those the Pearson coefficient between time and relative abundance in non-polluted microcosms was higher or equal to 0.9. Finally, the genes that did not show a Pearson coefficient between time and relative abundance in polluted samples higher than 0.9 were also removed. The remaining genes were subjected to a Leave-one-out cross-validation in order to determine whether their increase under gentamicin pollution over time was significant.

Gentamicin bioavailability and adsorption in soil

One gram of unpolluted soil or soil polluted at 1 µg/g, 100 µg/g or 1 mg/g of gentamicin was diluted in 10 ml of water to determine the bioavailable fraction, vortexed and stocked at -20°C until analysis. Then, tubes were centrifugated at 10000 rpm for 5 minutes to recover the bioavailable fraction of gentamicin. A 200 µl sample of supernatant was transferred to a polypropylene tube. Then, 200 µl of 75 mM sodium hexanesulfonate and 200 µl of 75 mM sodium heptanesulfonate were added to the tube, and after vortexing for 30 seconds were analyzed by HPLC-MS/MS[40].

Results

Gentamicin effect on soil bacterial communities

The highest tested concentration of gentamicin that could be considered as sub-inhibitory for enriched soil bacteria at a cursory level was 0.1 µg/ml since it did not cause significant overall growth inhibition and the growth curve approached that of the non-polluted soil bacteria enrichments (Figure 1). On the other hand, a delay of 15 hours on the offset of the curve was observed at 0.5 µg/ml of gentamicin and there was no visible growth in soil bacteria enrichments with 1 µg/ml of gentamicin (Figure 1a). In addition, significantly lower optical densities were measured at 600 nm (OD_{600}) in bacterial enrichments with 0.5 and 1 µg/ml of gentamicin when compared to the average of all samples (Figure 1b). Both concentrations were considered, therefore, to be inhibitory for enriched soil bacteria. The inhibitory concentration selected for samples undergoing DNA extraction was scaled up to 12 µg/ml

of gentamicin to ensure that inhibitory effects were observed. DNA was extracted from bacterial liquid enrichments with 0.1 µg/ml of gentamicin (a sub-inhibitory concentration), 12 µg/ml of gentamicin (an inhibitory concentration) and in control enrichments with no added gentamicin.

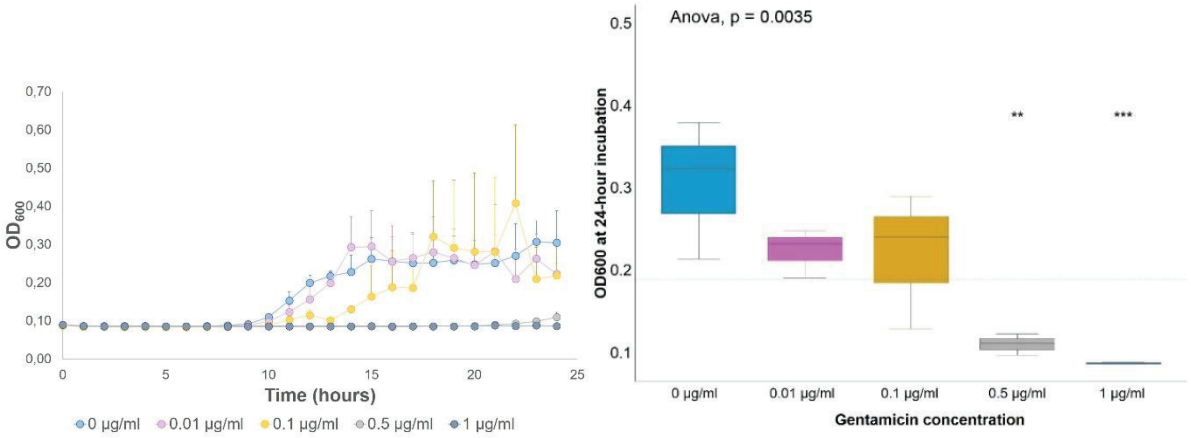


Figure 1. Gentamicin growth inhibition on soil bacterial enrichments. Optical density measured at 600 nm (OD_{600}) of cultivable soil bacteria incubated in 1:10 Tryptic Soy Broth (TSB) with different gentamicin concentrations for every hour during a 24-hour incubation at 29°C (left) and after 24-hour incubation (right). Significant differences between each group (each gentamicin concentration) and the average between all groups (horizontal dashed line) were determined by a t-test. **p-value ≤ 0.01; ***p-value ≤ 0.001. n=3.

Significantly lower DNA concentrations and biomass were detected in soil bacterial enrichments with 12 µg/ml of gentamicin (Figure 2), showing a clear inhibitory effect, whereas gentamicin concentrations of up to 1 mg/g did not show a significant decrease of DNA concentrations nor bacterial biomass in soil microcosms (Figure 3). Non-gentamicin soil microcosms showed a significantly lower number of 16S rRNA gene copies than soil with gentamicin after both a 2-day exposure and an 8-day exposure. Non-gentamicin soil microcosms also showed a significantly lower DNA concentration than soil with gentamicin after an 8-day exposure. The significant differences found between soil microcosms with and without added gentamicin may reflect, rather than a growth stimulation induced by gentamicin, small differences in the community composition between soil triplicates polluted at the same gentamicin concentration and exposure times. The differences in the bacterial size of gentamicin-polluted soil microcosms may also account for handling error during DNA extraction and quantitative-PCR (qPCR) amplification. Thus, no obvious effect of gentamicin on the size of the bacterial community can be observed.

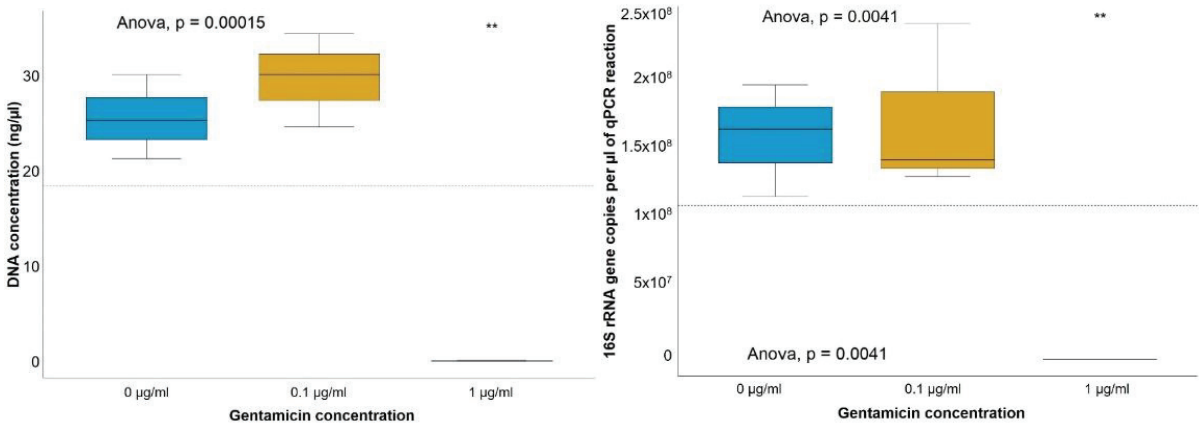


Figure 2. Gentamicin effect on soil bacterial growth in 1:10 TSB medium enrichments after 24-hour exposure. DNA concentrations (left) and number of 16S rRNA gene copies/µl of qPCR reaction (right). Significant differences between each group (each gentamicin concentration) and the average between all groups (horizontal dashed line) were determined by a t-test. **p-value≤0.01. n=3.

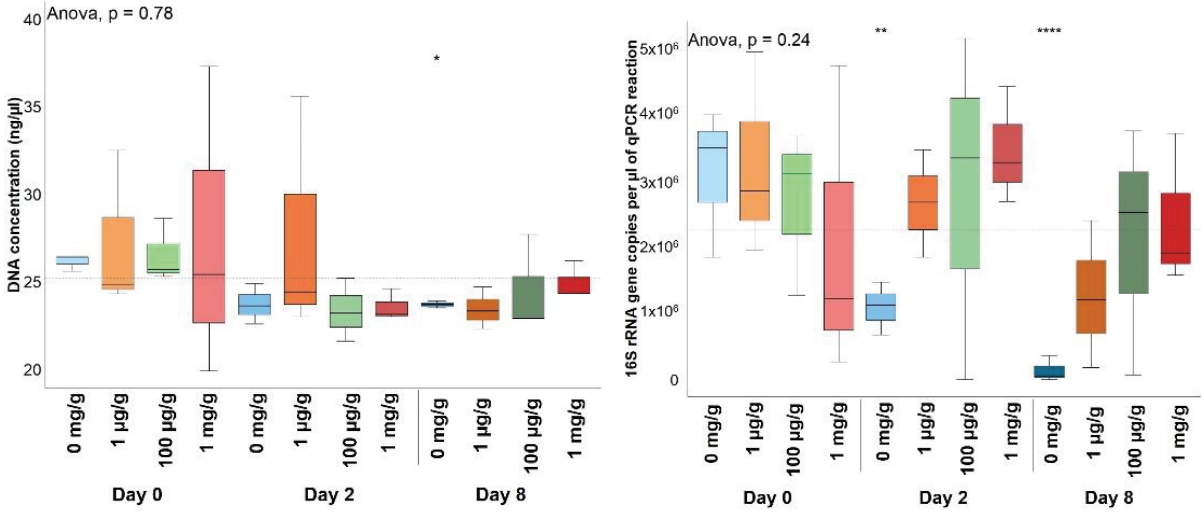


Figure 3. Bacterial dynamics over time in soil microcosms at different gentamicin concentrations. DNA concentrations (left)and number of 16S rRNA gene copies/µl of qPCR reaction obtained from non-polluted soil microcosms or microcosms polluted with gentamicin at 1 µg/g, 100 µg/g or 1 mg/g (right). D0: 0-day exposure; D2: 2-day exposure; D8: 8-day exposure. Significant differences between each group (each concentration and each exposure time) and the average between all groups (horizontal line) were determined by a t-test *p-value ≤0.05. **p-value≤0.01. ****p-value≤0.0001. n=3.

Five genera represented almost 100% of the total communities extracted from soil bacteria enrichments in 1:10 Tryptic Soy Broth (TSB) medium (Figure 4). *Propinimicrobium* was the most prevalent genus, representing between 65 and 90% of the total in non-gentamicin (control) enrichments and almost 100% in enrichments at gentamicin concentrations of 0.1 µg/ml. Enrichments at 12 µg/ml of gentamicin could not be sequenced, since DNA concentrations were very low and the 16S rRNA gene could not be amplified. Whereas the composition of the bacterial communities in non-polluted enrichments varied between triplicates, gentamicin at 0.1 µg/ml reduced those differences (Figure S1a in Supplementary Information). In addition, a significantly lower genus richness was detected in gentamicin contaminated enrichments when compared to non-contaminated controls (Figure S2 in Supplementary Information). These results suggest that gentamicin, even at sub-inhibitory concentrations, inhibits some members of the community in bacterial enrichments.

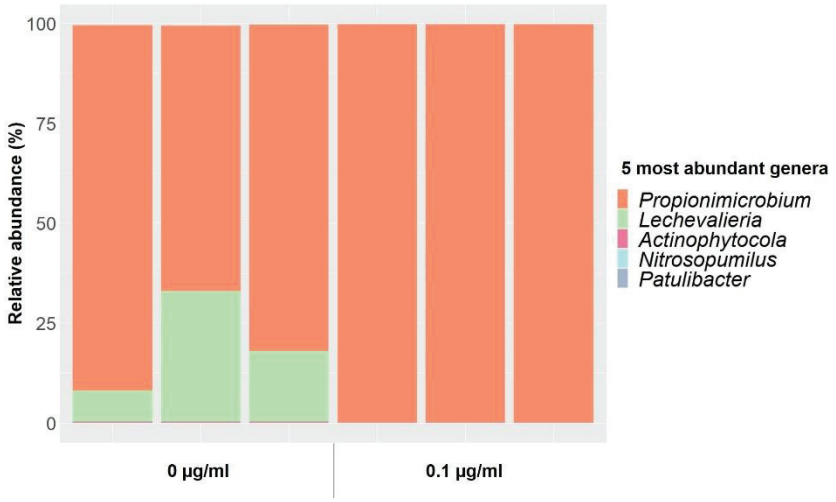


Figure 4. Relative abundances of the five most abundant genera from soil bacteria enrichments in 1:10 TSB medium incubated at 29°C for 24 hours at gentamicin concentrations of 0 and 0.1 ug/mL. Triplicates are plotted individually.

On the other hand, the relative abundance of the 24 most abundant genera in soil microcosms experienced only little changes under gentamicin pollution over time (Figure 5), and both the overall composition of the bacterial communities (Figure S1b in Supplementary Information) and genus richness (Figure S2 in Supplementary Information) remained stable after 8-day exposure. None of the gentamicin concentrations used in this study had an overall inhibitory effect on soil microcosms and even concentrations at mg/g levels showed little effect on the composition of soil bacterial communities after 8-day exposure, while 12 µg/ml of gentamicin had clear inhibitory effects on the culturable fraction of the soil microbiome. This may partly be due to differences in the communities present in soil microcosms and bacterial enrichments (Figure S1c in Supplementary Information). The genus richness measured in bacterial enrichments was significantly lower than that of most of the communities present in soil microcosms (Figure S2 in Supplementary Information). Therefore, the proportion of culturable soil bacteria does not represent the diversity present in soil microcosms.

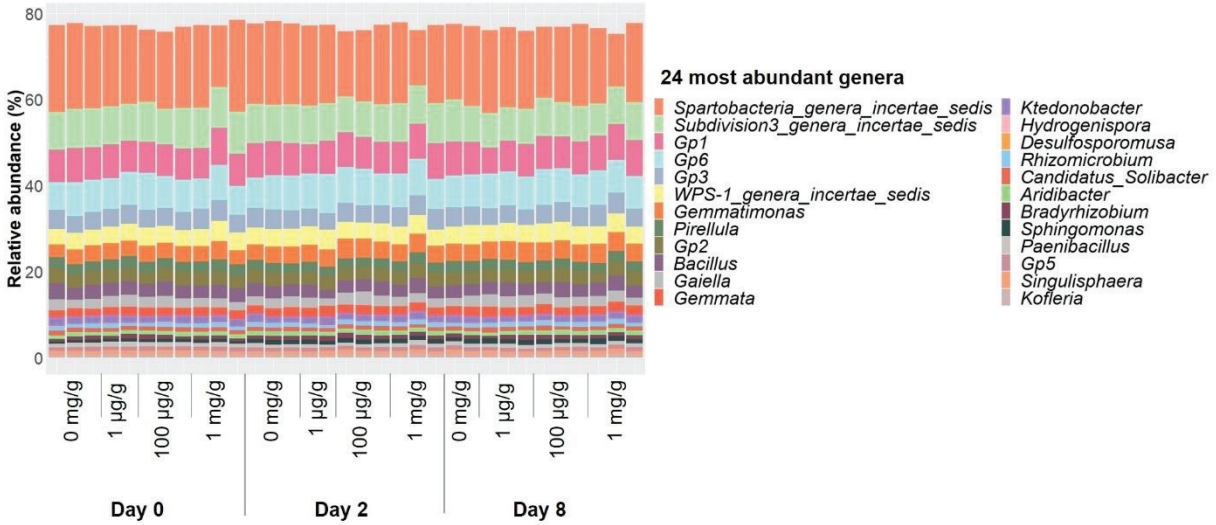
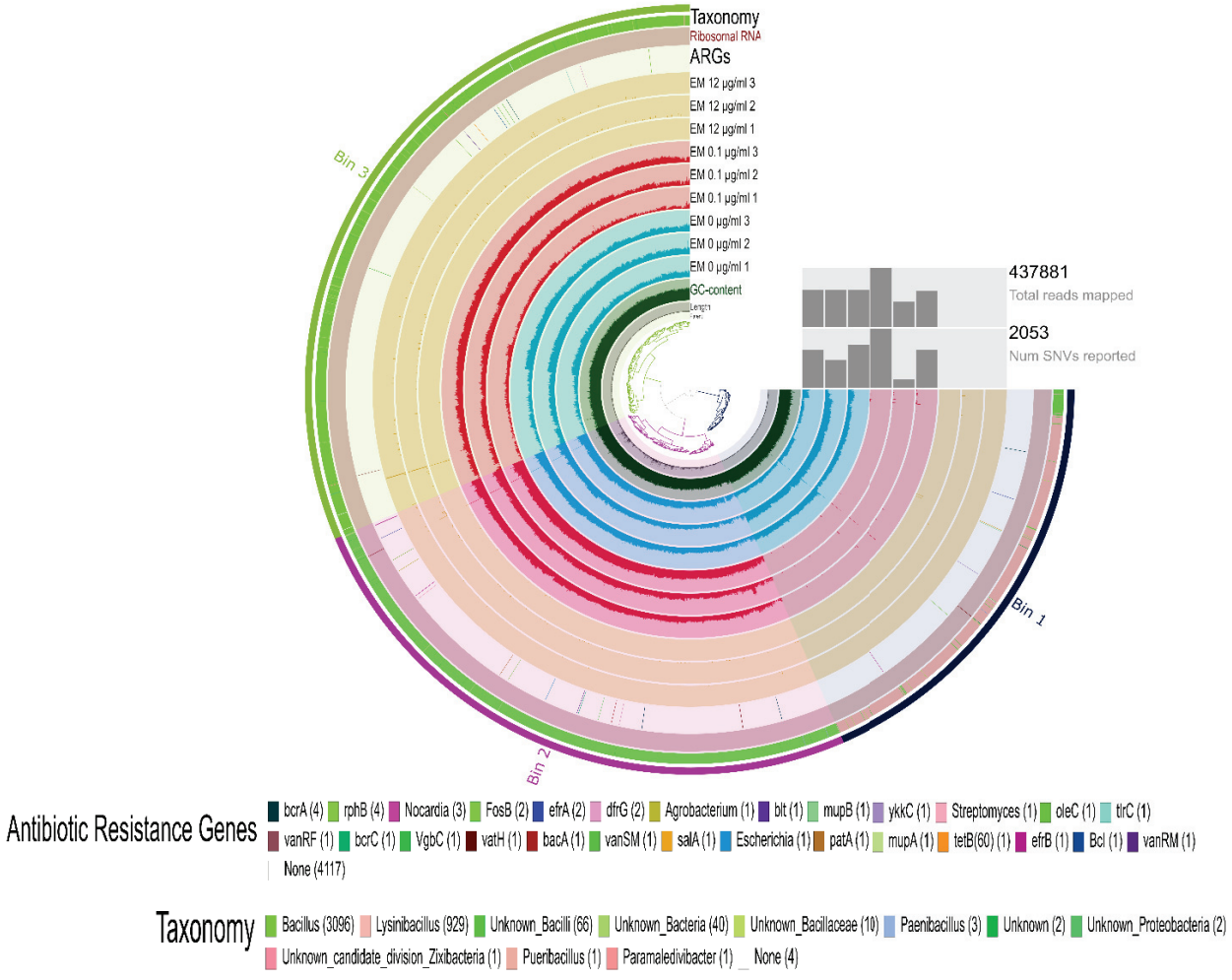


Figure 5. **Relative abundances of the 24 most abundant genera in the total communities from non-polluted soil microcosms or microcosms with 1 µg/g, 100 µg/g or 1 mg/g of gentamicin**. Triplicates are plotted individually.

Gentamicin effect on the soil resistome

Regarding soil bacterial enrichments, samples at 12 µg/ml were generally absent in the metagenomics assembly, while samples from non-polluted enrichments and enrichments polluted at 0.1 µg/ml of gentamicin showed a high alignment rate (Figure 6). Two metagenome assembled genomes (MAGs) were obtained from the assembly. The first one had a completion of 81.69% and a redundancy of 9.86%, and belonged to the genus *Bacillus.* The second one had a completion of 60.56% and a redundancy of 2.82%, and belonged to the genus *Lysinbacillus.* Both MAGs were virtually absent in samples at 12 µg/ml of gentamicin and their presence did not show any obvious link to gentamicin. Many contigs (39) contained different (27) ARGs (Figure 6). However, none of these genes was specific to gentamicin. On the other hand, only 137 contigs were co-assembled from soil microcosms and the maximum alignment rate obtained when mapping metagenomic reads against the contigs was of 0.08%. No ARGs were identified when blasting the contigs against the CARD database and filtering at 60% identity and 33 amino acid length.



Bin	Source	Taxonomy	Total Size	Num Contigs	N50	GC Content	Compl.	Red.	SCG Domain
Bin_2	*anvi-interactive*	*Bacillus*	5.03 Mb	1,396	4,864	35.37%	81.69%	9.86%	*bacteria*
Bin_1	*anvi-interactive*	*Lysinibacillus*	1.46 Mb	1,023	1,376	38.37%	60.56%	2.82%	*bacteria*
Bin_3	*anvi-interactive*	*Bacillus*	2.80 Mb	1,736	1,634	36.08%	0.00%	0.00%	*blank*

Figure 6. Assembly of metagenomic samples obtained from soil bacterial enrichments. EM: Enriched Medium. 1,2,3: triplicates from each condition.

Two genes encoding for components of multidrug efflux pumps, *muxB* and *adeF,* were found to significantly increase their relative abundance in non-assembled metagenomic sequences from bacterial enrichments polluted at 12 µg/ml of gentamicin (Table 1). However, none of these efflux pumps is related to gentamicin efflux. Thus, although gentamicin causes significant effects on the bacterial communities of soil bacterial enrichments, a selection for genes conferring resistance to gentamicin was not obvious. Regarding soil microcosms, 26 genes showed an increased relative abundance over exposure time in non-assembled metagenomic sequences from gentamicin-polluted soils (Table S2 in Supplementary Information). 5 of these genes were related to aminoglycoside resistance, 11 were related to multidrug efflux, 5 to tetracycline efflux and 5 to other mechanisms. However, when applying the Leave-One-Out cross-validation to each of these genes, none of these increases in relative abundance was found significant. In addition, although almost all samples from soil microcosms and non-contaminated bacterial enrichments contained genes coding for enzymes of the AAC(6') aminoglycoside acetyltransferase family that had been overlooked by metagenomic sequencing, these were low-abundant and a significant increase on their abundance with gentamicin present was not detected (Figure 7). Therefore, gentamicin, even at high concentrations on the order of mg/g, did not induce any significant change in the resistome of soil bacteria.

[Gentamicin]	**0 µg/ml**			**0.1 µg/ml**			**12 µg/ml**		
Sample/Gene	SC1	SC2	SC3	SC4	SC5	SC6	SC7	SC8	SC9
muxB	0	4,9E-6	0	0	5,4E-6	0	1,5E-4	6,7E-5	1,3E-4
adeF	0	0	0	0	0	0	5,1E-5	9,3E-5	6,4E-5

Table 1. Antibiotic resistance genes (ARGs) that increase their average relative abundance in soil enrichments in 1:10 TSB medium with different gentamicin concentrations after 24-hours incubation (Pearson coefficient > 0.9). Both genes showed a significantly higher abundance at 12 µg/ml of gentamicin compared to the rest of culture enrichments (p-values of 0.046 and 0.032 for *muxB* and *adeF*, respectively).

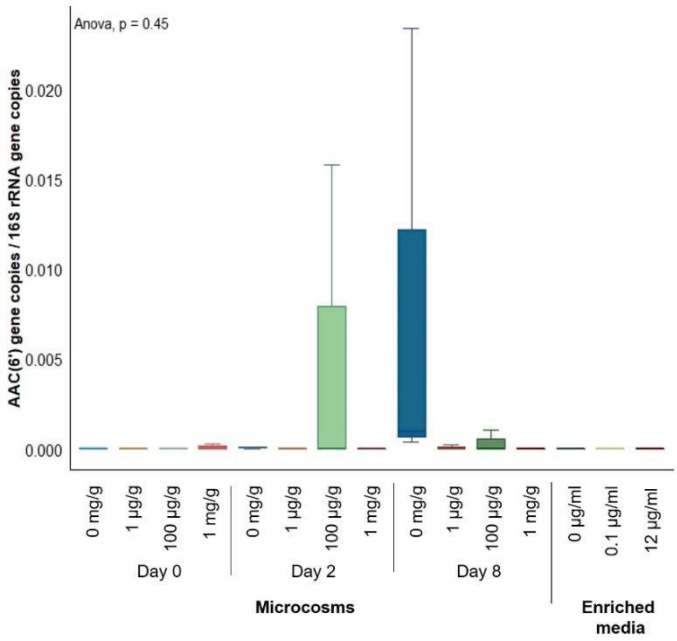


Figure 7. Dynamics of gentamicin resistance gene abundance in soil microcosms and enriched media with and without gentamicin. Number of copies of gentamicin resistance genes from the AAC(6') aminoglycoside acetyltransferase family per µl of qPCR reaction normalized by the number of copies of the 16S rRNA gene per µl of qPCR reaction. D0: 0-day exposure; D2: 2-day exposure; D8: 8-day exposure. No significant differences were found between conditions (each gentamicin concentration and each exposure time for soil microcosms and each gentamicin concentration after 24-hour exposure for bacterial enrichments). n=3.

The available fraction of gentamicin in soils polluted at the concentrations used in this study was measured by High-Performance Liquid Chromatography/Mass Spectometry (HPLC-MS/MS) to estimate the percentage of gentamicin that is adsorbed onto soil particles (Table S3 in Supplementary Information). No trace of gentamicin was detected in any of the measured samples, suggesting that it is almost completely adsorbed onto soil particles. Although a small percentage of gentamicin can escape this adsorption, it was under the quantification limit (10 ng/ml) of the technique used in this study and below the inhibition threshold determined with enriched media. Thus, even when high concentrations of gentamicin were added to soil, the bioavailable fraction was sub-inhibitory. The lack of bioavailability of gentamicin in soil partially accounts for the differences observed between the response of soil bacteria in microcosms and enriched media, where gentamicin was not adsorbed onto soil particles and had an immediate effect on soil bacteria.

Discussion

The main goal of this study was to compare the response of soil bacteria to the antibiotic gentamicin in soil microcosms and enriched media using a metagenomics approach. The effects of gentamicin pollution on the soil microbiome and resistome were analyzed at three levels: overall effects on bacterial growth, impact on the community composition and potential selection for ARGs. As hypothesized, clear differences were observed between the response to gentamicin of soil bacteria enriched in selective media and in soil microcosms. These differences probably lay in the limited bacterial diversity present in bacterial enrichments and in gentamicin adsorption onto soil particles.

The reduced richness detected in bacterial enrichments (Figure S2 in Supplementary Information) and the differences observed in bacterial composition between soil microcosms and bacterial enrichments

(Figure S1c in Supplementary Information) support the concerns about the use of culture-based studies for antibiotic surveillance in terrestrial ecosystems. The majority of the bacteria present in soil are uncultured[41,42] and these uncultured soil bacteria are a reservoir of ARGs[6]. In addition, in complex environments such as soil, other selective forces such as nutrient availability or predation are likely to take place[43]. The impact of the environmental and bacterial interactions on the development of antibiotic resistance is critical and should not be overlooked[44–46]. Therefore, these results demonstrate the need to analyze the effects of antibiotics using microcosms and field studies when possible, since the response observed using culture-based approaches does not necessarily reflect what happens in the environment.

In addition, this research illustrates how the effects of the same antibiotic are strongly dependent on the environmental matrix. Concentrations that were overall sub-inhibitory in bacterial enrichments perturbed the bacterial community structure without changing their resistome. Concentrations that were roughly four orders of magnitude higher than the inhibition threshold did not cause significant short-term effects on soil bacteria in their environmental matrix due to gentamicin adsorption onto soil particles. Thus, this study demonstrated the need to take environmental physico-chemical properties into account in antibiotic surveillance studies in terrestrial ecosystems and to systematize antibiotic concentration measurement, since both antibiotic structure and soil properties affect the behavior of antibiotics in the receiving environment[47]. Furthermore, these results demonstrated the limitations of the terms “sub-inhibitory” and “inhibitory” in complex environments. These terms were initially defined in single cultures and they may not be descriptive of the dose-response relationships taking place in complex communities *in situ*[48], where some members of the community may be inhibited even at concentrations below the overall inhibitory threshold.

Finally, gentamicin resistance genes were detected by qPCR (Figure 7), even though they had been undetected during metagenomics analyses. Although they provide a more complete version of the environmental microbiome than culture-based experiments, metagenomic techniques are still biased and do not uncover all soil bacterial diversity[49]. Sequencing technologies such as the one used in this study (MiSeq sequencing) may not sequence deeply enough to obtain MAGs from samples as rich and diverse as soil. Deeper sequencing technologies should be used in this kind of matrixes in order to obtain a more accurate picture of the soil microbiome[50] and identify the bacterial hosts of environmental ARGs.

Different antibiotic-soil combinations may cause different effects on the soil microbiome. However, this study evaluated the short-term effects of gentamicin after a single dose. Antibiotic sequestration being a reversible process[11], the bioavailability of gentamicin could increase at longer exposure times and cause long-term perturbations on the soil microbial communities. In addition, several studies have shown that environmental factors, such as particle size or mineral composition, create specific microenvironments adequate for the growth of specific bacterial taxa[51–53]. This generates a spatial heterogeneity in the soil microbial communities at a small scale. Moreover, antibiotic concentrations in the environment likely form gradients in soil and these gradients are likely associated with soil resistome heterogeneity at a small scale. The different populations in soil may respond differently to the same antibiotic according to their physiological state[54]. Thus, the dilution of local heterogeneity during DNA extraction at higher sample sizes might hide changes at a micro-scale. Further studies should analyze the effects of antibiotics in soil microenvironments and account for differences related to soil spatial heterogeneity, since studies designed to observe general changes in soil microcosms after antibiotic addition might overlook any event happening at a local scale and, therefore, underestimate the risk associated with antibiotics in soil.

References

1. Delmont, T. O. *et al.* Accessing the soil metagenome for studies of microbial diversity. *Appl. Environ. Microbiol.* **77,** 1315–1324 (2011).

2. Nesme, J. & Simonet, P. The soil resistome: A critical review on antibiotic resistance origins, ecology and dissemination potential in telluric bacteria. *Environ. Microbiol.* **17,** 913–930 (2015).

3. Monier, J. M. *et al.* Metagenomic exploration of antibiotic resistance in soil. *Curr. Opin. Microbiol.* **14,** 229–235 (2011).

4. Naidoo, Y., Valverde, A., Cason, E. D., Pierneef, R. E. & Cowan, D. A. A clinically important, plasmid-borne antibiotic resistance gene (β-lactamase TEM-116) present in desert soils. *Sci. Total Environ.* **719,** 137497 (2020).

5. Perron, G. G. *et al.* Functional characterization of bacteria isolated from ancient arctic soil exposes diverse resistance mechanisms to modern antibiotics. *PLoS One* **10,** 1–19 (2015).

6. Riesenfeld, C. S., Goodman, R. M. & Handelsman, J. Uncultured soil bacteria are a reservoir of new antibiotic resistance genes. *Environ. Microbiol.* **6,** 981–989 (2004).

7. Willms, I. M. *et al.* Discovery of Novel Antibiotic Resistance Determinants in Forest and Grassland Soil Metagenomes. *Front. Microbiol.* **10,** 460 (2019).

8. Luo, W., Xu, Z., Riber, L., Hansen, L. H. & Sørensen, S. J. Diverse gene functions in a soil mobilome. *Soil Biol. Biochem.* **101,** 175–183 (2016).

9. Jiang, X. *et al.* Dissemination of antibiotic resistance genes from antibiotic producers to pathogens. *Nat. Commun.* **8,** 1–7 (2017).

10. Hoelzer, K. *et al.* Antimicrobial drug use in food-producing animals and associated human health risks: What, and how strong, is the evidence? *BMC Vet. Res.* **13,** 1–38 (2017).

11. Jechalke, S., Heuer, H., Siemens, J., Amelung, W. & Smalla, K. Fate and effects of veterinary antibiotics in soil. *Trends Microbiol.* **22,** 536–545 (2014).

12. Blau, K. *et al.* Manure and Doxycycline Affect the Bacterial Community and Its Resistome in Lettuce Rhizosphere and Bulk Soil. *Front. Microbiol.* **10,** (2019).

13. Wolters, B. *et al.* Bulk soil and maize rhizosphere resistance genes, mobile genetic elements and microbial communities are differently impacted by organic and inorganic fertilization. *FEMS Microbiol. Ecol.* **94,** (2018).

14. Kang, Y. *et al.* High diversity and abundance of cultivable tetracycline-resistant bacteria in soil following pig manure application. *Sci. Rep.* **8,** 1489 (2018).

15. Jechalke, S. *et al.* Increased abundance and transferability of resistance genes after field application of manure from sulfadiazine-treated pigs. *Appl. Environ. Microbiol.* **79,** 1704–1711 (2013).

16. Heuer, H. *et al.* IncP-1ε plasmids are important vectors of antibiotic resistance genes in agricultural systems: Diversification driven by class 1 integron gene cassettes. *Front. Microbiol.* **3,** 1–8 (2012).

17. Heuer, H. *et al.* Accumulation of sulfonamide resistance genes in arable soils due to repeated application of manure containing sulfadiazine. *Appl. Environ. Microbiol.* **77,** 2527–2530 (2011).

18. Surette, M. & Wright, G. D. Lessons from the Environmental Antibiotic Resistome. *Annu. Rev. Microbiol.* **71,** 309–329 (2017).

19. Pan, M. & Chu, L. M. Adsorption and degradation of five selected antibiotics in agricultural soil. *Sci. Total Environ.* **545–546,** 48–56 (2016).

20. Albero, B., Tadeo, J. L., Escario, M., Miguel, E. & Pérez, R. A. Persistence and availability of veterinary antibiotics in soil and soil-manure systems. *Sci. Total Environ.* **643,** 1562–1570 (2018).

21. Thiele-Bruhn, S. Pharmaceutical antibiotic compounds in soils - A review. *J. Plant Nutr. Soil Sci.* **166,** 145–167 (2003).

22. Kemper, N. Veterinary antibiotics in the aquatic and terrestrial environment. *Ecol. Indic.* **8,** 1–13 (2008).

23. Brandt, K. K. *et al.* Ecotoxicological assessment of antibiotics: A call for improved consideration of microorganisms. *Environ. Int.* **85,** 189–205 (2015).

24. Manaia, C. M. *et al.* Antibiotic resistance in wastewater treatment plants: Tackling the black box. *Environ. Int.* **115,** 312–324 (2018).

25. Durso, L. M. & Cook, K. L. Impacts of antibiotic use in agriculture : what are the benefits and risks ? *Curr. Opin. Microbiol.* **19,** 37–44 (2014).

26. Cycon, M., Mrozik, A. & Piotrowska-Seget, Z. Antibiotics in the soil environment - Degradation and their impact on microbial activity and diversity. *Front. Microbiol.* **10,** (2019).

27. Heuer, H. *et al.* Gentamicin resistance genes in environmental bacteria: Prevalence and transfer. *FEMS Microbiol. Ecol.* **42,** 289–302 (2002).

28. Josephson, L., Houle, P. & Haggerty, M. Stability of dilute solutions of gentamicin and tobramycin. *Clin. Chem.* **25,** 298–300 (1979).

29. Muyzer, G., Hottentrager, S., Teske, A. & Wawer, C. Denaturing gradient gel electrophoresis of PCR-amplified 16S rDNA—a new molecular approach to analyse the genetic diversity of mixed microbial communities. In *Molecular microbial ecology manual;* (eds. Akkermans, A., van Elsas, J. & de Bruijn, F.) 1–23 (Kluwer Academic Publishers, 1995).

30. Watanabe, K., Kodama, Y. & Harayama, S. Design and evaluation of PCR primers to amplify bacterial 16S ribosomal DNA fragments used for community fingerprinting. *J. Microbiol. Methods* **44,** 253–262 (2001).

31. Klindworth, A. *et al.* Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res.* **41,** 1–11 (2013).

32. Masella, A. P., Bartram, A. K., Truszkowski, J. M., Brown, D. G. & Neufeld, J. D. PANDAseq: Paired-end assembler for illumina sequences. *BMC Bioinformatics* **13,** 31 (2012).

33. Wang, Q., Garrity, G. M., Tiedje, J. M. & Cole, J. R. Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* **73,** 5261–5267 (2007).

34. Oksanen, J. *et al.* vegan: Community Ecology Package. R package version 2.4-2. *Community Ecol. Packag.* **2.5-6,** 1–296 (2019).

35. Minoche, A. E., Dohm, J. C. & Himmelbauer, H. Evaluation of genomic high-throughput sequencing data generated on Illumina HiSeq and Genome Analyzer systems. *Genome Biol.* **12,** (2011).

36. Li, D., Liu, C. M., Luo, R., Sadakane, K. & Lam, T. W. MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* **31,** 1674–1676 (2015).

37. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9,** 357–359 (2012).

38. Eren, A. M. *et al.* Anvi'o: An advanced analysis and visualization platformfor 'omics data. *PeerJ* **2015,** 1–29 (2015).

39. Alcock, B. P. *et al.* CARD 2020: Antibiotic resistome surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids Res.* **48,** D517–D525 (2020).

40. Guironnet, A., Wiest, L., Sanchez-Cid, C., Vogel, T. M. & Vulliet, E. Aminoglycosides analysis optimization using Ion pairing Liquid Chromatography coupled to tandem Mass Spectrometry and application on wastewater samples. *J. Chromatogr.* Under review.

41. Rappé, M. S. & Giovannoni, S. J. The Uncultured Microbial Majority. *Annu. Rev. Microbiol.* **57,** 369–394 (2003).

42. Schloss, P. D. & Hadelsman, J. Status of the Microbial Census. *Microbiol. Mol. Biol. Rev.* **68,** 686–691 (2004).

43. Bengtsson-Palme, J. & Larsson, D. G. J. Concentrations of antibiotics predicted to select for resistant bacteria: Proposed limits for environmental regulation. *Environ. Int.* **86,** 140–149 (2016).

44. Hughes, D. & Andersson, D. I. Evolutionary Trajectories to Antibiotic Resistance. *Annu. Rev. Microbiol.* **71,** 579–596 (2017).

45. Cairns, J. *et al.* Ecology determines how low antibiotic concentration impacts community composition and horizontal transfer of resistance genes. *Commun. Biol.* **1,** 1–8 (2018).

46. Bergk Pinto, B., Maccario, L., Dommergue, A., Vogel, T. M. & Larose, C. Do Organic Substrates Drive Microbial Community Interactions in Arctic Snow? *Front. Microbiol.* **10,** 1–13 (2019).

47. Zhi, D. *et al.* Current progress in the adsorption, transport and biodegradation of antibiotics in soil. *J. Environ. Manage.* **251,** (2019).

48. Andersson, D. I. & Hughes, D. Microbiological effects of sublethal levels of antibiotics. *Nat. Rev. Microbiol.* **12,** 465–478 (2014).

49. Delmont, T. O., Simonet, P. & Vogel, T. M. Describing microbial communities and performing global comparisons in the omic era. *ISME J.* **6,** 1625–1628 (2012).

50. Delmont, T. O., Franqueville, L., Jacquiod, S., Simonet, P. & Vogel, T. M. Soil Metagenoic Exploration of the Rare Biosphere. In *Handbook of Molecular Microbial Ecology. Volume I: Metagenomics and Complementary Approaches.* **2011,** 287–298. doi:10.1128/AEM.01931-08

51. Hemkemeyer, M., Dohrmann, A. B. & Christensen, B. T. Bacterial Preferences for Specific Soil Particle Size Fractions Revealed by Community Analyses. *Front. Microbiol.* **9,** (2018).

52. Neumann, D., Heuer, A., Hemkemeyer, M., Martens, R. & Tebbe, C. C. Response of microbial communities to long-term fertilization depends on their microhabitat. *FEMS Microbiol. Ecol.* **86,** 71–84 (2013).

53. Martens, R. & Tebbe, C. C. Artificial soil studies reveal domain-specific preferences of microorganisms for the colonisation of different soil minerals and particle size fractions €. *FEMS Microbiol. Ecol.* **90,** 770–782 (2014).

54. Bernier, S. P. & Surette, M. G. Concentration-dependent activity of antibiotics in natural environments. *Front. Microbiol.* **4,** (2013).

Chapter IV. Gentamicin pollution at sub-inhibitory concentrations induces a response in the environmental microbiome and resistome

Abstract

Residual concentrations of antibiotics are often released to the environment after wastewater treatment and through waste from pharmaceutical factories, farms, and aquacultures. These antibiotics might impose a selective pressure on the environmental resistome and induce the selection for ARB and ARG that can be disseminated to human and animal microbiomes. Residual concentrations of antibiotics are often sub-inhibitory and have traditionally been considered as innocuous. However, an increasing body of evidence shows that sub-inhibitory concentrations of antibiotics might select for resistance and induce a wide range of processes related to resistance development *in vitro.* This is, to the best of our knowledge, the first study to determine the effects of sub-inhibitory concentrations of antibiotics on the environmental microbiome. Given the higher complexity of the microbial community and the environmental matrix in different ecosystems, antibiotics could have considerably different effect than that observed in microbial cultures in the laboratory. River water microcosms were dosed with both sub-inhibitory and inhibitory concentrations of gentamicin as determined previously based on bacterial growth in enriched media. DNA and RNA were extracted after 0, 1 and 2-day exposure, and the effects of gentamicin were assessed at three different levels: a) overall effects on bacterial growth and transcription, b) effect on the composition of the total and active communities, and c) effect of the abundance of *aac(6')* and *aadA* aminoglycoside resistance gene transcripts. Furthermore, the genes associated to specific taxa were identified by metagenomic sequencing and both the assembly of short reads and the hybrid assembly of short and long reads. A sub-inhibitory concentration (gentamicin at 50 ng/ml) provoked a significant increase in overall transcription levels without inhibiting growth. Furthermore, a shift in the total and active communities was identified in water microcosms with gentamicin at that concentration after only a 1-day exposure with an concomitant increase *Limnohabitans* abundance. The gentamicin resistance gene (*aac(6')-Ib8*) was identified in the partially-assembled genome of *Limnohabitans.* Finally, the abundance and transcription of genes belonging to the *aac(6')* family identified in class 1 integrons from water microcosms increased under gentamicin pollution at 50 ng/ml. Thus, gentamicin at sub-inhibitory concentrations induced a response (identified via mobile genetic elements) in the environmental microbiome and resistome, a change in community composition and activity, and a selection of genes conferring resistance to gentamicin. This is the first study showing evidence of the effects of sub-inhibitory concentrations of antibiotics *in situ* and establishing dose-response relationships between antibiotic dose measured by HPLC-MS/MS and the magnitude of the resistome response. The scope of antibiotic selection under sub-inhibitory concentrations of antibiotics and the mechanisms underlying this process might provide the basis for understanding antibiotic resistance dispersion in low antibiotic concentration ecosystems.

Introduction

Most of the antibiotics used in human therapy and animal production are excreted unaltered to the environment[1]. These bioactive antibiotics enter wastewater treatment plants (WWTP) and agricultural fields and groundwater, and stimulate resistance development and horizontal gene transfer (HGT). Thus, even though the majority of antibiotic resistant bacteria (ARB) and their associated antibiotic resistance genes (ARGs) present in wastewater influents are removed during treatment[2,3], some of them can be released to the environment[4,5,6]. Thus, antibiotics remaining in WWTP effluents can reach the receiving aquatic environment[7]. In addition, they can contaminate food crops through the use of treated wastewater for irrigation[8]. Residual antibiotic concentrations might also reach fresh water through the waste of antibiotic manufacturing plants, farms and aquacultures[9]. The selective pressure imposed by environmental pollution with residual concentrations of antibiotics due to anthropogenic activities might result in the development and selection of antibiotic resistance in environmental settings[10,11] and the subsequent dissemination of ARB and ARGs from the environment to animal and human microbiomes[12,13]. Nonetheless, the scope of this phenomenon remains unclear[14–16] .

Antibiotic concentrations found in anthropogenically-polluted environments are often sub-inhibitory[17] (*i.e*. too low to cause a *visible* growth inhibition of susceptible bacteria in culture[18]). However, sub-inhibitory concentrations might cause a perturbation of the microbial community. They could slow down bacterial growth[19] and the selective pressure they exert might be sufficient to offset the cost of resistance, and thereby, contribute to resistance selection[20]. An increasing body of evidence shows that sub-inhibitory concentrations of antibiotics might select for ARB[19,21] and induce a wide range of responses including quorum sensing, persistence, biofilm formation and the expression of genes involved in antibiotic resistance, virulence and the SOS response[22]. Furthermore, sub-inhibitory concentrations of antibiotics might select for mutants with a lower fitness cost than the ones selected at inhibitory concentrations[19] and might be sufficient for the maintenance of multidrug resistance plasmids[23]. These concentrations might also stimulate ARG integration in mobile genetic elements (MGEs) and dissemination through horizontal gene transfer (HGT)[24–27]. Therefore, at least some antibiotics might select for resistance in a dose-independent manner and contribute to resistance development at concentrations below the minimal inhibitory concentration (MIC) of susceptible bacteria as much as at inhibitory concentrations[20].

Published studies working to elucidate the effects of sub-inhibitory concentrations of antibiotics on bacterial communities and their associated resistome base their findings on culture-based approaches. Bacteria need to possess genes involved in antibiotic resistance that are translated to proteins in order to grow in antibiotic-polluted media. Therefore, culture-based approaches provide phenotypical evidence of antibiotic resistance and the mechanisms involved. However, these approaches deal with a reduced portion of the bacterial community under established *in vitro* conditions and do not necessarily account for the complexity of natural environments[28], where bacterial growth and interactions are influenced by several factors. Experiments with natural microbial communities can be monitored by environmental DNA sequencing (*i.e*., metagenomics approaches). Metagenomic studies avoid culturing biases and provide access to a much wider proportion of the environmental microbiome without altering the original conditions of the environmental matrix. Nevertheless, the genes and taxa identified using metagenomics approaches are not necessarily actively responding to the antibiotic. Thus, the combination of metagenomic studies with RNA-based analyses should provide

a better understanding of the taxa and genes involved in the response to sub-inhibitory concentrations of antibiotics.

In this study, we examined whether sub-inhibitory concentrations of antibiotics induce a response at the taxonomical and functional level in the aquatic environment and established relationships between antibiotic dose and the magnitude of the response using gentamicin as a model antibiotic. Gentamicin is an aminoglycoside used in both human therapy and veterinary medicine. Residual concentrations of excreted gentamicin have been detected in wastewater treatment plants[29]. Gentamicin at sub-inhibitory concentrations has been shown to induce resistance development in pure cultures[30]. Moreover, aminoglycoside resistance genes are widely distributed in both chromosomal and plasmid class 1 integrons[31], which are considered major drivers in antibiotic resistance dissemination[32]. Thus, using a combination of metagenomics and RNA-based analyses, the bacteria and genes involved in the environmental response to gentamicin at sub-inhibitory concentrations in river water microcosms were identified and class 1 integron cassettes were sequenced to screen for aminoglycoside resistance genes. The response to sub-inhibitory concentrations of gentamicin was compared to that observed under gentamicin pollution at inhibitory concentrations. We hypothesized that inhibitory concentrations of gentamicin would induce a stronger response than sub-inhibitory concentrations, but that the latter would induce shifts in the composition of the microbial communities and increase the abundance and transcription of genes related to aminoglycoside resistance found in class 1 integrons. To the best of our knowledge, this is the first study showing that sub-inhibitory concentrations of gentamicin induce a response in the environmental microbiome and resistome and establishing links between antibiotic dose and the magnitude of the response.

Materials and methods

Sampling and determination of gentamicin effect on river water enrichments

Rhône river water was sampled in Lyon (45°45'08.3"N 4°50'11.3"E). In order to determine which gentamicin concentrations had sub-inhibitory effects on Rhône river water, 200 µl of river water containing bacteria in R2A medium without antibiotics or with gentamicin at 10, 50, 100, 500 or 1000 ng/ml were transferred to a 96-well culture plate and incubated for 45 hours at 29°C under continuous shaking in the MultiSkan GO Plate Reader (Thermo Scientific). OD_{600} was measured every hour (Figure S1 in Supplementary Information). Based on these results, 10 and 50 ng/ml were identified as sub-inhibitory concentrations. Since both 500 ng/ml and 1 µg/ml were identified as inhibitory, an intermediary inhibitory concentration of 800 ng/ml was the one selected for this study.

Microcosm experiment using river water polluted with gentamicin

One-liter microcosms of Rhône river water were prepared in polypropylene containers in order to avoid gentamicin adsorption on glass. Then, 0, 10, 50 or 800 ng/ml of gentamicin were added to the river water and incubated for 48 hours at room temperature. Microcosms were gently shaken at 35 rpm to emulate river currents. Triplicates were made for each gentamicin concentration. After 0, 24- and 48-hour exposure, DNA and RNA were extracted from water microcosms. Gentamicin concentration was measured for every sample and exposure time by HPLC-MS/MS (Table S1 in Supplementary information).

DNA extraction from water microcosms

For each extraction, 90 ml of sample were filtered through a 0.2 µm filter using a vacuum pump. Then, the bacteria on the filter were resuspended in 1 ml of CTAB Buffer (Promega) and heated for 5 minutes at 95 °C. After vortexing for 1 minute, 300 µl of supernatant were added to 300 µl of Lysis Buffer (Promega) and purified using the Maxwell® RSC Instrument and the Maxwell® RSC PureFood GMO and Authentication Kit (Promega). DNA was eluted in 100 µl of Elution Buffer (Promega).

RNA extraction from water microcosms

For each extraction, 90 ml of sample were filtered through a 0.2 µm filter using a vacuum pump. Then, the bacteria on the filter were resuspended in 1 ml of CTAB Buffer (Promega) with 2% 1-thioglycerol and vortexed for 1 minute. 200 µl of supernatant were added to 200 µl of Lysis Buffer (Promega) and purified using the Maxwell® RSC Instrument and the Maxwell® RSC miRNA Tissue Kit (Promega), which includes DNase treatment. RNA was eluted in 60 µl of nuclease-free water.

DNA quantification and 16S rRNA gene quantitative PCR (qPCR) assays

DNA concentrations were measured using the Qubit Fluorometer and the Qubit dsDNA HS Assay Kit (ThermoFisher). Then, the size of the total bacterial community was estimated by quantifying the V3 region of the 16S rRNA gene by qPCR using the "universal" primers 341F (5'-CCT ACG GGA GGC AGC AG- 3') and 534R (5'-ATT ACC GCG GCT GCT GGC A-3')[33,34]. qPCR assays were carried out using the Corbett Rotor-Gene 6000 (QIAGEN) in a 20 µl reaction volume containing GoTaq qPCR Master Mix (Promega), 0.75 µM of each primer and 2 µl of DNA. Two non-template controls were also included in all the assays. Standard curves for all the assays were obtained using 10-fold serial dilutions of a linearized plasmid pGEM-T Easy Vector (10^2 to 10^7 copies) containing the 16S rRNA gene of *Pseudomonas aeruginosa* PAO1. Cycling conditions for qPCR amplification were 95 °C for 2 minutes followed by 30 cycles of 95 °C for 15 seconds, 60 °C for 30 seconds and 72 °C for 30 seconds. Melting curves were generated after amplification by increasing the temperature from 60 °C to 95 °C. Then, ANOVA tests andt-student tests between each group (each gentamicin concentration and each exposure time) and the average between all groups were performed using the ggpubr package in R.

RNA quantification, retrotranscription and 16S rRNA gene quantitative PCR (qPCR) assays

RNA was quantified using the Quantus Fluorometer and the QuantiFluor RNA System (Promega). Then, RNA samples were diluted by 10x in nuclease-free water and retrotranscribed using the SuperScript III RT (Invitrogen). cDNA was analyzed by qPCR assays and statistical analyses as explained in the previous paragraph in order to estimate the size of the active bacterial community.

16S rRNA gene and cDNA sequencing and analysis

The V3-V4 hypervariable regions of the bacterial 16S rRNA gene were amplified using the Titanium Taq DNA Polymerase (Takara Clontech) and forward (5'-TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG CCT ACG GGN GGC WGC AG-3') and reverse (5'-GTC TCG TGG GCT CGG AGA TGT GTA TAA GAG ACA GGT CTC GTG GGC TCG GAG ATG TGT ATA AGA GAC AGG ACT ACH VGG GTA TCT AAT CC-3')[35] primers. DNA was amplified as follows: 95 °C for 3 minutes followed by 30 cycles of 95 °C for 30 seconds, 55 °C for 30 seconds and 72 °C for 30 seconds and a final extension step at 72°C for 5 minutes. cDNA was amplified using the Platinum Taq DNA Polymerase (Invitrogen) and the primers described above. cDNA was amplified as follows: 94 °C for 2 minutes followed by 30 cycles of 94 °C for 30 seconds, 55 °C for 30 seconds and 72 °C for 30 seconds and a final extension step at 72°C for 5 minutes. DNA and cDNA libraries were prepared from amplified products based on Illumina's "16S Metagenomics Library Prep Guide" (15044223 Rev. B) using the

Platinum Taq DNA Polymerase (Invitrogen) and the Nextera XT Index Kit V2 (Illumina). DNA and cDNA were sequenced with a 15% PhiX spike-in using the MiSeq System and the MiSeq Reagent Kit v2 (Illumina). Reads were trimmed to meet a quality score of Q20. Then, pair-ended reads were assembled using PANDAseq[36] at a sequence length between 410 and 500 bp and an overlap length between 20 and 100 bp using the rdp_mle algorithm. Finally, each of the DNA sequences was annotated to the genus level using the Ribosome Data Project (RDP) database and the RDP Bayesian classifier[37] using an assignment confidence cut-off of 0.6. The genera that had less than 10 associated sequences in the ensemble of DNA or cDNA sequences were removed. Finally, the relative abundances of the remaining genera were assessed and normalized per number of copies of the 16S rRNA gene per L of water.

Illumina metagenomics sequencing and analysis

Metagenomics libraries were prepared from ≤1 ng of DNA obtained from water microcosms using the Nextera XT Library Prep Kit and Indexes (Illumina), as detailed in Illumina's "Nextera XT DNA Library Prep Kit" reference guide (15031942 v03). DNA sequencing with a 1% PhiX spike-in was performed using the MiSeq System and the MiSeq Reagent Kit v2 (Illumina). Metagenomic reads were filtered according to the criteria described by Minoche *et al*[38]. Then, reads were co-assembled using MEGAHIT[39] to generate contigs. Reads were mapped onto the contigs using Bowtie2[40] to generate BAM files. Data regarding the co-assembly of Illumina reads can be found in Table S2 (Supplementary Information). Profiles were created for each individual sample and merged using the anvi'o[41] metagenomic workflow. The assembled contigs were blasted against the CARD antibiotic gene database[42] to identify ARGs. Results were filtered at an amino acid identity percentage of 60%, 33 amino acid length and an e-value of 10^{-5}. The best hit was used. Finally, contigs were binned based on their differential coverage across samples using anvi'o[41], and the bins were refined based on differential coverage and sequence composition. Bins with <50% completion and ≥10% redundancy were discarded, since they were considered low-quality metagenome assembled genome (MAG) drafts according to Bowers *et al*[43]. A summary of the remaining bins is shown in Figure S2 (Supplementary Information).

Oxford Nanopore metagenomics sequencing and analysis

One hundred ng of pooled DNA extracted from the three microcosms polluted at 50 ng/ml of gentamicin after 2-day exposure were sequenced using the Oxford Nanopore MiniON and the Oxford Nanopore Ligation Sequencing Kit. Libraries were prepared as detailed in Oxford Nanopore's "1D Genomic DNA by Ligation" protocol. Short-reads sequenced using Illumina's MiSeq System and filtered according to the criteria described by Minoche *et al.* were co-assembled to long-reads obtained from the Oxford Nanopore sequencing using Unicycler[44]. Information regarding the hybrid assembly of short and long reads can be found in Table S3 (Supplementary Information). The rest of sequencing analysis, from the mapping of Illumina reads onto the assembled contigs to the refinement of bins, was described in the previous section.

Sequencing of chromosomal class 1 integron cassettes

Since aminoglycoside resistance genes are often integrated in class 1 integrons, chromosomal class 1 integron cassettes were amplified and sequenced in order to identify the ARGs that they encoded. Specific primers for the integrase *int1* (HS463a: 5'-CTGGATTTCGATCACGGCACG-3' and HS464: 5'-ACATGCGTGTAAATCATCGTCG-3')[45] were used to detect class 1 integrons in water DNA samples. Three µl of DNA were amplified using the Platinum Taq DNA Polymerase (Thermofisher) as follows: 94 °C for

2 minutes followed by 35 cycles of 94 °C for 30 seconds, 60 °C for 30 seconds and 72 °C for 90 seconds and a final extension step at 72°C for 5 minutes. All water DNA samples contained class 1 integrons. Then, the integron gene cassette arrays were amplified from water DNA using MRG284 (5'-GTTACGCCGTGGGTCGATG-3') and MRG285 (5'- CCAGAGCAGCCGTAGAGC-3') primers[46]. These primers amplified the entire cassette array in chromosomal class 1 integrons and Tn402-like elements from the aatI1 site to a conserved region beyond the final *attC* site in the cassette array. Annealing was performed at 65°C for 2 minutes instead of 60°C for 30 seconds. Then, amplicons were cleaned-up using AMPure XP beads (Beckman-Coulter) and sequenced as explained in the "Illumina metagenomic sequencing and analysis" section.

Metagenomics reads were trimmed using the Fastq Quality Trimmer tool of the FASTX-Toolkit. Nucleotides that did not meet a minimum quality score of Q20 were trimmed from the sequences and sequences shorter than 100 nucleotides after trimming were removed. R1 and R2 reads were blasted separately against the CARD database using Diamond[47]. The obtained results were filtered at a minimum identity of 90%, a minimum length of 50 amino acids and an e-value of 10^{-10}. The best hit was used. Results obtained from R1 and R2 blasts were identical. Therefore, the analysis was continued using only R1. The abundance of the aminoglycoside resistance genes encoded by class 1 integrons was normalized by sequencing depth (Figure S3 in Supplementary Information).

Quantification of aminoglycoside resistance genes by qPCR

Thirty-two aminoglycoside resistance genes coding for *aadA* aminoglycoside nucleotidyltransferases and *AAC*(6') aminoglycoside acetyltransferases were identified in integron cassette arrays. Since abundance calculations of aminoglycoside resistance genes based on the sequencing of amplified integron cassettes lacks accuracy, two primer pairs were designed to target conserved regions of the identified *aadA* (17 genes) and the *aac(6')* (15 genes) genes, respectively. Information regarding these primers is in Table S4 (Supplementary Information). Then, *aadA* and *aac(6')* genes were quantified by qPCR in water DNA and cDNA samples in order to evaluate the abundance and transcription of aminoglycoside resistance genes under gentamicin pollution at different concentrations and exposure times. The Corbett Rotor-Gene 6000 (QIAGEN) amplified target genes in a 20 µl reaction volume containing GoTaq qPCR Master Mix (Promega), 0.75 µM of each primer and 2 µl of DNA or cDNA. Two non-template controls were also included in all the assays and non-retrotranscribed controls were included in cDNA amplifications. Standards for qPCR assays were obtained from water DNA samples using *aadA* and *aac(6')* primers. Standards were cloned and transformed using the TOPO TA cloning Kit (Invitrogen) and normalized to 10^8 copies/µl. Standard curves were made using 10-fold serial dilutions of the standards ($10^7 – 10^2$ copies/µl). Cycling conditions for qPCR amplification were 95 °C for 2 minutes followed by 35 cycles of 95 °C for 30 seconds, 55 °C for 30 seconds and 72 °C for 30 seconds. Melting curves were generated after amplification by increasing the temperature from 60 °C to 95 °C. The number of copies per µl of qPCR reaction obtained from the amplification of *aadA* and *aac(6')* genes or transcripts were normalized by the copies of the 16S rRNA gene or transcripts per µl of qPCR reaction to assess their relative abundance in water. Average values and the percentage of standard deviation calculated for each condition are shown in Table S5 (Supplementary Information). Since gentamicin pollution at 800 ng/ml provoked a higher increase on gene abundance and transcription than sub-inhibitory concentrations, dose-response graphs were created only for sub-inhibitory concentrations for better visualization of the effect of gentamicin pollution at sub-inhibitory concentrations on aminoglycoside resistance gene abundance and transcription.

Results

Effect of gentamicin on the growth of enriched bacterial communities from river water

At 10 and 50 ng/ml of gentamicin, the offset of the growth curve was delayed. At 50 ng/ml, slower growth rates were observed (Figure S1 in Supplementary Information). However, similar OD_{600} than in non-polluted controls were detected after 45-hours growth at both concentrations. Therefore, 10 ng/ml and 50 ng/ml of gentamicin did not inhibit overall growth in the enriched fraction of water bacteria. Concentrations from 100 ng/ml cause a strong growth inhibition in water bacteria. Indeed, no growth was observed in enriched media polluted with gentamicin at 500 ng/ml nor at 1 µg/ml after 45-hours incubation. Therefore, a concentration of 800 ng/ml of gentamicin was chosen as an inhibitory concentration control, whereas 10 ng/ml and 50 ng/ml were the sub-inhibitory concentrations selected for this study.

Gentamicin at 50 ng/ml increased global transcription levels without inhibiting overall bacterial growth

Based on ANOVA statistical tests, significant ($p \leq 0.05$) differences between the conditions (different gentamicin concentrations and exposure times) were found for DNA and RNA yield and number of copies of the 16S rRNA gene and transcript (Figure 1). In addition, whereas the size of the bacterial community was significantly lower in microcosms polluted at 800 ng/ml after 2-day exposure than those at the other concentrations (p-values<0.05) for both DNA quantity and number of copies of the 16S rRNA gene), both the quantity of DNA and the number of copies of the 16S rRNA gene increased slightly over time. In the non-polluted microcosms and microcosms with gentamicin at sub-inhibitory concentrations (10 or 50 ng/ml), the bacterial concentration increased significantly (Figure 1). Significantly higher bacterial transcription levels were detected in microcosms polluted up to 50 ng/ml of gentamicin based on the number of copies of 16S rRNA transcripts (p-value<0.0001). Therefore, gentamicin at 50 ng/ml did not inhibit bacterial transcription nor overall bacterial growth in the river water microcosms.

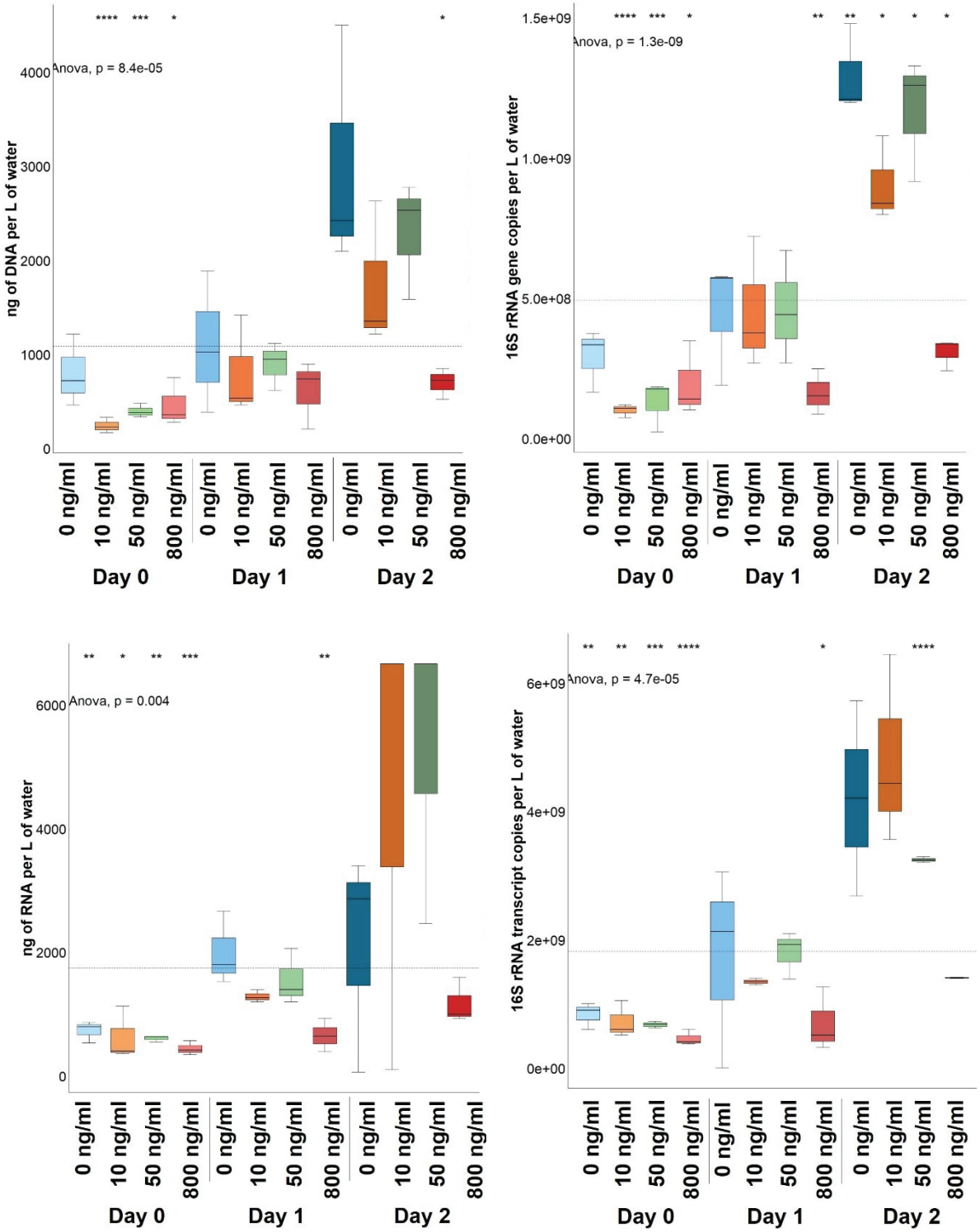


Figure 1. Bacterial dynamics over time at different gentamicin concentrations in river water microcosms. Yield of extracted DNA in ng per L (top left), 16S rRNA gene copies per L (top right), yield of extracted RNA in ng per L (bottom left), 16S rRNA transcript copies per L (bottom right) obtained from non-polluted water or water polluted with gentamicin at 10, 50 or 800 ng/ml after 0, 1 or 2-day exposure. qPCR efficiency=1.07. R^2 linearity coefficient=0.996. Significant differences between each group (each gentamicin concentration and each exposure time) and the average between all groups (horizontal dashed line) were determined by a t-test. *p-value ≤0.05. **p-value≤0.01. ***p-value ≤0.001. ****p-value≤0.0001. n=3.

Gentamicin at 50 ng/ml increased the abundance of Limnohabitans in the total and active communities of water microcosms after 1-day exposure

The average relative abundance of the 24 most abundant genera amongst 16S rRNA gene and cDNA sequences represented between 65 and 90% of the total communities and between 75% to 95% of the active communities sequenced from water microcosms. All gentamicin concentrations tested affected (although differently for each concentration) the composition of total and active communities in the river water microcosms. First, the abundance of several genera, including *Zooglea, Sphingohabdus* and *Polynucleobacter,* was slightly lower in the total and active communities from microcosms polluted at 10 ng/ml of gentamicin after a 2-day exposure compared to non-polluted controls (Figure 2). In addition, the abundance of *Limnohabitans* and *Pseudomonas* transcripts increased over time at this gentamicin concentration. However, this concentration does not seem to have a big impact on community composition. In microcosms polluted at 50 ng/ml of gentamicin, a shift in community composition was observed after a 1-day exposure, when the average relative abundance of *Limnohabitans* increased from 6.28% to 34.64% of the total communities and from 8.91% to 47.68% of the active communities. This increase became more evident after a 2-day exposure, when *Limnohabitans* represented 58.59% of the total communities and 60.36% of the active communities present in microcosms polluted at 50 ng/ml of gentamicin, whereas its relative abundance in the other microcosms was relatively stable over time and concentrations. This increase was also observed after normalizing the number of copies of the 16S rRNA gene per liter of water. Other genera such as *Pseudomonas* and *Zooglea* were less abundant in microcosms polluted at 50 ng/ml of gentamicin after a 2-day exposure than in non-polluted controls. On the other hand, at 800 ng/ml, a concentration that inhibited growth in water microcosms, the abundance and transcription of virtually all genera was reduced compared to the rest of microcosms. However, *Pseudomonas* increased its abundance and transcription over time and thus seems to be able to grow under gentamicin pollution at 800 ng/ml. In addition, even though *Rhodoferax* abundance was lower in microcosms with 800 ng/ml of gentamicin than in the other microcosms, its transcription levels increased over time.

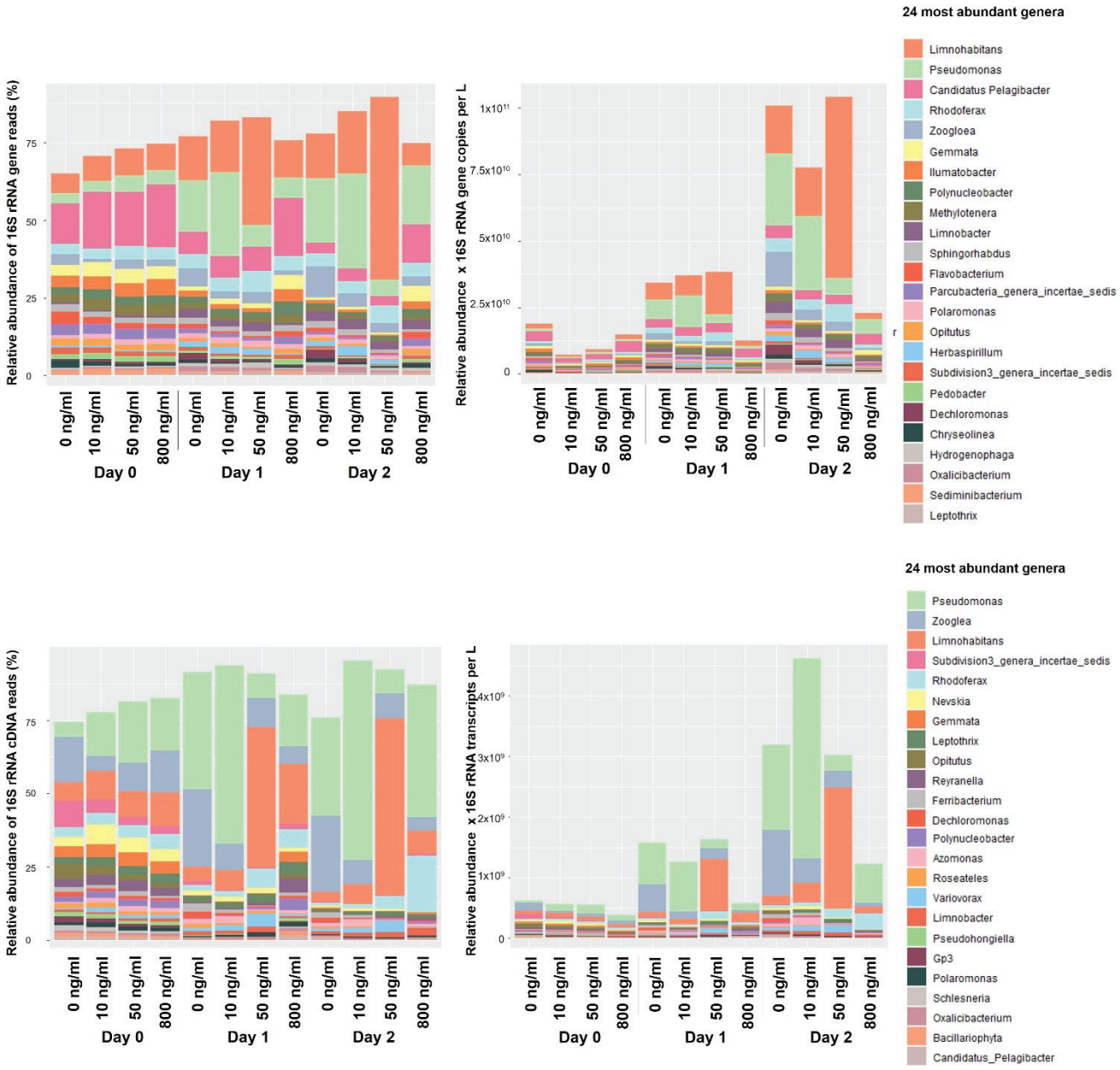


Figure 2. Changes in the relative abundance and transcription levels of the 24 most abundant bacteria in river water microcosms. Average relative abundance of 16S rRNA gene sequences (top left); average relative abundance normalized per number of copies of 16S rRNA gene per L of water (top right); average relative abundance of 16S rRNA cDNA sequences (bottom left); average relative abundance normalized per number of copies of 16S rRNA transcripts per L of water (bottom right) from non-polluted water microcosms and microcosms polluted at 10, 50 or 800 ng/ml of gentamicin after 0, 1 or 2-day exposure. n=3.

Hybrid assembly of short and long reads of a gentamicin resistance gene in a partial genome reconstruction from Limnohabitans

Short metagenomic reads obtained from water microcosms were co-assembled in order to obtain metagenome-assembled genomes (MAGs) and identify their associated ARGs. Three bins with more than 50% completion and less than 10% redundancy were assembled from Illumina metagenomic reads (Figure S2 in Supplementary Information). The first bin had a completion of 60.2%, a redundancy of 4.2, contained 683 splits, and had a length of 1.86 Mbp. Its coverage increased over time in

metagenomic sequences from all microcosms except for the ones contaminated at 800 ng/ml of gentamicin. It was identified as *Burkholderiaceae.* Two hits from the CARD antibiotic database were found in this MAG. One belonged to the *mexW* gene that encodes for the MexW protein of the MexW-OprM multidrug efflux pump. The other one belonged to the *ugd* gene that is involved in peptide resistance. The second bin was identified as *Pseudomonas*. It contained 602 splits, had a completion of 67.6%, a redundancy of 1.4%, and a length of 852 Kbp. Its coverage increased over time in all microcoms except for the ones contaminated at 50 ng/ml of gentamicin. Four hits from the CARD antibiotic database were identified in this bin. All hits were related to drug efflux (*opmH* for triclosan efflux and *mexW, mexF* and *mvaT* related to multidrug efflux). Finally, the third bin, which had a completion of 56.3%, a redundancy of 5.6%, a length of 2.04 Mbp, and contained 880 splits was identified as *Limnohabitans.* Its coverage increased over time in microcosms polluted at 50 ng/ml of gentamicin similar to that shown by the 16S rRNA gene sequencing. However, no associated ARGs were identified in this bin. Therefore, the DNA from microcosms polluted at 50 ng/ml of gentamicin after 2-day exposure was sequenced by Oxford Nanopore to determine whether *Limnohabitans* contained genes related to antibiotic resistance within its genome. Two bins with more than 50% completion and less than 10% redundancy were identified as *Limnohabitans.* The first one had a coverage of 81.7%, a redundancy of 9.9%, 311 splits, and a length of 2.99 Mbp. The *aac(6')-Ib8* gene, which codes for the homonymous protein, an aminoglycoside acetyltransferase involved in gentamicin resistance, was identified in this MAG. The second bin had a coverage of 54.9%, a redundancy of 2.8%, a length of 4.45 Mbp, and 1423 splits. A hit from the *mexK* gene encoding for the MexK component of the MexJK multidrug efflux pump was identified in this MAG.

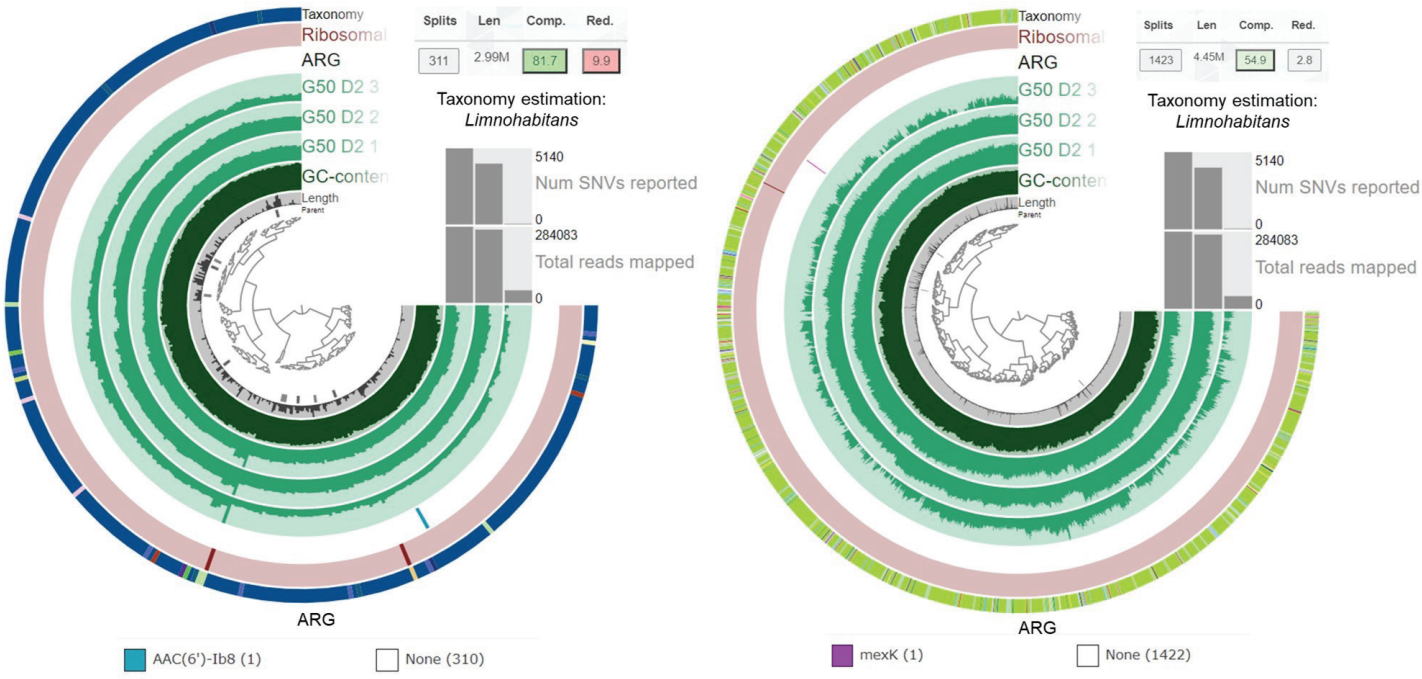


Figure 3. Two *Limnohabitans* MAGS co-assembled with Oxford Nanopore long and Illumina Miseq short metagenomic reads of DNA from water microcosms polluted at 50 ng/ml of gentamicin after 2-day exposure.

Gentamicin at 50 ng/ml increased the abundance and transcription of aminoglycoside resistance genes from the AAC(6') family, which were present in class 1 integrons cassettes

Aminoglycoside resistance genes found in sequenced class 1 integrons belonged to two main families, the *aadA* genes coding for aminoglycoside nucleotidyltransferases, which confer resistance to spectinomycin and streptomycin[48], and the *aac(6')* genes coding for *AAC*(6') acetyltransferases, which mediate resistance to several aminoglycosides including gentamicin[48] (Figure S3 in Supplementary Information). The 17 *aadA* genes and 15 *aac(6')* genes were found in class 1 integrons. While the relative abundance of the ensemble of *aadA* gene-related reads generally remained constant regardless of gentamicin concentration or exposure time, the relative abundance of *aac(6')* gene-related reads showed an increase in gentamicin-contaminated microcosms over time.

To avoid metagenomic biases, the relative abundance of these genes and their transcripts was determined by qPCR. Whereas in non-polluted controls and in microcosms polluted at 10 ng/ml of gentamicin the relative abundance of *aac(6')* genes and transcripts remains stable over time, both experience an increase over exposure time in microcosms polluted at 50 ng/ml of gentamicin (Figure 4). After 2-day exposure, the average relative abundance of *aac(6')* genes and transcripts in microcosms contaminated at 50 ng/ml was two orders of magnitude higher than that observed in non-polluted controls and samples contaminated at 10 ng/ml of gentamicin, although the response varied between triplicates. The gentamicin concentration decreased from an average of 52 ng/ml to 32 ng/ml after 2-day exposure. On the other hand, the relative abundance of *aac(6')* genes and transcripts was between 10 and 100 times higher at gentamicin concentration of 800 ng/ml than at 50 ng/ml (Table S5 in Supplementary Information).The relative abundance of *aadA* genes and transcripts remained relatively constant over exposure time and gentamicin concentration. A slightly higher abundance of *aadA* genes was observed over time at 10 and 50 ng/ml of gentamicin than in non-polluted controls, but this is not followed by a similar increase in transcription levels.

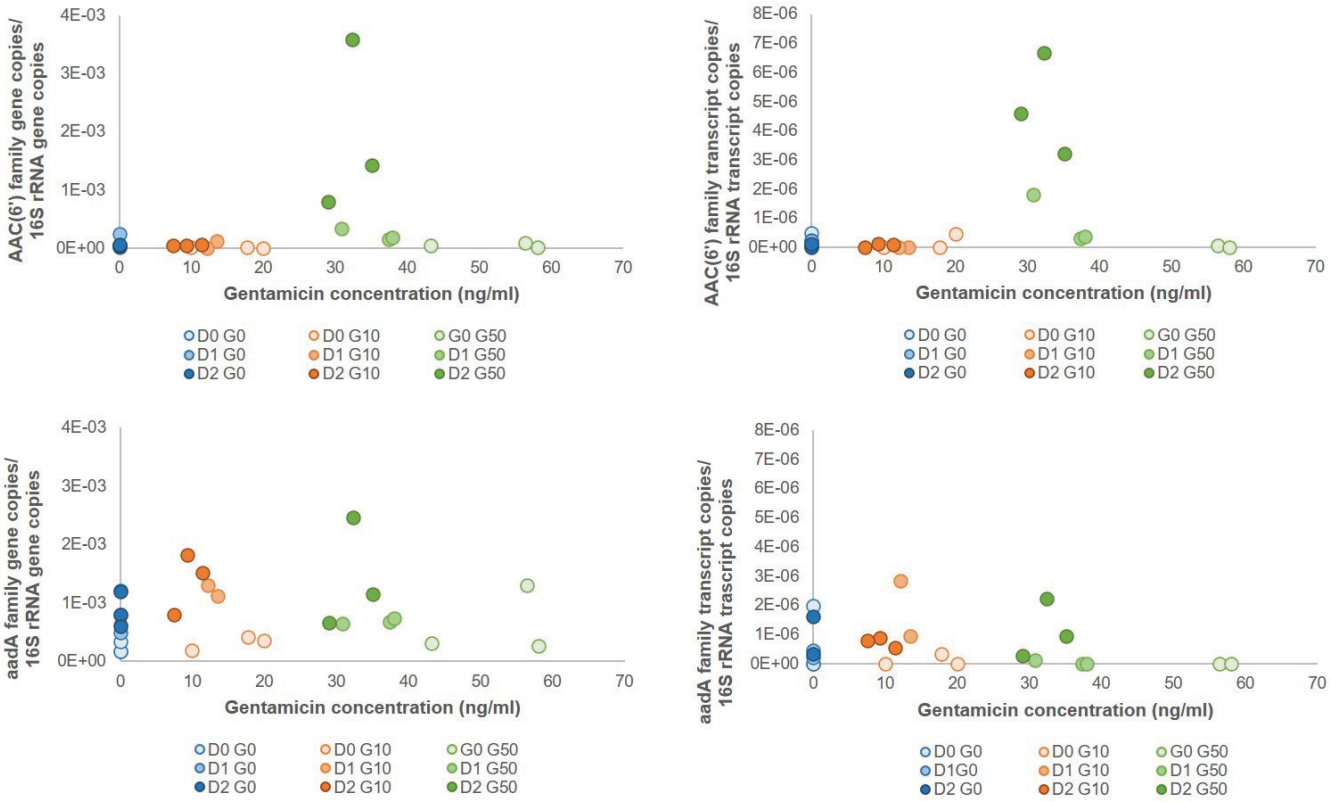


Figure 4. Dose-response linking sub-inhibitory gentamicin concentrations and relative abundance of gene copies (top left) and transcript copies (top right) of the genes from the *aac*(6') family and relative abundance of gene copies (bottom left) and transcript copies (bottom right) of the genes from the *aadA* family. Blue: 0 ng/ml of gentamicin. Orange: 10 ng/ml of gentamicin. Green: 50 ng/ml of gentamicin. Dot color gets darker with exposure time (DO; day 0, D1: day 1; D2; day 2). *aac*(6') family gene primers: qPCR efficiency=0.99; R^2 linearity coefficient =0.997. *aadA* family gene primers: qPCR efficiency=1; R^2 linearity coefficient =0.996.

Discussion

The aim of this study was to determine whether sub-inhibitory concentrations of gentamicin induced a response in the environmental microbiome and to evaluate the magnitude of that response as a function of antibiotic concentration. An increasing body of evidence suggests that sub-inhibitory concentrations of antibiotics might stimulate changes in the bacterial communities and their associated resistome and the links between antibiotic dose and community response in the environmental microbiome need to be elucidated[12,49]. The environmental impact of gentamicin at both inhibitory and sub-inhibitory concentrations determined *in vitro* was evaluated by assessing its effects on overall growth and transcription levels, the composition of total and active bacterial communities, and the abundance and transcription of ARGs in water microcosms.

The overall bacterial transcription levels increased over time without growth inhibition in river water with gentamicin at 50 ng/ml (putatively sub-inhibitory). This is consistent with previous findings that showed that sub-inhibitory concentrations of antibiotics induced changes in overall transcriptional levels *in vitro*[22,25,50]. This increase in transcription levels could be related to a higher energy consumption in sensitive bacteria and/or to an adaptive response of resistant bacteria. In any case, sub-inhibitory concentrations of gentamicin affected the bacterial transcriptome without inhibiting

overall growth. This effect on transcription demonstrated one of the mechanisms that sub-inhibitory antibiotic concentrations might have in diverse ecosystems.

In addition, subinhibitory gentamicin concentrations shifted the composition of both total and active communities with a drastic increase in *Limnohabitans* (Figure 2). Possibly gentamicin at subinhibitory concentrations favors different bacteria than at inhibitory concentrations due to differences in resistance levels and mechanisms. For example, *Pseudomonas,* which dominated at inhibitory concentrations where *Limnohabitans* were missing, might not be as competitive when *Limnohabitans* was not inhibited. Thus, *Limnohabitans* would dominate taking advantage of the possible inhibition of some competitors for nutrients. This is consistent with overall growth remaining high at these "subinhibitory" concentrations, although some members of the community are clearly not favored. Therefore, the use of the term "sub-inhibitory", which has been defined in single-cultures, might not be accurate to explain what takes place in complex microbial communities in their natural setting.

One of the *Limnohabitans* MAGs obtained from the hybrid assembly of short and long reads contained a putative gentamicin resistance gene (*aac(6')-Ib8*) and could explain in part its dominance at measurable, but generally subinhibitory gentamicin concentrations. This supports the hypothesis that *Limnohabitans* can actively respond to gentamicin, survive at sub-inhibitory concentrations and dominant sensitive bacteria. If these genes explain *Limnohabitans* lack of inhibition at 50 ng of gentamicin/ml, their lack of resistance at 800 ng/ml might be due the inefficiency of these genes to protect at such high concentrations. The hybrid assembly of short and long reads produced a more complete and better-defined vision of the genome of *Limnohabitans* than that using short reads alone. This approach should be routinely used for the resistome metagenomic analysis of complex communities, since sequence depth plays a crucial role in the identification of ARG[51,52] and long-read sequences increase the contiguity of plasmid assemblies[53].

The isolation of *Limnohabitans* from bacterial enrichments and the sequencing of its genome should provide a more complete reconstruction of its genome and a deeper analysis of the resistance mechanisms associated with this taxon. In addition, although the *aac(6')-Ib8* gene has been found in several human pathogens, such as *Enterobacter cloacae, Salmonella enterica, Klebsiella pneumoniae, Escherichia coli* or *Pseudomonas aeruginosa,* to the best of our knowledge, it has not been previously associated with *Limnohabitans.* However, the data gathered in this study does not differentiate between *Limnohabitans* from Rhône river water having the resistance gene or having acquired it in the laboratory through HGT. Further research should answer this question and provide a deeper understanding of both the genetic background associated with *Limnohabitans* dominance at sub-inhibitory gentamicin concentrations and the underlying mechanism(s) of gene acquisition by *Limnohabitans.*

Two genera seemed to survive at inhibitory concentration. First, *Pseudomonas* benefited from the high gentamicin dose although no specific gentamicin resistance gene was found in its reconstructed genome. Antibiotic resistant *Pseudomonas* have been identified in both the environment and in clinics[54,55] and its resistance is often related to multidrug resistance[56,57]. In this study, several efflux pumps conferring resistance to other antibiotics were found in the partial reconstruction of the *Pseudomonas* genome (Figure S2 in Supplementary Information), although the possibility that these pumps have some use against gentamicin has not been reported yet. However, these pumps or other mechanisms that have not been identified in this study, including gentamicin resistance genes, might be responsible for *Pseudomonas* survival at inhibitory concentrations of gentamicin. Second, the

increased activity of *Rhodoferax* without an increase in growth might indicate that it was consuming energy to survive in the presence of gentamicin and its activity was not related to growth but rather to survival under antibiotic pressure. Thus, *Rhodoferax* could possess a mechanism that provides moderate resistance to gentamicin, sufficient to enable its survival without promoting its growth. However, since its genome could not be assembled from metagenomic reads, there is no evidence to support that this genus had specific gentamicin resistance or any other antibiotic resistance mechanism.

Finally, aminoglycoside resistance genes belonging to the *aac(6')* acetyltransferase and the *aadA* nucleotidyltransferase families were identified in class 1 integrons cassette arrays. Although the relative abundance and transcription levels of *aac(6')* genes was lower at sub-inhibitory concentrations than in microcosms with 800 ng/ml of gentamicin (Table S5 in Supplementary Information), bacterial biomass was significantly reduced in microcosms at this inhibitory concentration. Therefore, the selective pressure imposed by inhibitory concentrations of gentamicin inhibited most of the members of the community and resulted in a higher abundance of gentamicin resistance genes per 16S rRNA genes, whereas gentamicin at 50 ng/ml selected for *aac(6')* genes in water microcosms without inhibiting overall bacterial growth. Therefore, gentamicin at sub-inhibitory concentrations may select for gentamicin resistance genes in the environment and contribute to the burden of antibiotic resistance. On the other hand, the slightly higher abundances of *aadA* family in the presence of gentamicin could be a consequence a co-selection of these genes with *aac(6')* genes located in the same integron cassettes. Since *aadA* genes did not show an active response to gentamicin addition, they did not appear to confer resistance to gentamicin, although they are often found in class 1 integrons[58].

In conclusion, this study was, to the best of our knowledge, the first to show that sub-inhibitory concentrations of antibiotics induced a response in the environmental microbiome and resistome *in situ* and to establish relationships between antibiotic dose and the selective force for aminoglycoside resistance genes. In addition, the genes that increased their abundance and transcription at sub-inhibitory gentamicin concentrations were present in class 1 integrons and might be mobilized and disseminated across the environmental microbiome and into the human and animal microbiomes. These results support the concern that sub-inhibitory concentrations of antibiotics affect resistance profiles in the environment[18].

This study was based on a single environmental model and a single antibiotic. Future studies should compare the response of different environments to the same antibiotic and establish patterns between the composition of the environmental matrix and/or environmental conditions and the observed response. In addition, a combination of antibiotics at sub-inhibitory concentrations might reflect more accurately the situation occurring in freshwater environments receiving treated wastewater.Therefore, studies analyzing the effect of combined antibiotics at sub-inhibitory concentrations on the environmental microbiome and resistome are needed to fill knowledge gaps and provide a more realistic risk assessment. In addition, this study analyzed a short-term environmental response to gentamicin, but did not provide any results concerning the response at longer exposure times nor the persistence of resistance without antibiotic pressure. These elements are key to gain a more thorough understanding of the risk posed by sub-inhibitory concentrations of antibiotics and should be evaluated in the future.

Additional research

Elucidating the resistance mechanisms associated to Limnohabitans: perspectives

The results obtained from this study show an increase of abundance of *Limnohabitans* over exposure time and an increased abundance and transcription of gentamicin resistance genes at concentrations that are sub-inhibitory in culture (50 ng/ml). In addition, a gentamicin resistance gene was identified in a MAG estimated to be *Limnohabitans.* These results supported the hypothesis that *Limnohabitans* is resistant to gentamicin at 50 ng/ml and can benefit from the death of susceptible members of the water bacterial community and use the available nutrients to grow in presence of the antibiotic. These results supported the hypothesis that gentamicin at sub-inhibitory concentrations selects for ARB and ARGs in the environment. However, the data gathered in this study was not sufficient to confirm this hypothesis. This section discusses the approaches carried out to analyze the MAG of *Limnohabitans,* its associated resistome and mobilome as well as to discuss perspectives and approaches that should be used to help answer this question.

Evaluation of the genetic context of the *aac(6')-Ib8* gene identified in the *Limnohabitans* bin

After co-assembling long and short metagenomic reads, a *Limnohabitans* MAG containing a hit for a gentamicin resistance gene, *aac(6')-Ib8*, was obtained. However, the sequence that aligned the gene when contigs where blasted against the CARD database had only 132 nucleotides, whereas the complete gene is 678 nucleotides long. Thus, this hit was obtained from a partial protein alignment (61% identity) of the first 39 amino acids of the *aac(6')-Ib8* gene, as shown on Figure LH1. The contig containing the sequence identified as the *aac(6')-Ib8* was inspected to analyze its genetic context (Figure LH2). A gap was found downstream from the 132 nucleotide bases that aligned the *aac(6')-Ib8* gene and the region downstream the gap was blasted against the CARD database in order to determine whether it contained another fraction of the gene: the sequence did not blast against any gene on the CARD database. Secondly, the 132-nucleotide sequence identified as an *aac(6')-Ib8* hit was blasted against a set of *Limnohabitans* genomes and assembled genomes downloaded from the NCBI database. However, no alignments were found. Finally, the set of *Limnohabitans* genomes and assembled genomes was blasted against the CARD database in order to identify the ARG associated to publicly available *Limnohabitans* sequences. No aminoglycoside resistance genes were found in these genomes: hits were obtained from the *ugd* gene, involved in peptide antibiotic resistance, and the *qacH* gene, a gene codifying for a subunit of the qac multidrug efflux pump, commonly associated with class 1 integrons. Therefore, the presence of the *aac(6')-Ib8* gene in the *Limnohabitans* genome could not be confirmed using these approaches. In addition, the *aac(6')-Ib8* gene has been found in several *Enterobacteriaceae,* such as *Enterobacter cloacae, Enterobacter hormaechei, Salmonella enterica, Klebsiella pneumoniae, Escherichia coli, Pseudomonas aeruginosa* and *Pseudomonas putida*, and in some genera from the *Burkholderiales* order belonging to the *Achromobacter, Delftia, Sphingomonas* and *Sphingobium* genera, but it has not been previously identified in *Limnohabitans.* Thus, the presence of the *aac(6')-Ib8* gene in the *Limnohabitans* MAG identified by metagenomic sequencing could mean a recent acquisition of the aminoglycoside resistance gene by *Limnohabitans* under gentamicin pressure, but it could as well be a sequencing artifact, or *Limnohabitans* could already have acquired the gene before the beginning of the experiment. The metagenomic sequencing carried out in this study does not provide enough information to confirm or exclude the hypothesis that the *Limnohabitans* contains a gentamicin resistance gene nor to determine whether it has been acquired through HGT during the experiment.

```
 Score = 51.6 bits (122),  Expect = 6e-11, Method: Compositional matrix adjust.
 Identities = 24/39 (62%), Positives = 31/39 (79%), Gaps = 0/39 (0%)
 Frame = +1

Query  1    MPIKPGPKPIAKSTGKPDKRRRDNKETPKNTPDLKPHKH  117
            M +KPGPK IA+STG+PD+R+RDNK+TP NT  L   K+
Sbjct  1    MSLKPGPKRIAESTGQPDQRQRDNKKTPGNTDKLGITKY  39
```

Figure LH1. Protein alignment between the sequence from co-assembled contigs identified as the *aac(6')-Ib8* gene (query) and the sequence of the *aac(6')-Ib8* gene obtained from the CARD database (sbjct).

In addition, the genes that were present in the same contig as the 132-nucleotide sequence were inspected. Only one matched the nr database. Probably the putative genes that did not match were only partly sequenced and the sequenced region was not sufficient to identify the genes. On the other hand, a site-specific DNA recombinase related to the DNA invertase Pin was located near the gene (green arrow in Figure LH2). It was commonly associated to species from the *Burkholderiaceae* family, but no prior association to *Limnohabitans* was described. However, when extracting the sequence identified as the recombinase from the assembled contigs and blasting it against the set of genomes and assembled genomes from *Limnohabitans,* it aligned at 93.5% identity and 1003 nucleotides length to a *Limnohabitans parvus* genome and at 74.2% identity and 508 nucleotides length to a *Limnohabitans sp. Hippo4 LimB-Hippo4-C7* genome. Thus, *Limnohabitans* genomes could encode an homologous of that protein that could have mediated the recombination of the *aac(6')-Ib8* gene in its genome -assuming that the gene is actually in the *Limnohabitans* genome-. However, the gathered information is insufficient to confirm this hypothesis. A PCR amplification and sequencing of the regions flanking the 132 nucleotide sequence identified as part of the *aac(6')-Ib8* gene could help elucidating whether the whole gene is present in water microcosms and whether its flanked by recombinases or by MGEs such as class 1 integrons, since the gene seems to be present in integrons from the water environment according to the metagenomic sequencing of the integron cassettes. However, determining whether the sequenced regions belong to *Limnohabitans might* not be straightforward. Furthermore, no DNA remains from microcosms polluted at 50 ng/ml of gentamicin at 2-day exposure. Thus, the isolation and sequencing of *Limnohabitans* from enriched media could be an alternative approach to answer the questions raised during this analysis.

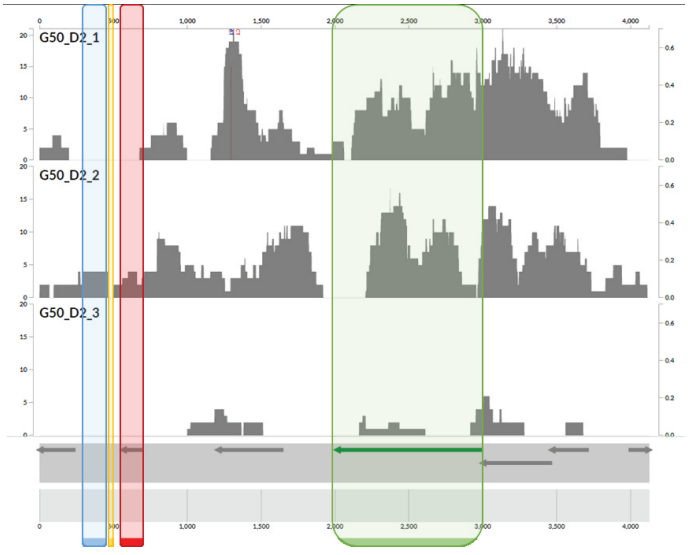


Figure LH2. Contig containing the sequence identified as the *aac(6')-Ib8* gene. Red: 132-nucleotide region that aligned the *aac(6')-Ib8* gene. Green: recombinase related to the DNA invertase Pin. Yellow: gap downstream the 132-nucleotide sequence. Blue: region downstream the gap that was blasted against the CARD database.

Limnohabitans isolation from bacterial enrichments and metagenomic sequencing

In order to understand why *Limnohabitans* increased its abundance in the total and active communities of water microcosms 50 ng/ml of gentamicin, to determine whether the *aac(6')-Ib8* gene its present in its genome and to identify the possible recombinases or MGEs, *Limnohabitans* should be isolated from enriched media and sequenced. Since the sequencing carried out on water microcosms was not sufficient to obtain the whole genome of *Limnohabitans,* isolating it from the rest of environmental microorganisms could improve the sensitivity of metagenomic techniques and provide a whole genome reconstruction. A preliminary approach consisted of isolating colonies from R2A agar plates inoculated with the samples that were used to determine gentamicin effect on water and sequencing their 16S rRNA gene using the Sanger method. However, although *Limnohabitans* grows well on R2A medium, *Arthrobacter* was largely predominant in water bacterial enrichments (Figure LH3) and it was the only genus identified from 6 different colonies. Therefore, a more thorough screening of *Limnohabitans* from the colonies plated on R2A agar should be carried out to increase the probability of sequencing the genome of *Limnohabitans.* Alternatively, *Limnohabitans* acclimatation and enrichment in increasing doses of NSY medium[59], which is commonly used to culture this genus[60,61], could be carried out if no *Limnohabitans* colonies are obtained from R2A medium.

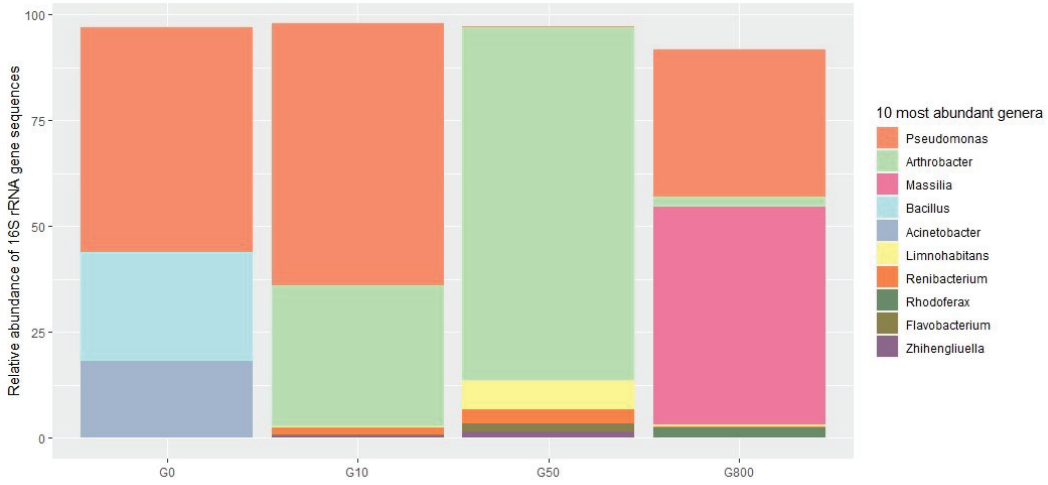


Figure LH3. Average relative abundance of the 16S rRNA gene of the 10 most abundant genera identified in bacterial enrichments. DNA was extracted from bacterial enrichments incubated for 48 hours at 29°C in R2A medium. Then, the 16S rRNA gene was sequenced and analyzed as explained in the Materials and Methods section. PCR amplification was performed in 25 cycles instead of 30. n=3.

Evaluation of the relative abundance of class 1 integrons and clinical class 1 integrons

qPCR of class 1 integrons

In order to determine whether the selection for aminoglycoside resistance genes was related to an increase in the relative abundance of integrons, a qPCR was carried out using HS463a/HS464 primers (information about this primers can be found in the Materials and Methods section). Standards for qPCR assays were obtained from water DNA, cloned and transformed using the TOPO TA cloning Kit (Invitrogen) and normalized to 10^8 copies/µl. Standard curves were made using 10-fold serial dilutions of the standards (10^7 – 10^2 copies/µl). Amplification was performed as explained in the Materials and Methods section. The reaction had an efficiency of 1.02 and a R^2 coefficient of 0.98. Results were submitted to a Leave-One-Out cross-validation, and no significant increase over was detected on the relative abundance of the *int1* gene in any of the gentamicin-polluted microcosms (Figure Int1). However, the integration rate of aminoglycoside resistance genes could increase under gentamicin pollution without impacting the abundance of *int1* genes. On the other hand, the increase of the relative abundance of aminoglycoside resistance genes observed in Figure 8 was not necessarily related to genes belonging to integron cassettes. Given that the sequencing of integrons carried out in this study was based on previously amplified products, calculations on the integration rate of aminoglycoside resistance genes -as performed in Figure S3 (Supplementary Information)- was not accurate. Therefore, with the elements analysed in this study, the question remains whether there was an increase of aminoglycoside resistance gene integration under gentamicin pollution. However, with the sequencing of class 1 integrons, at least some of the genes that are being selected for were shown to be present in class 1 integron cassette arrays.

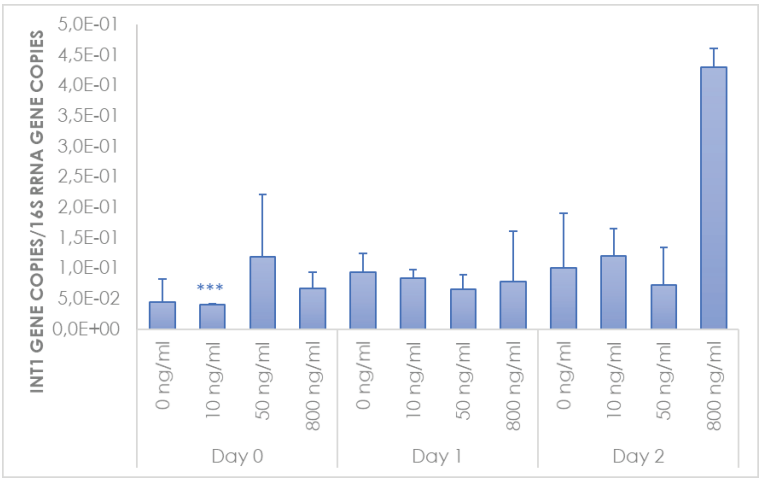


Figure Int1. Number of copies of the *int1* gene per µl of qPCR reaction normalized by the number of copies of the 16S rRNA gene per µl of reaction in water microcosms. The *int1* gene was amplified from water microcosms using HS463a/HS464 primers. Results were submitted to a Leave-One-Out cross-validation. ***: p-value > 0.001. n=3.

qPCR of clinical class 1 integrons

A second question that was raised during the analysis of these results was whether the relative abundance of clinical 1 integron-integrase genes *(intI1)* showed a different tendency despite that the relative abundance of class 1 integron-integrase genes *(int1)* did not increase under gentamicin pollution over time. If the abundance of clinical class 1 integrons increased over time in gentamicin-polluted microcosms, this would be an indicative of potential risks for human health. Therefore, a qPCR was carried out using the primers intI1F165 (5'-CGA ACGAGTGGCGGAGGGTG-3') and intI1R476 (5'-TAC CCGAGAGCTTGGCACCCA-3'), as described by Gillings *et al.*[62] The efficiency of the reaction was of 0.95 and the R^2 linearity coefficient was of 0.998 (Figure Int2). However, all samples at day 0 and microcosms polluted at 800 ng/ml after 1-day exposure had to be excluded from the analysis, since they presented 2 curves in the melting curve analysis (Figure Int3). Therefore, this analysis could not be considered as conclusive, since it lacked accurate information for more than a third of the samples.

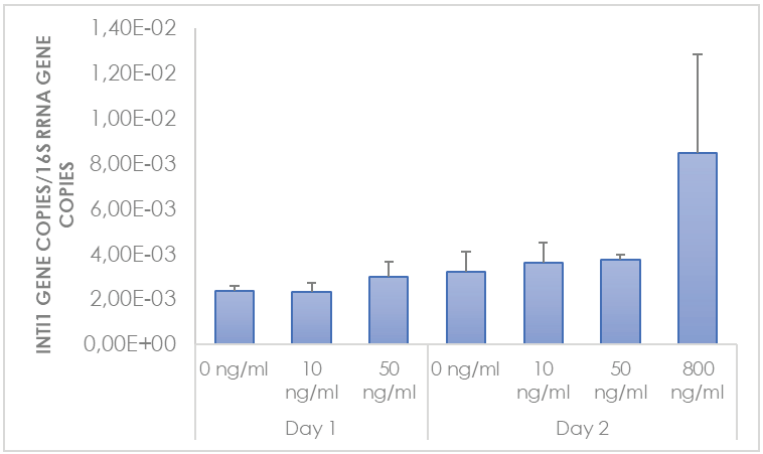


Figure Int2. Number of copies of the *intI1* gene per µl of qPCR reaction normalized by the number of copies of the 16S rRNA gene per µl of reaction in water microcosms. The *intI1* gene was amplified from water microcosms using intI1F165 and intI1R476 primers. Results were submitted to a Leave-One-Out cross-validation. No significant differences were found. n=3.

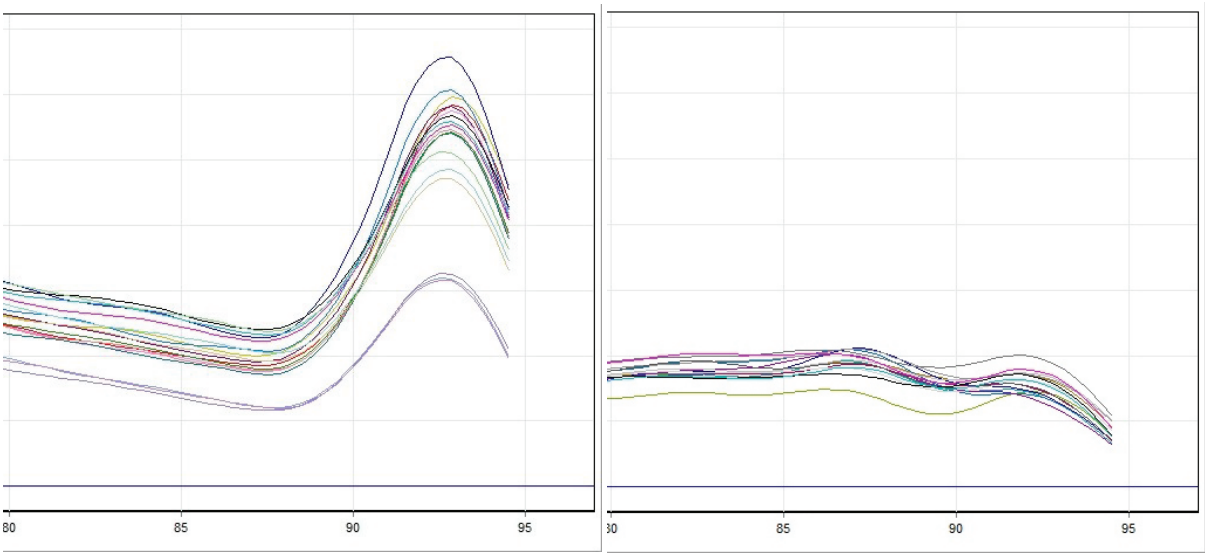


Figure Int3. Melting curve of *intI1* gene standards (left) and samples at day 0 and samples polluted at 800 ng/ml after 1-day exposure (right), from 80 to 95°C. Lower temperatures have been excluded from the pictures for better visualization of the curves.

References

1. Kemper, N. Veterinary antibiotics in the aquatic and terrestrial environment. *Ecol. Indic.* **8,** 1–13 (2008).

2. Pérez-Cobas, A. E. *et al.* High-throughput quantification of antibiotic resistance genes from an urban wastewater treatment plant. *FEMS Microbiol. Ecol.* **92,** (2016).

3. Li, B., Ju, F., Cai, L. & Zhang, T. Profile and Fate of Bacterial Pathogens in Sewage Treatment Plants Revealed by High-Throughput Metagenomic Approach. *Environ. Sci. Technol.* **49,** 10492–10502 (2015).

4. Cai, L. & Zhang, T. Detecting human bacterial pathogens in wastewater treatment plants by a high-throughput shotgun sequencing technique. *Environ. Sci. Technol.* **47,** 5433–5441 (2013).

5. Zhang, X. X., Zhang, T. & Fang, H. H. P. Antibiotic resistance genes in water environment. *Appl. Microbiol. Biotechnol.* **82,** 397–414 (2009).

6. Li, B., Zhang, T., Xu, Z. & Fang, H. H. P. Rapid analysis of 21 antibiotics of multiple classes in municipal wastewater using ultra performance liquid chromatography-tandem mass spectrometry. *Anal. Chim. Acta* **645,** 64–72 (2009).

7. Rodriguez-Mozaz, S. *et al.* Occurrence of antibiotics and antibiotic resistance genes in hospital and urban wastewaters and their impact on the receiving river. *Water Res.* (2015). doi:10.1016/j.watres.2014.11.021

8. Gudda, F. O. *et al.* Antibiotic-contaminated wastewater irrigated vegetables pose resistance selection risks to the gut microbiome. *Environ. Pollut.* **264,** 114752 (2020).

9. Danner, M., Robertson, A., Behrends, V. & Reiss, J. Antibiotic pollution in surface fresh waters : Occurrence and effects. *Sci. Total Environ.* **664,** 793–804 (2019).

10. Jechalke, S., Heuer, H., Siemens, J., Amelung, W. & Smalla, K. Fate and effects of veterinary antibiotics in soil. *Trends Microbiol.* **22,** 536–545 (2014).

11. Kalasseril, S., Paul, R., J, R. K. V & Pillai, D. Investigating the impact of hospital antibiotic usage on aquatic environment and aquaculture systems : A molecular study of quinolone resistance in Escherichia coli. *Sci. Total Environ.* **748,** 141538 (2020).

12. Ashbolt, N. J. Human Health Risk Assessment (HHRA) for Environmental Development and Transfer of Antibiotic Resistance. *Environ. Health Perspect.* **121,** 993–1002 (2013).

13. Bengtsson-Palme, J., Kristiansson, E. & Larsson, D. G. J. Environmental factors influencing the development and spread of antibiotic resistance. *FEMS Microbiol. Rev.* 68–80 (2017). doi:10.1093/femsre/fux053

14. Manaia, C. M. Assessing the Risk of Antibiotic Resistance Transmission from the Environment to Humans: Non-Direct Proportionality between Abundance and Risk. *Trends in Microbiology* **25,** (2017).

15. Manaia, C. M., Macedo, G., Fatta-Kassinos, D. & Nunes, O. C. Antibiotic resistance in urban aquatic environments: can it be controlled? *Appl. Microbiol. Biotechnol.* **100,** 1543–1557 (2016).

16. Durso, L. M. & Cook, K. L. Impacts of antibiotic use in agriculture : what are the benefits and risks ? *Curr. Opin. Microbiol.* **19,** 37–44 (2014).

17. Almakki, A., Jumas-Bilak, E., Marchandin, H. & Licznar-Fajardo, P. Antibiotic resistance in urban runoff. *Sci. Total Environ.* **667,** 64–76 (2019).

18. Andersson, D. I. & Hughes, D. Microbiological effects of sublethal levels of antibiotics. *Nat. Rev. Microbiol.* **12,** 465–478 (2014).

19. Gullberg, E. *et al.* Selection of resistant bacteria at very low antibiotic concentrations. *PLoS Pathog.* **7,** 1–9 (2011).

20. Murray, A. K. *et al.* Novel Insights into Selection for Antibiotic Resistance in Complex Microbial Communities. *MBio* **9,** 1–12 (2018).

21. Chow, L., Waldron, L. & Gillings, M. R. Potential impacts of aquatic pollutants: Sub-clinical antibiotic concentrations induce genome changes and promote antibiotic resistance. *Front. Microbiol.* **6,** 1–10 (2015).

22. Bruchmann, J., Kirchen, S. & Schwartz, T. Sub-inhibitory concentrations of antibiotics and wastewater influencing biofilm formation and gene expression of multi-resistant Pseudomonas aeruginosa wastewater isolates. *Environ. Sci. Pollut. Res.* **20,** 3539–3549 (2013).

23. Gullberg, E., Albrecht, L. M., Karlsson, C., Sandegren, L. & Andersson, D. I. Selection of a Multidrug Resistance Plasmid by Sublethal Levels of Antibiotics and Heavy Metals. *mBio* **5,** 19–23 (2014).

24. Choung, S. *et al.* Transfer of antibiotic resistance plasmids in pure and activated sludge cultures in the presence of environmentally representative micro-contaminant concentrations. *Sci. Total Environ.* **468–469,** 813–820 (2014).

25. Shun-Mei, E. *et al.* Sub-inhibitory concentrations of fluoroquinolones increase conjugation frequency. *Microb. Pathog.* **114,** 57–62 (2018).

26. Jutkina, J., Rutgersson, C., Flach, C. F. & Joakim Larsson, D. G. An assay for determining minimal concentrations of antibiotics that drive horizontal transfer of resistance. *Sci. Total Environ.* **548–549,** 131–138 (2016).

27. Jutkina, J., Marathe, N. P., Flach, C. F. & Larsson, D. G. J. Antibiotics and common antibacterial biocides stimulate horizontal transfer of resistance at low concentrations. *Sci. Total Environ.* **616–617,** 172–178 (2018).

28. Murray, A. K. *et al.* Novel insights into selection for antibiotic resistance in complex microbial communities. *MBio* **9,** 1–12 (2018).

29. Le-minh, N., Khan, S. J., Drewes, J. E. & Stuetz, R. M. Fate of antibiotics during municipal water recycling treatment processes. *Water Res.* **44,** 4295–4323 (2010).

30. George, J. & Halami, P. M. Sub-inhibitory concentrations of gentamicin triggers the expression of aac(6′)Ie-aph(2″)Ia, chaperons and biofilm related genes in Lactobacillus plantarum MCC 3011. *Res. Microbiol.* **168,** 722–731 (2017).

31. Zhang, A. N. *et al.* Conserved phylogenetic distribution and limited antibiotic resistance of class 1 integrons revealed by assessing the bacterial genome and plasmid collection. *Microbiome* **6,** 1–14 (2018).

32. Gillings, M. R. Integrons: Past, Present, and Future. *Microbiol. Mol. Biol. Rev.* **78,** 257–277 (2014).

33. Muyzer, G., Hottentrager, S., Teske, A. & Wawer, C. Denaturing gradient gel electrophoresis of PCR-amplified 16S rDNA—a new molecular approach to analyse the genetic diversity of mixed microbial communities. In *Molecular microbial ecology manual;* (eds. Akkermans, A., van Elsas, J. & de Bruijn, F.) 1–23 (Kluwer Academic Publishers, 1995).

34. Watanabe, K., Kodama, Y. & Harayama, S. Design and evaluation of PCR primers to amplify bacterial 16S ribosomal DNA fragments used for community fingerprinting. *J. Microbiol. Methods* **44,** 253–262 (2001).

35. Klindworth, A. *et al.* Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res.* **41,** 1–11 (2013).

36. Masella, A. P., Bartram, A. K., Truszkowski, J. M., Brown, D. G. & Neufeld, J. D. PANDAseq: Paired-end assembler for illumina sequences. *BMC Bioinformatics* **13,** 31 (2012).

37. Wang, Q., Garrity, G. M., Tiedje, J. M. & Cole, J. R. Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* **73,** 5261–5267 (2007).

38. Minoche, A. E., Dohm, J. C. & Himmelbauer, H. Evaluation of genomic high-throughput sequencing data generated on Illumina HiSeq and Genome Analyzer systems. *Genome Biol.* **12,** (2011).

39. Li, D., Liu, C. M., Luo, R., Sadakane, K. & Lam, T. W. MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* **31,** 1674–1676 (2015).

40. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9,** 357–359 (2012).

41. Eren, A. M. *et al.* Anvi'o: An advanced analysis and visualization platformfor 'omics data. *PeerJ* **2015,** 1–29 (2015).

42. Alcock, B. P. *et al.* CARD 2020: Antibiotic resistome surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids Res.* **48,** D517–D525 (2020).

43. Bowers, R. M. *et al.* Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. *Nat. Biotechnol.* **35,** 725–731 (2017).

44. Wick, R. R., Judd, L. M., Gorrie, C. L. & Holt, K. E. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput. Biol.* **13,** 1–22 (2017).

45. Holmes, A. J. *et al.* The gene cassette metagenome is a basic resource for bacterial genome evolution. *Environ. Microbiol.* **5,** 383–394 (2003).

46. Gillings, M. R., Xuejun, D., Hardwick, S. A., Holley, M. P. & Stokes, H. W. Gene cassettes encoding resistance to quaternary ammonium compounds: A role in the origin of clinical class 1 integrons? *ISME J.* **3,** 209–215 (2009).

47. Benjamin Buchfink, Xie, C. & Huson, D. H. Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* **12,** 59–60 (2015).

48. Ramirez .S, M. & Tolmasky .E, M. Aminoglycoside Modifing Enzymes. *Drug Resist. Updat.* **13,** 151–171 (2011).

49. Ben, Y. *et al.* Human Health Risk Assessment of Antibiotic Resistance Associated with Antibiotic Residues in the Environment: A Review. *Environ. Res.* **169,** 483–493 (2018).

50. Goh, E. B. *et al.* Transcriptional modulation of bacterial gene expression by subinhibitory concentrations of antibiotics. *Proc. Natl. Acad. Sci. U. S. A.* **99,** 17025–17030 (2002).

51. Gweon, H. S. *et al.* The impact of sequencing depth on the inferred taxonomic composition and AMR gene content of metagenomic samples. *Environ. Microbiomes* **14,** 1–15 (2019).

52. Zaheer, R. *et al.* Impact of sequencing depth on the characterization of the microbiome and resistome. *Sci. Rep.* **8,** 1–11 (2018).

53. George, S. *et al.* Resolving plasmid structures in enterobacteriaceae using the MinION nanopore sequencer: Assessment of MinION and MinION/illumina hybrid data assembly approaches. *Microb. Genomics* **3,** 1–8 (2017).

54. Peter, S. *et al.* Tracking of Antibiotic Resistance Transfer and Rapid Plasmid Evolution in a Hospital Setting by Nanopore Sequencing. *mSphere* **5,** 1–16 (2020).

55. Nonaka, L., Inubushi, A., Shinomiya, H., Murase, M. & Suzuki, S. Differences of genetic diversity and antibiotics susceptibility of Pseudomonas aeruginosa isolated from hospital, river and coastal seawater. *Environ. Microbiol. Rep.* **2,** 465–472 (2010).

56. Narciso-da-Rocha, C. & Manaia, C. M. Multidrug resistance phenotypes are widespread over different bacterial taxonomic groups thriving in surface water. *Sci. Total Environ.* **563–564,** 1–9 (2016).

57. Pumbwe, L. & Piddock, L. J. V. Two Efflux Systems Expressed Simultaneously in Multidrug-Resistant Pseudomonas aeruginosa. *Antimicrob. Agents Chemother.* **44,** 2861–2864 (2000).

58. Sultan, I. *et al.* Antibiotics, Resistome and Resistance Mechanisms: A Bacterial Perspective. *Front. Microbiol.* **9,** (2018).

59. Hahn, M. W. *et al.* Isolation of Novel Ultramicrobacteria Classified as Actinobacteria from Five Freshwater Habitats in Europe and Asia. *Appl. Environ. Microbiol.* **69,** 1442–1451 (2003).

60. Jezbera, J., Šimek, K. & Hahn, M. W. Limnohabitans planktonicus sp.nov., and Limnohabitans parvus sp . nov ., two novel planktonic Betaproteobacteria isolated from a freshwater reservoir. *Int. J. Syst. Evol. Microbiol.* **60,** 2710–2714 (2010).

61. Kasalický, V., Jezbera, J., Hahn, M. W. & Šimek, K. The Diversity of the Limnohabitans Genus, an Important Group of Freshwater Bacterioplankton, by Characterization of 35 Isolated Strains. *PLoS One* **8,** (2013).

62. Gillings, M. R. *et al.* Using the class 1 integron-integrase gene as a proxy for anthropogenic pollution. *ISME J.* **9,** 1269–1279 (2015).

Chapter V. Environmental and anthropogenic factors shape the snow microbiome and the antibiotic resistome

Abstract

Winter tourism can generate environmental pollution and affect microbial ecology in mountain ecosystems. This could stimulate the development of antibiotic resistance in snow and its dissemination through the atmosphere and through snow melting. Despite these potential impacts, the effect of winter tourism on ecosystem health and functioning remains to be elucidated. In this study, snow samples subjected to different levels of anthropogenic activities and surrounding vegetation were obtained from the Karkonosze National Park in the Sudety Mountains in Poland and the impact of vegetation and anthropogenic activity on the snow microbiome and resistome was evaluated using a metagenomics approach. We hypothesized that both surrounding vegetation and anthropogenic pollution generate an additional input of organic carbon, which supports the growth of a larger bacterial community and stimulates competition and ARG proliferation. DNA was extracted from filtered snow and the bacterial abundance was estimated by qPCR amplification of the 16S rRNA gene. In addition, bacterial community composition, genus richness and diversity were evaluated by the sequencing of the V3-V4 hypervariable region of the 16S rRNA gene. The size and composition of the antibiotic resistome and the ARG richness and diversity were analyzed using shotgun sequencing of total extracted DNA (metagenomics) sequences. Finally, metagenomic sequences were co-assembled to evaluate the resistome of assembled contigs. Both vegetation and anthropogenic effects increased bacterial numbers, induced changes in the antibiotic resistome, and increased the abundance of antibiotic resistance genes (ARGs). In addition, sites with surrounding vegetation showed a less diverse bacterial population with an increased abundance of bacteria that use the higher levels of organic carbon. This is the first study to show that anthropogenic activity influences the antibiotic resistome in alpine snow and to provide evidence of the impact of environmental and anthropogenic factors on snow microbial ecology, while highlighting the need to survey antibiotic resistance development in anthropogenically polluted sites.

Introduction

Mountains, and especially their snow cover, are sensitive indicators of climate change. Mountains support roughly one-third of all land-dwelling species and supply water for nearly half the global population[1]. Mountains are warming at rates similar to the Arctic[2] with impact on the environment (glacier shrinking, permafrost degradation, floods, changes in the distribution of species and ecosystems) and society (summer and winter tourism). One major source of uncertainty is the extent to which human activities interact with climatic factors to modify biogeochemical processes and ecosystem health and functioning. Among the activities known to affect alpine ecosystems, tourism[3–5] can have both positive and negative effects on mountain ecosystems, communities and economies[6]. Tourism activities often involve the development and intense use of tracks, paths and sport slopes by vehicles, non-motorized transport and pedestrian traffic. Visitor presence is also usually concentrated in small areas and contribute to increased noise and waste. The negative environmental effects of tourism can include devegetation, soil erosion, alteration of critical landscapes and water flows, water and air pollution, and wildlife relocation and behavioral changes[6]. The introduction of exotic and invasive species and diseases can also have a significant negative impact on the environment[7].

The snow cover on mountains is connected to the atmosphere, which is considered one of the main sources of microorganisms in snow and is responsible for the transport of atmospheric dust and its

inclusion in snow[8]. ARGs and pathogens contaminating the atmosphere can be deposited on Earth through snowfall, which could increase their dissemination[9,10]. In addition, microorganisms present in snow are better adapted to atmospheric transport that other microorganisms[11] and antibiotic resistant bacteria from contaminated snow could potentially be transported to remote environments. ARB and ARGs could also be transferred to freshwater through snow melting[12] and increase the risk of waterborne disease.

Despite these potential impacts, the effect of winter tourism on ecosystem health and functioning remains to be elucidated. The objective of this study was to address this lack of information by studying the impact of winter tourism on snow-covered mountain ecosystems. We hypothesized that anthropogenic activity would increase waste in snow environments and would support the growth of a more abundant community while stimulating interspecies competition and ARG production. Anthropogenic activity might shape the antibiotic resistome without necessarily changing snow bacterial communities, since human microbiome bacteria would be unlikely to survive in snow. We also hypothesized that surrounding vegetation would also affect the snow microbiome and resistome through bacterial growth on plant-derived organic matter. This increased nutrient availability would in turn stimulate competition and ARG production. In order to test this hypothesis, snow samples were obtained from two watersheds located in the Karkonosze National Park in the Sudety Mountains in Poland: an unaffected Czarny Kociol Jagniatkowski catchment and a Kociol Malego Stawu catchment with well documented human activities such as tourism and cottage development. Samples with different levels of surrounding vegetation were recovered from both catchments. Vegetation and anthropogenic activity effects on the snow microbiome (bacterial community numbers and composition, genus richness and diversity) and resistome (size and composition of the antibiotic resistome and ARG richness and diversity) were evaluated using metagenomics and 16SrRNA gene analyses.

Materials and methods

Snow sampling

Snow samples were obtained from two watersheds located in the Karkonosze National Park in the Sudety Mountains in Poland (Figure 1): an unaffected Czarny Kociol Jagniatkowski catchment and a Kociol Malego Stawu catchment with well documented human activities such as tourism and cottage development.

Samples with different levels of surrounding vegetation were recovered from both catchments. NP 1-10 samples from the Czarny Kociol Jagniatkowski catchment were obtained from open spaces, whereas NP 11-20 were sampled next to trees. S samples were obtained from paths in the Kociol Malego Stawu catchment with frequent human transit, SF were obtained from the surrounding forest areas, and LB were sampled from the highest point of the catchment in the open, less transited and windier areas with no surrounding vegetation. These samples, therefore, cover a range of human and vegetation interaction (Figure 2).

Samples were taken in March 2020 and left at room temperature until melted. The microbial fraction of snow was recovered by filtering melt water through 0.2 µm Nucleopore membranes (Whatman). Filters were then frozen at -20°C until DNA extraction. The coordinates and elevation of each sample, as well as the volume of sample that was filtered are shown in Table S1 (Supplementary Information).

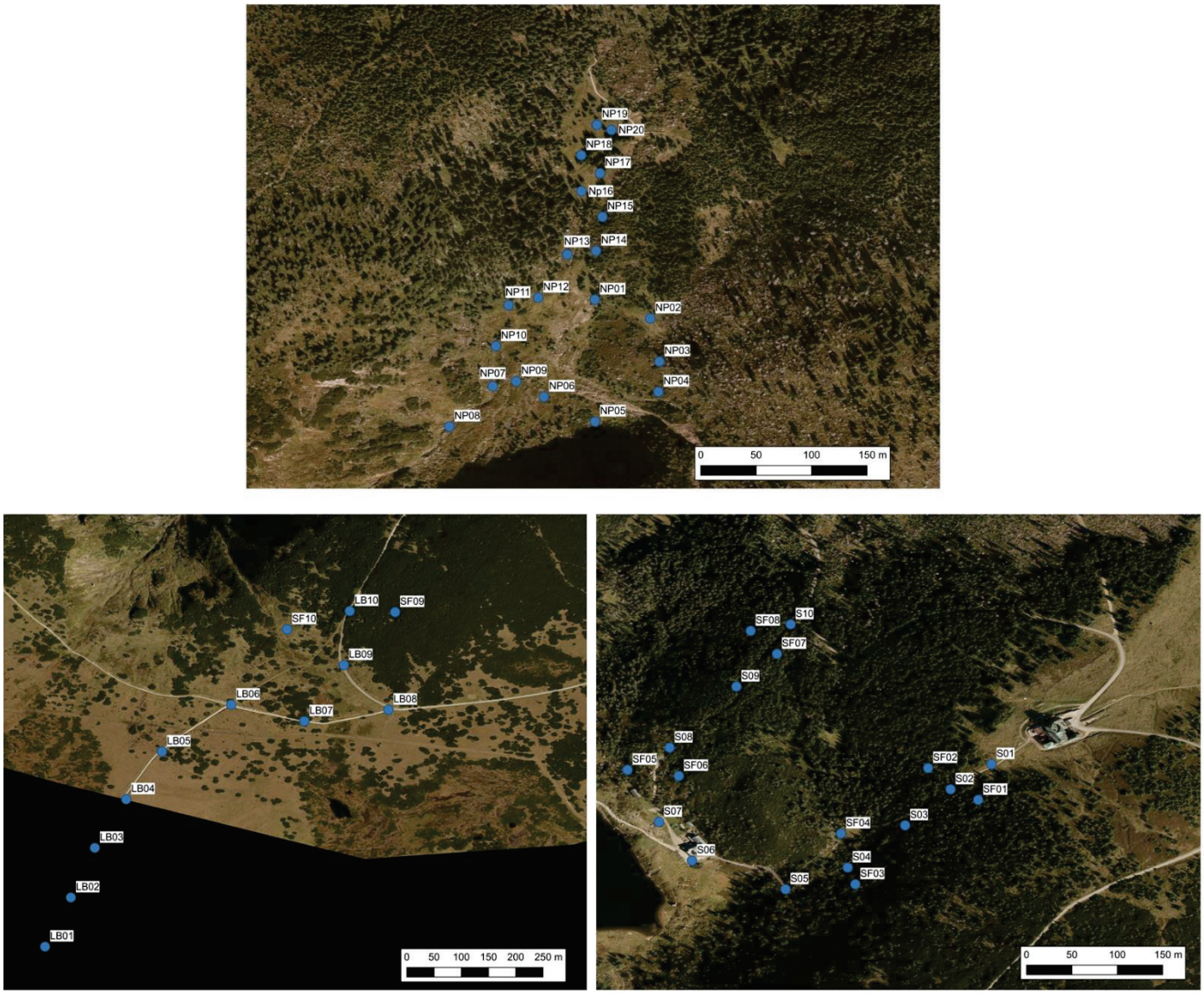


Figure 1. Map of the samples from the Czarny Kociol Jagniatkowski catchment (top) and the Kociol Malego Stawu catchment (bottom). LB1 to 3 were outside Polish territory in the Czech Republic and satellite images are not shown.

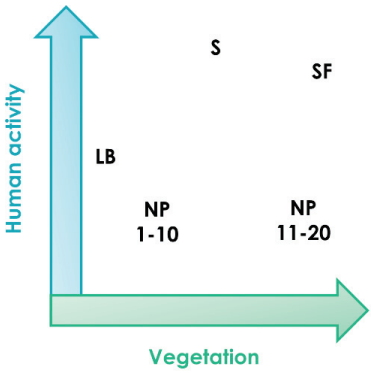


Figure 2 . Diagram illustrating the gradient of human activity and surrounding vegetation that characterized the five areas analyzed in this study.

DNA extraction, quantification and 16S rRNA gene amplification by qPCR

Microbial DNA was extracted from 0.2 µm filters using the DNeasy PowerWater Kit (QIAGEN) and eluted in 100 µl of Solution EB (QIAGEN). DNA concentrations were measured using the Qubit Fluorometer and the Qubit dsDNA HS Assay Kit (ThermoFisher). The size of the total bacterial community was estimated by quantifying the V3 region of the 16S rRNA gene by qPCR using the "universal" primers 341F (5'-CCT ACG GGA GGC AGC AG- 3') and 534R (5'-ATT ACC GCG GCT GCT GGC A-3')[13,14]. qPCR assays were carried out using the Corbett Rotor-Gene 6000 (QIAGEN) in a 20 µl reaction volume containing GoTaq qPCR Master Mix (Promega), 0.75 µM of each primer and 2 µl of DNA. Two non-template controls were also included in all the assays. Standard curves were obtained using 10-fold serial dilutions of a linearized plasmid pGEM-T Easy Vector (10^2 to 10^7 copies) containing the 16S rRNA *aeruginosa* PAO1. Cycling conditions for qPCR amplification were 95 °C for 2 minutes followed by 35 cycles of 95 °C for 15 seconds, 60 °C for 30 seconds and 72 °C for 30 seconds. Melting curves were generated after amplification by increasing the temperature from 60 °C to 95 °C. The number of copies of the 16S rRNA gene were normalized per liter of snow. ANOVA tests and pairwise t-student tests between sites were done using the ggpubr package in R.

16S rRNA gene sequencing and analysis

The V3-V4 hypervariable regions of bacterial 16S rRNA gene were amplified from 3 µl of DNA using the Platinum Taq Polymerase (Invitrogen) and forward 341F with Illumina overhang (5'-TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG CCT ACG GGN GGC WGC AG-3') and reverse 785F with Illumina overhang (5'-GTC TCG TGG GCT CGG AGA TGT GTA TAA GAG ACA GGT CTC GTG GGC TCG GAG ATG TGT ATA AGA GAC AGG ACT ACH VGG GTA TCT AAT CC-3') primers[15]. Amplification conditions were as follows: 95 °C for 3 minutes followed by 35 cycles of 95 °C for 30 seconds, 55 °C for 30 seconds and 72 °C for 30 seconds and a final extension step at 72°C for 5 minutes. DNA libraries were prepared from amplified products based on Illumina's "16S Metagenomics Library Prep Guide" (15044223 Rev. B) using the Platinum Taq DNA Polymerase (Invitrogen) and the Nextera XT Index Kit V2 (Illumina). DNA sequencing with a 15% PhiX spike-in was performed using the MiSeq System and the MiSeq Reagent Kit v2 (Illumina). Reads were trimmed to meet a quality score of Q20. Then, pair-ended reads were assembled using PANDAseq[16] at a sequence length between 410 and 500 bp and an overlap length between 20 and 100 bp, using the rdp_mle algorithm. Each of the DNA sequences was annotated to the genus level using the Ribosome Data Project (RDP) database and the RDP Bayesian classifier using an assignment confidence cut-off of 0.6[17]. Obtained sequencing depths are plotted in Figure S1 in Supplementary Material. Contaminant sequences from blanks and non-template controls were removed using the decontam package in R[18] and genera that had less than 10 sequences in the ensemble of reads were removed. Then, genus richness and diversity were calculated using the vegan package in R[19]. ANOVA tests and pairwise t-student tests between sites were used to compare genus richness and diversity using the ggpubr package in R. In addition, genus relative abundances were calculated and a NMDS analysis was performed using the vegan package in R. Finally, genus relative abundances were normalized by the number of copies of the 16S rRNA gene per liter of snow and statistical differences in genus abundance between sites were found using the DESeq2 package in R[20]. Log2FoldChange values were adjusted using the Approximate Posterior Estimation for generalized linear model or apeglm[21] and results with a log2FoldChange higher than ±2 and an adjusted p-value lower than 0.05 were plotted using the ggplot2 package in R.

Metagenomics sequencing and analysis

Metagenomics libraries were prepared from <1 ng of DNA using the Nextera XT Library Pep Kit and Indexes (Illumina), as detailed in Illumina's "Nextera XT DNA Library Prep Kit" reference guide (15031942 v03). DNA sequencing with a 1% PhiX spike-in was performed using the MiSeq System and

the MiSeq Reagent Kit v2 (Illumina). Sequencing depths are plotted on Figure S1 in Supplementary Material. Non-assembled reads were submitted to an ARG screening. Firstly, reads were trimmed using the Fastq Quality Trimmer tool of the FASTX-Toolkit. Nucleotides that did not meet a minimum quality score of Q20 were trimmed from the sequences, and sequences shorter than 100 nucleotides after trimming were removed. Then, reads from R1 and R2 were concatenated and blasted against the CARD antibiotic gene database[22] using Diamond[23]. The obtained results were filtered at a minimum identity of 60%, a minimum length of 33 amino acids and an e-value of $10e^{-5}$, and the best hit was chosen. Singletons and ARGs present in blanks were removed and ARG richness and diversity were calculated using the vegan package in R. Then, ARG copies were normalized per liter of snow. ANOVA tests and pairwise t-student tests between sites were used to compare total ARG reads per liter of snow, ARG richness and diversity using the ggpubr package inR. Then, ARGs were grouped by antibiotic class and their average abundance was compared between sites. In addition, statistical differences in ARG abundance between sites were found using the DESeq2 package in R. Log2FoldChange values were adjusted using the Approximate Posterior Estimation for generalized linear model or apeglm. Results with a log2FoldChange higher than ±2 and an adjusted p-value lower than 0.001 were plotted using the ggplot2 package in R.

In addition, metagenomic reads were filtered according to the criteria described by Minoche *et al*[24] and co-assembled using MEGAHIT[25] to generate contigs. Then, reads were mapped onto the contigs using Bowtie 2[26] to generate BAM files. Profiles were created for each individual sample and merged using the anvi'o metagenomic workflow[27]. In order to determine whether the assembled contigs contained ARGs, the merged profile was blasted against the CARD antibiotic resistance gene database using Diamond. The obtained results were filtered at a minimum identity of 60%, a minimum length of 33 amino acids and an e-value of $10e^{-5}$. The best-hit was selected. Samples were binned based on differential coverage and sequence composition. A bin was considered as a metagenome assembled genome (MAG) when it showed a completion higher than 50% and a redundancy lower than 10%.

Results

Bacterial abundance

The effects of vegetation and anthropogenic activity on bacterial abundance in snow were estimated by comparing the 16S rRNA gene copy number measured by qPCR across sampling site. Samples that had less surrounding vegetation showed a lower microbial biomass than those surrounded by vegetation (Figure 3). Samples from the LB site had significantly lower copies of the 16S rRNA gene per L of snow than samples from any other site. Samples from NP 1-10 had significantly lower copies of the 16S rRNA gene per L of snow than samples from NP 11-20. Although surrounding vegetation correlated to the abundance of bacterial communities in snow, the increase in bacterial biomass in SF compared to S was not significant and NP 1-10 had significantly less copies of the 16S rRNA gene than S. Thus, in order to evaluate the impact of anthropogenic activity on bacterial biomass regardless of vegetation, NP 1-20 were compared to S and SF (Figure 3). Bacterial abundance measured from the sites affected by anthropogenic activity in the Kociol Malego Stawu catchment (S-SF) was significantly higher than that detected in the Czarny Kociol Jagniatkowski catchment (NP 1-20).

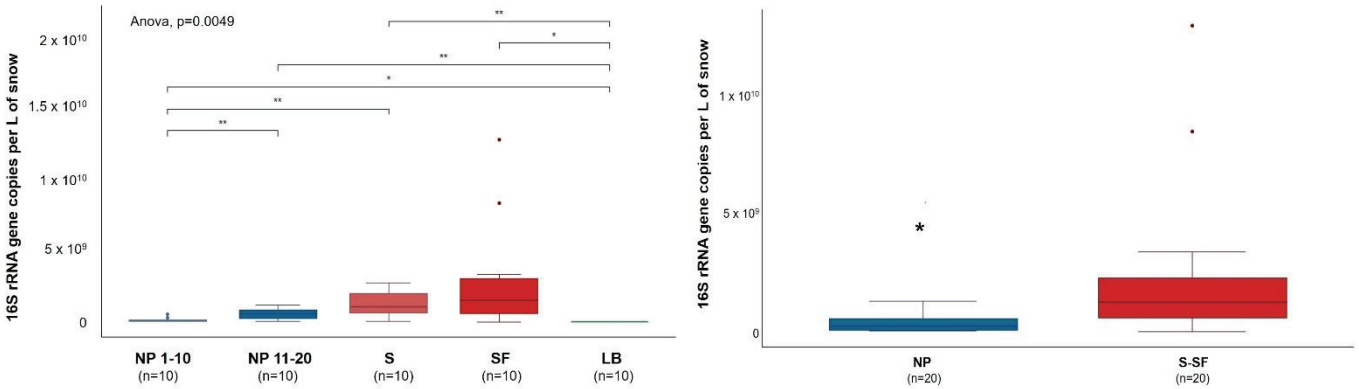


Figure 3. Estimates of microbial biomass based on qPCR of the 16S rRNA gene from all sites (left) and from NP and S-SF (right). Copies were normalized per L of water. qPCR efficiency=1.04. R^2 linearity coefficient=0.986. Significant differences between sites were determined by pairwise t-tests. *p-value ≤0.05. **p-value≤0.01.

Genus richness and diversity

Although no significant differences were found between sites in terms of bacterial richness, the diversity in samples from LB was significantly higher than that detected from any other site (Figure 4). Anthropogenic activity did not have any significant impact on genus richness and diversity (Figure 4).

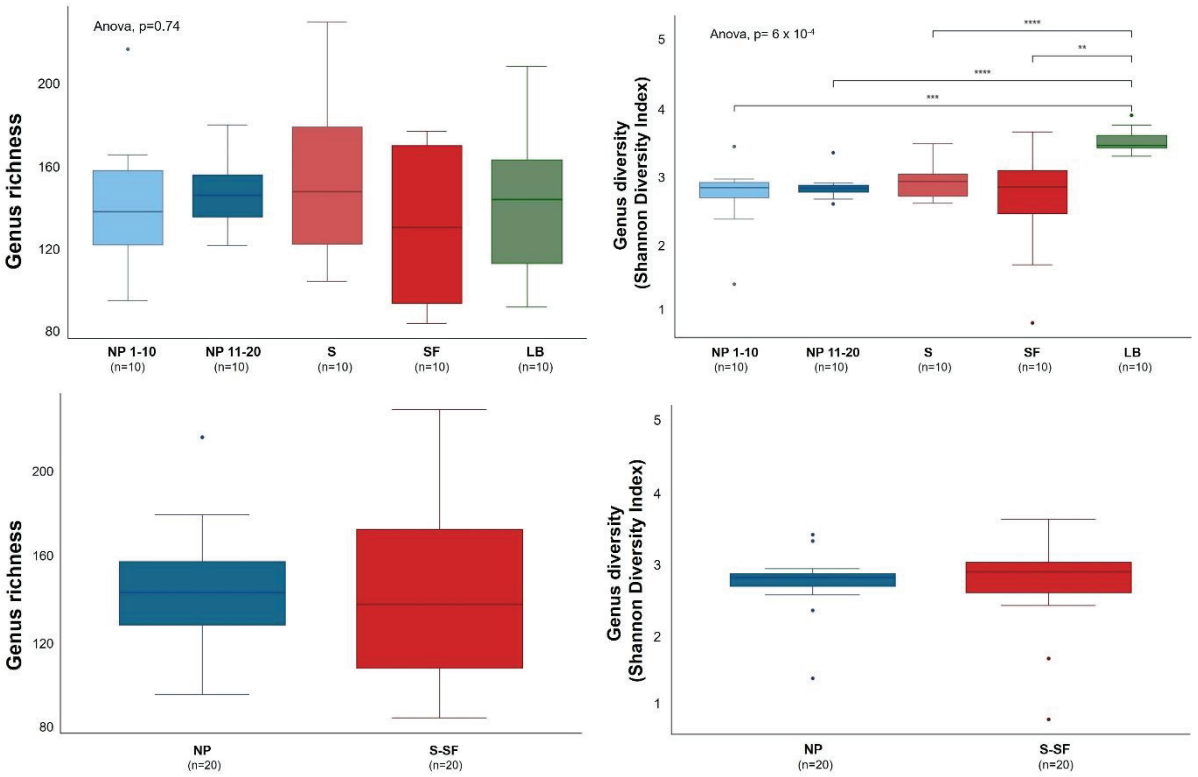


Figure 4. Genus richness (left) and diversity estimated using the Shannon Diversity Index (right) obtained from all sites (top) and from NP versus S-SF (bottom). Significant differences between sites were determined by pairwise t-tests. *p-value ≤0.05. **p-value≤0.01. ***p-value ≤0.001. ****p-value≤0.0001.

Community composition

A non-metric multidimensional scaling (NMDS) analysis was applied to all of the samples to determine whether vegetation and/or anthropogenic activity correlated with bacterial community composition (Figure 5). Most of the samples from NP 1-10, NP 11-20, S and SF formed a cluster, whereas all LB samples clustered separately from the rest. The only sample from SF that was included in the LB cluster was SF9, which is geographically closer to LB than to the rest of samples in SF (Figure 1).

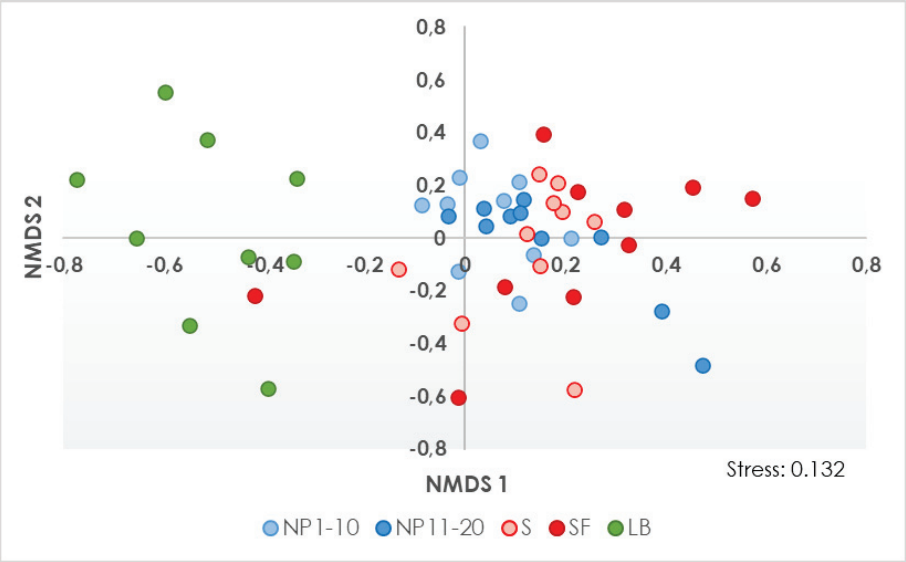


Figure 5. Bacterial community similarity (NMDS analysis) of snow from all sites.

Pairwise comparisons of genus abundance were performed between each site combination (Figure 6). No significantly more abundant genera were found in LB as compared to other sites. On the contrary, when compared to LB, all sites had significantly more abundant genera. Both *Acidipila* and *Acidicapsa,* which belong to the family *Acidobacteriaceae,* were found in a higher abundance at all sites than in LB (Figure 6A-D). Several members of the heterotrophic *Oxalobacteraceae* family (*Telluria, Massilia, Janthinobacterium, Paraherbaspirillum* and *Glaciimonas*) had higher abundances at NP 11-20 than at LB (Figure 6B) and all but *Paraherbaspirillum* and *Glaciimonas* were also at higher abundances at SF than at LB (Figure 6D). Some genera belonging to the *Sphingobacteriaceae* family were also found to be more abundant at other sites than at LB. *Nubsella* was more abundant at NP 12-20, S and SF (Figure 6B-D), whereas *Mucilaginibacter* was more abundant at S than at either LB or NP 1-10 (Figure 6C, F) and *Pedobacter* was more abundant at SF than at either LB or NP 1-10 (Figure 6D, G). This difference in genus abundance is not sufficient to change the overall similarity of S, SF, NP1-10, and NP11-20. No significant differences were found between NP 1-10 and NP 11-20 nor between S and NP 11-20. Insignificant differences were found between sites subjected to different levels of anthropogenic activity with only two genera showing significant shifts between sites (Figure 6E-I). This is consistent with the NMDS analysis.

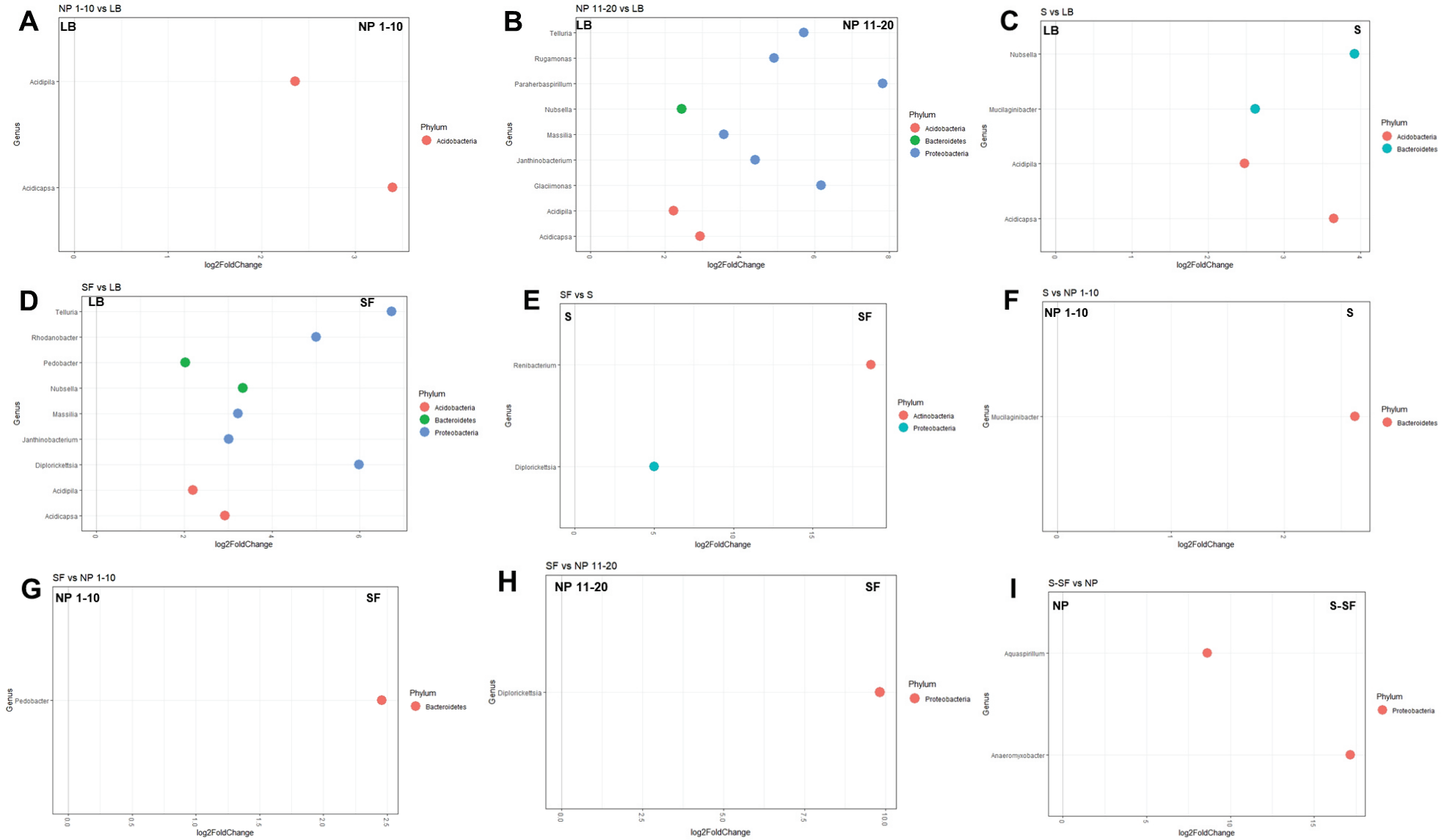


Figure 6. Genus abundance pairwise comparisons between all sites (A-H, n=10) and from NP 1-20 versus S-SF (I, n=20). Genus relative abundance was multiplied by the 16S rRNA gene copies per L of snow. Only results with a log2FoldChange higher than ±2 and an adjusted p-value lower than 0.05 are shown.

Antibiotic resistome size, richness and diversity

The number of ARG reads obtained from the metagenomic sequencing of snow DNA was normalized per liter of snow (Figure 7). Significantly less ARG reads were obtained from the sequencing of LB samples than from NP 11-20, S and SF. NP 1-10 had fewer ARG sequences than S. In addition, the size of the antibiotic resistome was significantly lower at NP 1-20 than at S and SF (Figure 7).

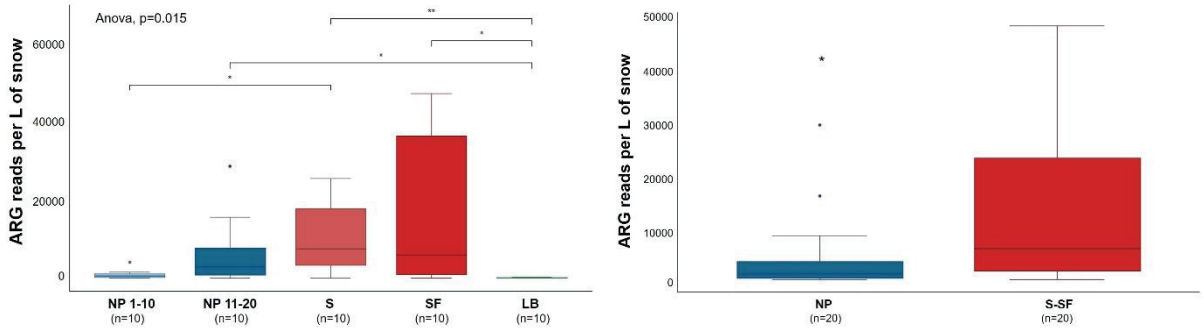


Figure 7. Size of the snow antibiotic resistome. Total ARG copies per L of snow from all sites (left) and from NP and S-SF (right). Significant differences between sites were determined by pairwise t-tests. *p-value ≤0.05. **p-value≤0.01.

In addition, the impact of vegetation and anthropogenic activity on ARG richness and diversity was evaluated. Both ARG richness and diversity were significantly lower at LB than at S and SF (Figure 8). No significant differences were found between NP 1-20 and S-SF.

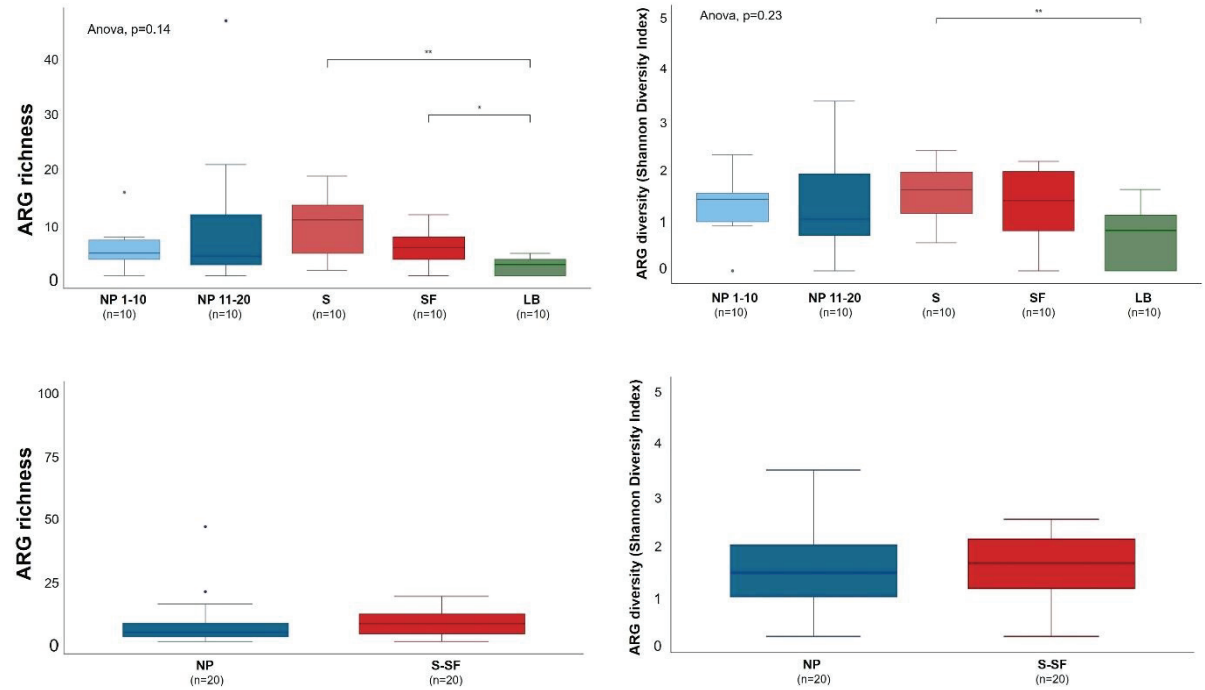


Figure 8. ARG richness (left) and diversity estimated using the Shannon Diversity Index (right) obtained from all sites (top) and from NP versus S-SF (bottom) Significant differences between sites were determined by pairwise t-tests. *p-value ≤0.05. **p-value≤0.01.

Antibiotic resistome composition

The composition of the antibiotic resistome at different sites was compared. Sites with less surrounding vegetation (*i.e.* NP 1-10 and LB) had a lower amount of ARGs from all antibiotic classes

than at any other site (Figure 9). In addition, S and SF had higher numbers of multidrug efflux genes than NP 11-20 and an increase in ARGs conferring resistance to antimicrobial peptides and triclosan was also observed at SF. When NP 1-20 were compared to S-SF, ARGs from virtually every antibiotic class were more abundant in anthropogenically-impacted sites (S-SF, Figure 9). This increase in abundance was particularly high in genes related to multidrug resistance.

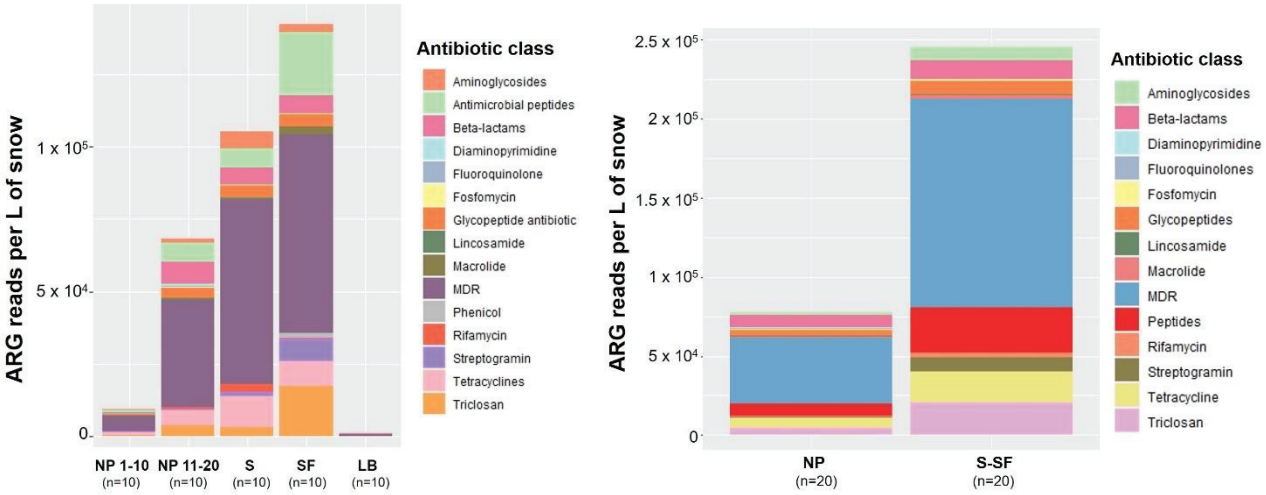


Figure 9. ARG copies per L of snow from all sites (left) and from NP vs S-SF (right) grouped by antibiotic class.

Gene abundance differences between sites were evaluated by pairwise comparisons using DESeq2 (Figure 10). All sites had several genes that were significantly more abundant than in LB, whereas no ARGs were significantly more abundant at LB (Figure 10 A-D). S showed the highest number of ARGs (14) with a significant increase in abundance as compared to LB, followed by SF (8), NP 11-20 (6) and NP 1-10 (5). In addition, the same number of ARG were significantly more abundant at NP 11-20 than at NP 1-10 and vice versa: *vanO, tet(43)* and *smeF* were more abundant at NP 11-20, whereas *tetA(46), LpeB* and *bcrA* were more abundant at NP 11-10 (Figure 10E). On the other hand, S showed a higher abundance of nine ARGs when compared to SF, which only had three genes that were more abundant than at S (Figure 10F). Furthermore, when comparing S and SF to NP 1-10 and NP 11-20 (Figure 10 G-J), more genes were found in a higher abundance at S or SF than at NP 1-10 or NP 11-20. These differences were higher between S and NP 1-10 or NP 11-20 than between SF and NP 1-10 or NP 11-20. Finally, the resistome of NP 1-20 was compared to that of S-SF (Figure 10K). *tetA(60), tetA(46)* and *LpeB* were more abundant at NP 1-20, whereas *tetO, tetB(58), rphB, Rm3, otrC, mexJ, FosA5, aadA25* and *aadA17* were found in a higher abundance at S and SF. These genes were involved in resistance to several antibiotic classes (tetracyclines, rifamycin, beta-lactams, fosfomycin, aminoglycosides and multidrug resistance).

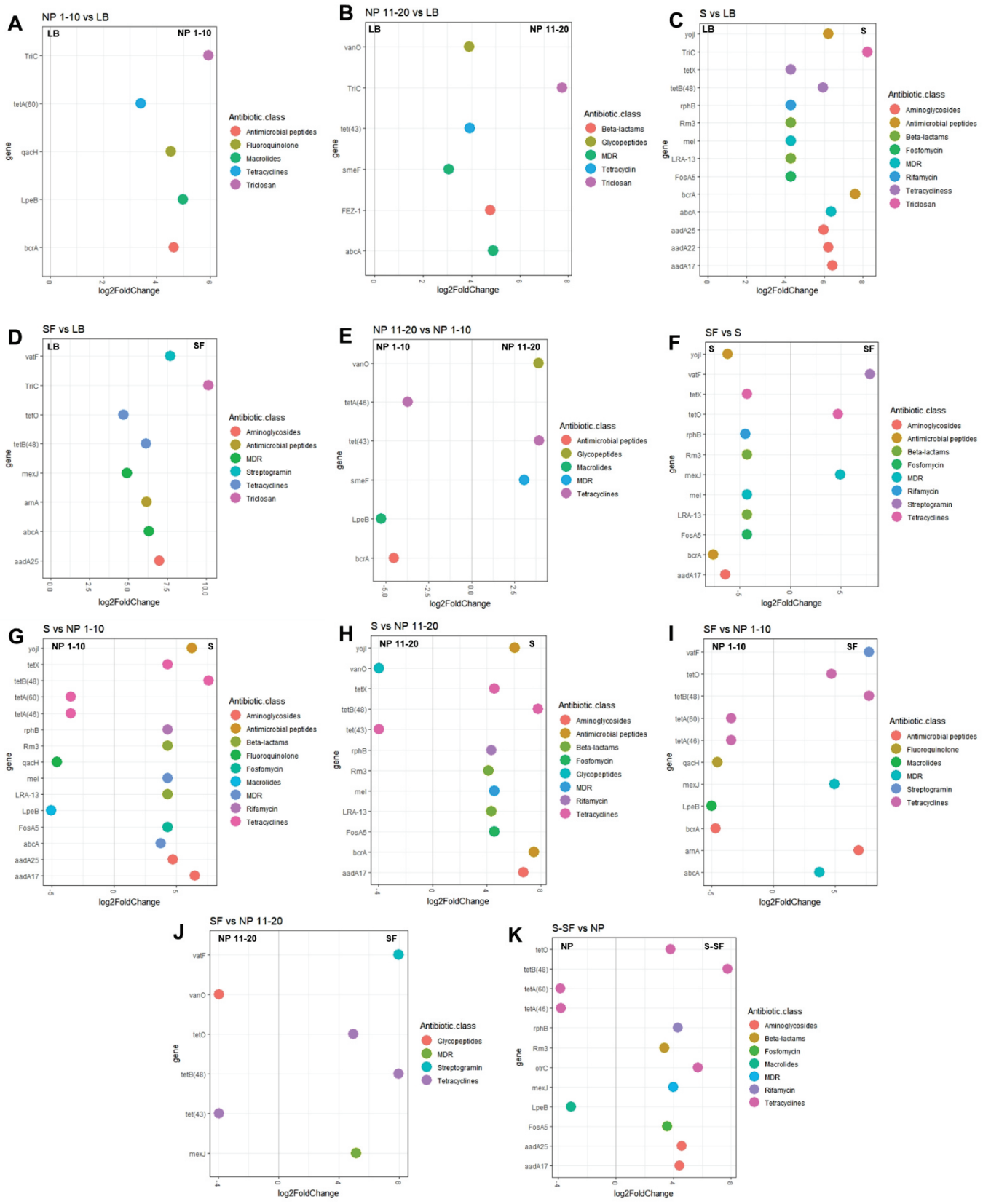


Figure 10. ARG abundance pairwise comparisons between all sites (A-J, n=10) and from NP 1-20 versus S-SF (K, n=20). Only results with a log2FoldChange higher than ±2 and an adjusted p-value lower than 0.001 are shown.

The co-assembly of metagenomics sequences produced a bin with a completion of 59.2% and a redundancy of 2.8% (Figure 11). This bin was putatively identified as *Rickettsiella isopodorum.* The bin showed a higher coverage in samples from SF and had one associated ARG, *vatF*. This streptogramin resistance gene was found more abundant at SF than anywhere else when pairwise comparisons were performed (Figure 6).

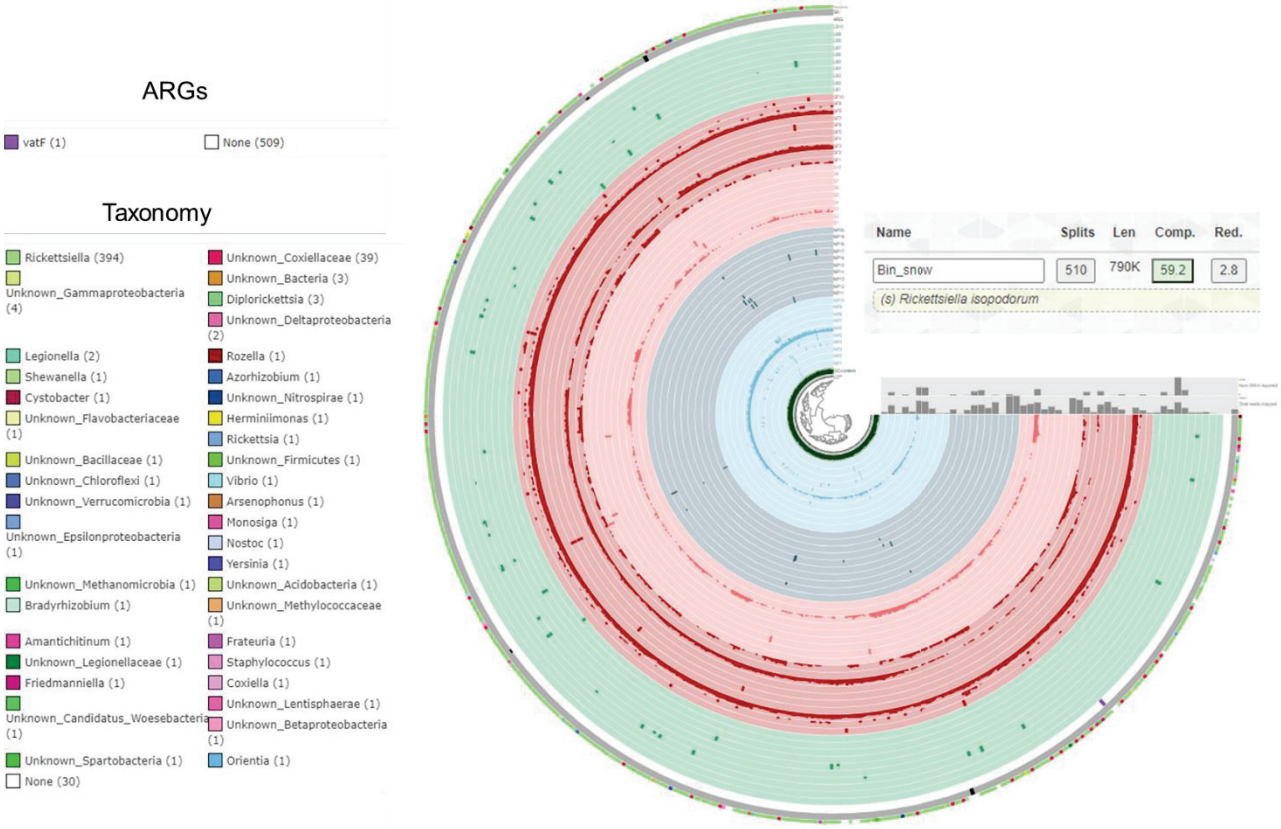


Figure 11. Bin obtained from the co-assembly of metagenomic sequences from all sites. Light blue: NP 1-10, dark blue: NP 11-20, light red: S. dark red: SF. Green: LB. n=10.

Discussion

The goal of this study was to determine whether environmental factors such as vegetation and anthropogenic activity influenced the size and composition of the snow bacterial communities and antibiotic resistome. Our results support the hypothesis that growth is limited at sites with lower levels of organic carbon and that sites with more surrounding vegetation seem to able to support a higher biomass and less diverse community. While no significant overall community differences were found between sites (other than the extreme LB site), several genera that use carbon as a nutrient source were more abundant in sites with higher levels of vegetation. Both *Acidipila* and *Acidicapsa,* which belong to the family *Acidobacteriaceae* and are able to degrade organic carbon and plant polysaccharides including cellulose, and some genera belonging to the *Sphingobacteriaceae* family, which is composed of chemo-organotrophic organisms, were found in a higher abundance at all sites other than in LB (Figure 6). In addition, the differences in genus abundance of genera that use organic carbon as a nutrient source increased along with differences in surrounding vegetation. We hypothesized that the higher abundance of heterotroph and chemo-organotroph organisms was related to higher organic carbon and other nutrient availability. The low microbial biomass site (LB) is

the site subjected to the most extreme environmental conditions (wind, cold) and lowest amount of surrounding vegetation. Therefore, lower amounts of organic carbon would reduce the growth of potentially copiotrophic organisms and produce a more even microbial community where no specific taxon can outgrow others, which is supported by the high diversity levels at LB (Figure 4). Thus, only the extreme conditions and the virtually complete absence of vegetation associated with LB samples seem to have an effect on bacterial diversity. The presence of vegetation could be inversely related to a higher diversity of bacteria. Snowpack microbial communities are seeded from the atmosphere through wet and dry deposition processes[8,28]. Since the snow is coupled to the atmosphere, factors that affect airborne microbial communities, such as surrounding ecosystems and local meteorological conditions, should change the snow microbial community structure[29,30]. At sites where the vegetation is scarce, microbial life would be expected to colonize snow mainly from the atmosphere and dust[31] and at sites with surrounding vegetation, the plants would contribute to the bacterial communities present in snow, providing a less even and less diverse community.

In addition, the more favorable conditions for bacterial growth observed at sites with higher levels of surrounding vegetation seem to provoke higher levels of antibiotic resistance at the genetic level. All sites with more surrounding vegetation had a more abundant antibiotic resistome, which was likely a consequence of the higher biomass. Sites surrounded by vegetation contained several ARGs that were more abundant than at the low vegetation extreme site (LB), whereas LB did not have any ARGs in a higher abundance than anywhere else (Figure 10). Thus, vegetation appeared to support the growth of a more abundant bacterial community, which in turn had a higher abundance of antibiotic resistance genes. Previous research showed that increased levels of organic acids induced competition and increased the abundance of ARGs in snow[32]. They also support the establishment of a more metabolically active community[33]. Therefore, surrounding vegetation may provide additional sources of organic acids and stimulate competition and ARG proliferation in snow.

Although surrounding vegetation correlated to the abundance of bacterial communities in snow, a significantly higher bacterial abundance was measured at sites associated with anthropogenic activity than at less frequented sites with similar levels of surrounding vegetation (Figure 3). Thus, both vegetation and anthropogenic activity influenced the abundance of the bacterial community in snow. Anthropogenic activity was related to a higher biomass community and induced changes in the snow antibiotic resistome while having a limited impact on the composition of the bacterial communities. Both sites subjected to anthropogenic activity (S and SF) were surrounded by vegetation, which translated to favorable conditions for bacterial growth. Therefore, the potential added effects of anthropogenic waste would not necessarily impact the core bacterial community. This is consistent with the lack of differences in genus richness and diversity and the overall composition of the bacterial communities despite the higher biomass. On the other hand, anthropogenic activities seemed to influence the size and composition of the antibiotic resistome without having a significant effect on ARG richness and diversity. A partial metagenome assembled genome (MAG) predominantly with sequences from an anthropogenically-impacted site (SF) was putatively identified as *Rickettsiella isopodorum,* a Gammaproteobacteria that is an intracellular pathogen of terrestrial isopod crustaceans[34] (Figure 11). This MAG had one associated ARG, *vatF*, that was first isolated from *Yersinia enterocolitica*[35]*,* a bacterium associated to the pet microbiome that can cause infections in humans. Since human microbiome bacteria are unlikely to survive in extreme environments, direct biological contamination of anthropogenically-impacted snow should not have a strong effect on the snow microbiome. Nevertheless, human activities could provide the snow microbiome with additional levels of organic carbon through increased waste dispersion. This in turn would stimulate bacterial growth, competition and ARG production in a similar way as the presence of vegetation. A previous study has

shown that anthropogenic waste can induce ARG spread without influencing community composition, and linked these changes to a higher macronutrient load in sewage[36].

In conclusion, both vegetation and anthropogenic activities induced changes in the snow microbiome and its associated resistome and showed that some micro-organisms are responding to environmental changes in alpine snow. Although several studies have illustrated the impact of anthropogenic activity on the environmental resistome[37–39], this is the first study to demonstrate that winter tourism induced changes in the snow resistome without affecting the core bacterial community structure. Thus, anthropogenic activity could be an indirect source of environmental pollution and stimulate the development of antibiotic resistance in the snow microbiome that might be subsequently disseminated through the atmosphere and snow melting. The genes that showed an increased abundance at sites contaminated by human activity could be considered as biomarkers of anthropogenic impact on alpine snow. However, this increase in abundance needs to be confirmed by qPCR analyses. This study provides insights into the potential impact of environmental and anthropogenic factors on snow microbial ecology and highlights the need for survey of antibiotic resistance development in sites affected by anthropogenic activities and the consequences that environmental pollution may have on antibiotic resistance dispersion.

References

1. Körner, C. & Paulsen, J. A world-wide study of high altitude treeline temperatures. *J. Biogeogr.* **31,** 713–732 (2004).

2. Mountain Research Initiative EDW Working Group. Elevation-dependent warming in mountain regions of the world. *Nat. Clim. Chang.* **5,** 424–430 (2015).

3. Mandaric, L. *et al.* Contamination sources and distribution patterns of pharmaceuticals and personal care products in Alpine rivers strongly affected by tourism. *Sci. Total Environ.* **590–591,** 484–494 (2017).

4. Barros, A., Aschero, V., Mazzolari, A., Cavieres, L. A. & Pickering, C. M. Going off trails: How dispersed visitor use affects alpine vegetation. *J. Environ. Manage.* **267,** 110546 (2020).

5. Ballantyne, M. & Pickering, C. M. Recreational trails as a source of negative impacts on the persistence of keystone species and facilitation. *J. Environ. Manage.* **159,** 48–57 (2015).

6. UNEP 2007 Annual Report. (2007).

7. Cowan, D. A. *et al.* Non-indigenous microorganisms in the Antarctic: Assessing the risks. *Trends Microbiol.* **19,** 540–548 (2011).

8. Maccario, L., Carpenter, S. D., Deming, J. W., Vogel, T. M. & Larose, C. Sources and selection of snow-specific microbial communities in a Greenlandic sea ice snow cover. *Sci. Rep.* **9,** 1–14 (2019).

9. Zhu, G. *et al.* Air pollution could drive global dissemination of antibiotic resistance genes. *ISME J.* (2020). doi:10.1038/s41396-020-00780-2

10. Shen, F. & Yao, M. Are We Biologically Safe with Snow Precipitation? A Case Study in Beijing. *PLoS One* **8,** (2013).

11. Harding, T., Jungblut, A. D., Lovejoy, C. & Vincent, W. F. Microbes in high arctic snow and implications for the cold biosphere. *Appl. Environ. Microbiol.* **77,** 3234–3243 (2011).

12. Rogers, S. O., Starmer, W. T. & Castello, J. D. Recycling of pathogenic microbes through survival in ice. *Med. Hypotheses* **63,** 773–777 (2004).

13. Muyzer, G., Hottentrager, S., Teske, A. & Wawer, C. Denaturing gradient gel electrophoresis of PCR-amplified 16S rDNA—a new molecular approach to analyse the genetic diversity of mixed microbial communities. In *Molecular microbial ecology manual;* (eds. Akkermans, A., van Elsas, J. & de Bruijn, F.) 1–23 (Kluwer Academic Publishers, 1995).

14. Watanabe, K., Kodama, Y. & Harayama, S. Design and evaluation of PCR primers to amplify bacterial 16S ribosomal DNA fragments used for community fingerprinting. *J. Microbiol. Methods* **44,** 253–262 (2001).

15. Klindworth, A. *et al.* Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res.* **41,** 1–11 (2013).

16. Masella, A. P., Bartram, A. K., Truszkowski, J. M., Brown, D. G. & Neufeld, J. D. PANDAseq: Paired-end assembler for illumina sequences. *BMC Bioinformatics* **13,** 31 (2012).

17. Wang, Q., Garrity, G. M., Tiedje, J. M. & Cole, J. R. Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* **73,** 5261–5267 (2007).

18. Davis, N. M., Proctor, D. M., Holmes, S. P., Relman, D. A. & Callahan, B. J. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *bioRxiv* **221499,** (2017).

19. Oksanen, J. *et al.* vegan: Community Ecology Package. R package version 2.4-2. *Community Ecol. Packag.* **2.5-6,** 1–296 (2019).

20. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15,** (2014).

21. Zhu, A., Ibrahim, J. G. & Love, M. I. Heavy-Tailed prior distributions for sequence count data: Removing the noise and preserving large differences. *Bioinformatics* **35,** 2084–2092 (2019).

22. Alcock, B. P. *et al.* CARD 2020: Antibiotic resistome surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids Res.* **48,** D517–D525 (2020).

23. Benjamin Buchfink, Xie, C. & Huson, D. H. Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* **12,** 59–60 (2015).

24. Minoche, A. E., Dohm, J. C. & Himmelbauer, H. Evaluation of genomic high-throughput sequencing data generated on Illumina HiSeq and Genome Analyzer systems. *Genome Biol.* **12,** (2011).

25. Li, D., Liu, C. M., Luo, R., Sadakane, K. & Lam, T. W. MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* **31,** 1674–1676 (2015).

26. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9,** 357–359 (2012).

27. Eren, A. M. *et al.* Anvi'o: An advanced analysis and visualization platformfor 'omics data. *PeerJ* **2015,** 1–29 (2015).

28. Maccario, L., Vogel, T. M. & Larose, C. Potential drivers of microbial community structure and function in Arctic spring snow. *Front. Microbiol.* **5,** (2014).

29. Els, N. *et al.* Comparison of Bacterial and Fungal Composition and Their Chemical Interaction in Free Tropospheric Air and Snow Over an Entire Winter Season at Mount Sonnblick, Austria. *Front. Microbiol.* **11,** (2020).

30. Tignat-Perrier, R. *et al.* Global airborne microbial communities controlled by surrounding landscapes and wind conditions. *Sci. Rep.* **9,** (2019).

31. Chuvochina, M. S., Alekhina, I. A., Normand, P., Petit, J. R. & Bulat, S. A. Three events of Saharan dust deposition on the Mont Blanc glacier associated with different snow-colonizing bacterial phylotypes. *Microbiology* **80,** 125–131 (2011).

32. Bergk Pinto, B., Maccario, L., Dommergue, A., Vogel, T. M. & Larose, C. Do Organic Substrates Drive Microbial Community Interactions in Arctic Snow? *Front. Microbiol.* **10,** 1–13 (2019).

33. Zhu, C. *et al.* Snow microbiome functional analyses reveal novel aspects of microbial metabolism of complex organic compounds. *Microbiologyopen* **9,** 1–12 (2020).

34. Wang, Y. D. & Chandler, C. Candidate pathogenicity islands in the genome of 'Candidatus Rickettsiella isopodorum', an intracellular bacterium infecting terrestrial isopod crustaceans. *PeerJ* **2016,** (2016).

35. Seoane, A. & García Lobo, J. M. Identification of a streptogramin a acetyltransferase gene in the chromosome of Yersinia enterocolitica. *Antimicrob. Agents Chemother.* **44,** 905–909 (2000).

36. Lehmann, K. *et al.* Trace levels of sewage effluent are sufficient to increase class 1 integron prevalence in freshwater biofilms without changing the core community. *Water Res.* **106,** 163–170 (2016).

37. Chen, B., Yang, Y., Liang, X., Zhang, T. & Li, X. Metagenomic Profiles of Antibiotic Human Impacted Estuary and Deep Ocean Sediments. *Environ. Sci. Technol.* **47,** 12753–12760 (2013).

38. Suzuki, S., Ogo, M., Koike, T., Takada, H. & Newman, B. Sulfonamide and tetracycline resistance genes in total- and culturable-bacterial assemblages in south african aquatic environments. *Front. Microbiol.* **6,** 1–8 (2015).

39. Czekalski, N., Sigdel, R., Birtel, J., Matthews, B. & Bürgmann, H. Does human activity impact the natural antibiotic resistance background? Abundance of antibiotic resistance genes in 21 Swiss lakes. *Environ. Int.* **81,** 45–55 (2015).

Chapter VI. Conclusions and perspectives

Anthropogenic environmental pollution with residual concentrations of antibiotics and biological pollutants (ARB and ARGs) is a major source of concern **(Chapter I)**. Antibiotics may impose a selective pressure in the environmental microbiome stimulating the development and dissemination of antibiotic resistance and potentially even at sub-inhibitory concentrations (discussion in **Chapter I**). In addition, biological contamination may contribute to the environmental resistome reservoir and organic nutrients present in anthropogenic waste may affect environmental microbial ecology and promote changes in the environmental resistome. In both cases, environmental antibiotic resistance is susceptible to transfer to the human and animal microbiome and further aggravate the antibiotic crisis that we are currently facing. On the other hand, interactions between both chemical and biological pollutants and the receiving environmental microbiome are influenced by environmental factors and the physicochemical properties of the environmental matrix. Research is needed in order to determine i) the scope of the response of the environmental microbiome and resistome to both chemical and biological pollutants and the relationship between this response and environmental factors, ii) the mobilization and dissemination potential of antibiotic resistance in the environment, and iii) the risk of transfer of environmental ARB and ARGs to the human and animal microbiome.

The main objective of this PhD was to provide elements to help respond to major questions regarding antibiotic resistance development in the environment subjected to anthropogenic pollution. This research was divided into three main axes. First, a methodological study aimed to answer the question *"is taxonomic and functional/ARG richness discovery in soil dependent on nucleic extraction methods or sequencing depth?"* Second, the response to gentamicin pollution at both inhibitory and sub-inhibitory concentrations was evaluated in two environments, soil and freshwater, in order to answer the questions *"can sub-inhibitory concentrations of gentamicin induce a response in the environmental microbiome and resistome?", "how is the link between gentamicin dose and the magnitude of this response?"* and *"is this response consistent between environments?"* Finally, a field-study aimed to provide insights into the question *"does anthropogenic activity impact microbiome and resistome composition in cold environments?"* This work is summarized and discussed here.

Effect of DNA/RNA extraction and sequencing depth on richness discovery

A methodological study was considered to be necessary in order to assess the biases related to soil DNA and RNA extraction and sequencing prior to the evaluation of the effects of antibiotic pollution **(Chapter II).** Since soil is a particularly challenging matrix due to its high microbial diversity and its heterogeneous physicochemical properties that induce biases in DNA and RNA extraction. The effect of the choice of an extraction method on the outcoming results is not new. Yet, this study addresses a critical aspect of the growing use of high throughput sequencing to characterize different environments. Taxonomic, ARG and functional richness measured by several DNA extraction methods were compared using an Illumina MiSeq sequencing approach. In addition, the taxonomic richness revealed by different RNA extraction methods was evaluated. The main outcome of this study was that sequencing depth has a stronger effect on taxonomic and functional richness discovery than the choice of an extraction method and that sequence rarefaction at an equal depth lead to the assignation of unique genus and functional classes that were detected by more than one method before rarefaction. Therefore, the choice of a DNA or RNA extraction method seemed less relevant than the need for optimization of high throughput sequencing techniques to study soil microbial communities and their associated functions. On the other hand, ARG richness discovery was strongly influenced by sample variation in ARG content. Only one method (Maxwell 1) showed a higher and more consistent richness

discovery than the others in La Côte de Saint André soil, (the soil selected for antibiotic pollution experiments). Therefore, although this increased richness is arguably exclusively due to DNA extraction and caution should be taken before recommending this method for resistome discovery analyses in soil, this method was chosen for DNA extraction in the evaluation of antibiotic pollution in soil.

The limitations linked to Illumina MiSeq sequencing depth have been considered all through this thesis. Higher bacterial diversity observed in soil microcosms than in soil bacterial enrichments was probably due to the lower quality of the assembly obtained from soil microcosm metagenomic sequences, underlining the need of deeper sequencing technologies to cover soil bacterial diversity and accurately analyze the soil resistome using metagenomic approaches **(Chapter III).** In addition, long read sequencing using the Oxford Nanopore technology and the hybrid assembly of short and long reads showed an improvement of the completion and definition of MAGs from freshwater and demonstrated Nanopore importance in the metagenomic analysis of the resistome of complex communities **(Chapter IV).** Furthermore, qPCR analyses were performed to complement Illumina MiSeq sequencing when assessing ARG abundance **(Chapters III and IV)** and the need of further qPCR validation of differences in ARG abundance observed through metagenomic sequences was unfortunately limited when these analyses could not been carried out due to time limitations **(Chapter V).**

Impact of antibiotic anthropogenic pollution on resistance development in the environment

Microcosm studies were performed to evaluate the effects of antibiotic pollution on the environmental microbiome and resistome **(Chapters III and IV)**. Gentamicin was chosen as a model antibiotic, since it is used in both human therapy and food production. Gentamicin resistance genes are abundant and diverse in several environmental settings and sub-inhibitory concentrations of gentamicin have been shown to stimulate resistance selection and biofilm formation *in vitro*[1] as well as resistance dissemination[2]. In addition, this aminoglycoside could be analyzed by HPLC-MS/MS, and help establish links between gentamicin dose and the magnitude of the response observed in the environmental microbiome and resistome.

Gentamicin sub-inhibitory and inhibitory concentrations in soil and water bacteria enriched in culture media were determined by evaluating bacterial growth. Since the response to gentamicin pollution was expected to have a much lower magnitude in soil than in water due to gentamicin adsorption onto soil particles, different approaches were used depending on the receiving environment. On the one hand, soil microcosms were polluted with inhibitory concentrations of gentamicin. The difference in the response between the communities in enriched media and environmental settings at several exposure times was evaluated using a combination of culture-dependent and metagenomic approaches. In addition, the bioavailable fraction of gentamicin in soil was measured by HPLC-MS/MS **(Chapter III)**. On the other hand, river water microcosms were polluted with two sub-inhibitory concentrations and an inhibitory concentration to evaluate the magnitude of the response of the environmental microbiome and resistome. The effects of gentamicin at both inhibitory and sub-inhibitory concentrations at different exposure times were evaluated by assessing its effects on overall growth, transcription levels, the composition of total and active bacterial communities, and the abundance and transcription of ARGs in water microcosms. Class 1 integrons were sequenced and screened for aminoglycoside resistance genes and dose-response relationships were established

between gentamicin concentrations measured in water by HPLC-MS/MS and the abundance and expression of aminoglycoside resistance genes measured by qPCR/RT-qPCR **(Chapter IV)**.

These two studies illustrate how the effects of the same antibiotic on different environments are strongly dependent of environmental factors and physicochemical properties. Whereas in soil bacterial enrichments, an inhibitory concentration of 12 µg of gentamicin per ml of medium significantly reduced the size of the bacterial communities and altered their composition, in soil microcosms, where gentamicin is highly adsorbed onto soil particles and bacterial diversity is higher, inhibitory concentrations up to 1 mg of gentamicin per gram of soil did not have significant effects on bacterial growth nor on the composition of the soil microbiome and resistome after 8-days exposure **(Chapter III)**. With the river water microcosms, a sub-inhibitory concentration of 50 ng of gentamicin per ml of water, which did not affect the size of the bacterial community, caused a significant increase of overall transcription levels after 2-days exposure and induced shifts in the composition of both total and active communities and increased the abundance of *Limnohabitans* after only 1-day exposure **(Chapter IV)**. In addition, the hybrid assembly of short and long metagenomic reads sequenced using the Illumina MiSeq and the Oxford Nanopore technologies, respectively, were used to assemble a *Limnohabitans* MAG that contained a partial sequence of *aac(6')-Ib8,* a gene encoding for an aminoglycoside acetyltransferase involved in gentamicin resistance. Finally, gentamicin pollution at 50 ng/ml induced an increase in the abundance and transcription of *aac(6')* genes after 2-days exposure, although the magnitude of this response was lower than at 800 ng/ml, an inhibitory gentamicin concentration that causes the inhibition of most of the members of the community and the selection of gentamicin resistant clones. These genes are present in class 1 integrons and might, therefore, be mobilized and disseminated across the environmental microbiome and to the human and animal microbiome. Whereas gentamicin concentrations at the ng/ml scale induced changes in the freshwater microbiome and resistome after 2-days exposure without inhibiting overall growth, the same antibiotic at concentrations up to 1 mg/g did not affect soil bacterial communities nor their resistome after 8-days exposure. This demonstrates the need for antibiotic pollution surveillance in a wide range of environmental settings and the systematic measurement of antibiotic concentrations and characterization of environmental physicochemical properties as well as temporal series to account for seasonal changes in the environmental setting and its associated microbiome.

These studies also demonstrate the limitations of the terms "sub-inhibitory" and "inhibitory" in complex environments. Since these terms were defined in single cultures, they may not be accurate to illustrate what takes place in complex environmental communities. Regarding soil, these results illustrate the differences between the response to gentamicin pollution between bacterial enrichments and soil microcosms, where gentamicin adsorption, a higher bacterial diversity and selective forces, such as nutrient availability and predation are likely to take place. In addition, sub-inhibitory concentrations that do not cause an overall growth inhibition in freshwater may inhibit some members of the community and stimulate the growth of others, thereby changing community composition and potentially selecting for antibiotic resistance. Therefore, the term "sub-inhibitory" should be used carefully, since the research carried out in this PhD shows that sub-inhibitory concentrations may stimulate resistance development *in vivo* and inhibitory concentrations may not impact the environmental resistome, depending on the interactions between the antibiotic, the environmental matrix and its microbiome. Thus, these results demonstrate the need of analyzing the effects of antibiotic pollution using microcosm approaches and field studies when possible, since culture-based studies target a reduced proportion of the environmental microbiome[3] and they do not take environmental factors into account.

These two studies compared the impact of the same antibiotic on different environmental matrixes. Furthermore, the study on freshwater microcosms is, to the best of our knowledge, the first to show that sub-inhibitory concentrations of antibiotics induce a response in the environmental microbiome and resistome *in vivo* and to establish relationships between antibiotic dose and the selective force for gentamicin resistance genes. These findings are critical, since they support the concern that sub-inhibitory concentrations of antibiotics may select for resistance in the environment and, therefore, deserve more attention when assessing the risks associated to antibiotic environmental pollution.

However, further research is needed to answer several questions that were raised during this PhD and that should provide useful insights into the development of resistance in gentamicin-polluted environments and the associated mechanisms. First, the spatial heterogeneity in the soil microbiome and resistome created by environmental factors and antibiotic concentration gradients may create a different response to the same antibiotic depending on the microniche population and physiological state. Therefore, the dilution of local heterogeneity during DNA extraction at higher sample sizes could hide changes at a microenvironment scale and underestimate the risk associated with soil antibiotic pollution. Future studies should analyze the effects of antibiotics in soil microniches and account for differences related to soil spatial heterogeneity. On the other hand, the *Limnohabitans* MAG reconstructed from water metagenomic sequences was not complete and thus reduced the potential evaluation of the genetic context and mobility potential of the gentamicin resistance gene that it contained. Further research, including the isolation of *Limnohabitans* from bacterial enrichments and the sequencing of its genome, should provide a deeper understanding of both the genetic background associated to *Limnohabitans* dominance at sub-inhibitory concentrations of gentamicin and the underlying mechanisms of gene acquisition by *Limnohabitans.*

Finally, these studies evaluated the short-term effects of a single dose of a single antibiotic in a single environmental matrix. This approach leaves several unanswered questions. First, the bioavailability of gentamicin could increase at several doses and with longer exposure times in soil and cause long-term perturbations on microbial communities. The study on freshwater does not provide any insights on the response at longer exposure times nor the persistence of resistance in the absence of antibiotic pressure. Second, different soil and freshwater environments could respond differently to the same antibiotic. Patterns should be established to link the composition of the environmental matrix and/or environmental conditions and the observed response (or lack of). Third, a combination of antibiotics at sub-inhibitory concentrations may reflect more accurately the actual environmental contamination. The combination of several antibiotics at sublethal levels might alter the outcome of resistance selection *in vitro*[4]. Therefore, studies analyzing the effect combined antibiotics (and co-selective agents such as metals and biocides) at sub-inhibitory concentrations on the environmental microbiome and resistome are needed to fill knowledge gaps and provide more realistic information for risk assessment. Finally, these experiments focused on the effects of gentamicin pollution on antibiotic resistance development in the environment and they were not designed to analyze mobilization and dissemination of antibiotic resistance in the environment and the probability of transfer to the human and animal microbiomes. All these elements are key to gain a more thorough understanding of the risk posed by antibiotic environmental pollution.

Impact of anthropogenic activities on resistance development in the environment

The field study carried out in **Chapter V** aimed to evaluate the extent to which anthropogenic activities and environmental factors affect the microbiome and resistome of snow-covered mountain ecosystems. Snow samples were obtained from two watersheds located in the Karkonosze National Park in the Sudety Mountains in Poland. One of them was subjected to human activities such as tourism and cottage development and the other one was hypothetically unaffected by local anthropogenicactivity. In addition, snow with different levels of surrounding vegetation were sampled from both catchments. The impact of vegetation and anthropogenic activity on the snow microbiome and resistome was evaluated using a metagenomics and qPCR approach. Bacterial community size, composition, genus richness and diversity were the parameters chosen to determine the effects of anthropogenic and environmental factors on the snow microbiome, whereas their effects on the snow resistome were assessed by the determination of the size and composition of the antibiotic resistome, as well as ARG richness and diversity.

The main hypothesis of this research was that both environmental and anthropogenic factors influence snow ecology and induce changes in the snow microbiome and antibiotic resistome by providing bacterial communities with higher levels of organic carbon and other nutrients. This would support the growth of a more abundant bacterial community that would increase the abundance of the antibiotic resistome. Previous research performed by our group has shown that increased levels of organic acids induced competition and increased the abundance of ARGs in snow[5]. These increased concentrations also supported the establishment of a more metabolically active community[6]. Therefore, both anthropogenic and environmental sources of organic carbon could stimulate competition and ARG proliferation in snow.

The results obtained from this study support that hypothesis. First, higher levels of surrounding vegetation increased bacterial biomass, reduced genus diversity and shaped the composition of bacterial communities in snow. Several genera that were more abundant at sites with higher levels of surrounding vegetation use carbon as a nutrient source and the differences in genus abundance increased with differences in surrounding vegetation. Furthermore, the size of the antibiotic resistome increased at sites with more surrounding vegetation and this increase was related to the increase in bacterial biomass. Finally, the number and abundance of ARGs conferring resistance to all antibiotic classes identified in snow were higher at sites with higher levels of surrounding vegetation.

This research shows that anthropogenic activity induced changes in the snow antibiotic resistome while having a limited impact on the composition of the bacterial communities. Although higher bacterial biomass levels were detected at sites subjected to anthropogenic activity, anthropogenic factors did not show significant influence of genus richness, diversity and overall community composition. The differences found in genus abundance along the anthropogenic activity gradient were much lower than the ones observed along the surrounding vegetation gradient. Since sites subjected to anthropogenic activity were surrounded by vegetation and human microbiome bacteria are unlikely to survive in extreme environments, the added effects of anthropogenic waste may not necessarily affect the core bacterial community. However, the increase in bacterial community size in anthropogenic-polluted sites was followed by an increase in the size of the antibiotic resistome and the number of ARG reads associated with several antibiotic classes. Finally, nine genes involved in

resistance to six antibiotic classes were found more abundant at anthropogenic polluted sites than in pristine snow. Thus, these results support the hypothesis that human activities generate increased amounts of waste that provide the snow microbiome with additional organic carbon and stimulate competition and ARG proliferation similar to the presence of vegetation.

In conclusion, this research shows that some micro-organisms respond to environmental changes due to the surrounding vegetation and the local anthropogenic activity. This is the first study to demonstrate that winter tourism induces changes in the snow resistome without affecting the core bacterial communities. Anthropogenic activities could pollute surrounding snow through the input of organic matter present in waste and stimulate the development of antibiotic resistance in the snow microbiome that might be subsequently disseminated through the atmosphere and by snow melting. This highlights the need for a survey of antibiotic resistance development at anthropogenic polluted sites and the consideration of organic sources of pollution in addition to biological pollutants (ARB and ARGs). However, the increase in the abundance of specific ARGs at anthropogenic polluted sites should be validated by qPCR and further studies should evaluate the dissemination potential of these genes.

References

1. George, J. & Halami, P. M. Sub-inhibitory concentrations of gentamicin triggers the expression of aac(6')Ie-aph(2")Ia, chaperons and biofilm related genes in Lactobacillus plantarum MCC 3011. *Res. Microbiol.* **168,** 722–731 (2017).

2. Jutkina, J., Marathe, N. P., Flach, C. F. & Larsson, D. G. J. Antibiotics and common antibacterial biocides stimulate horizontal transfer of resistance at low concentrations. *Sci. Total Environ.* **616–617,** 172–178 (2018).

3. Rappé, M. S. & Giovannoni, S. J. The Uncultured Microbial Majority. *Annu. Rev. Microbiol.* **57,** 369–394 (2003).

4. Gullberg, E., Albrecht, L. M., Karlsson, C., Sandegren, L. & Andersson, D. I. Selection of a Multidrug Resistance Plasmid by Sublethal Levels of Antibiotics and Heavy Metals. *mBio* 5(5):e01918-14. (2014). doi:10.1128/mBio.01918-14.

5. Bergk Pinto, B., Maccario, L., Dommergue, A., Vogel, T. M. & Larose, C. Do Organic Substrates Drive Microbial Community Interactions in Arctic Snow? *Front. Microbiol.* **10,** 1–13 (2019).

6. Zhu, C. *et al.* Snow microbiome functional analyses reveal novel aspects of microbial metabolism of complex organic compounds. *Microbiologyopen* **9,** 1–12 (2020).

Annex. Supplementary Information

Chapter II. Optimization of nucleic acid extraction from soil. Evaluation of the effect of nucleic acid extraction and sequencing depth on taxonomic, functional and ARG discovery

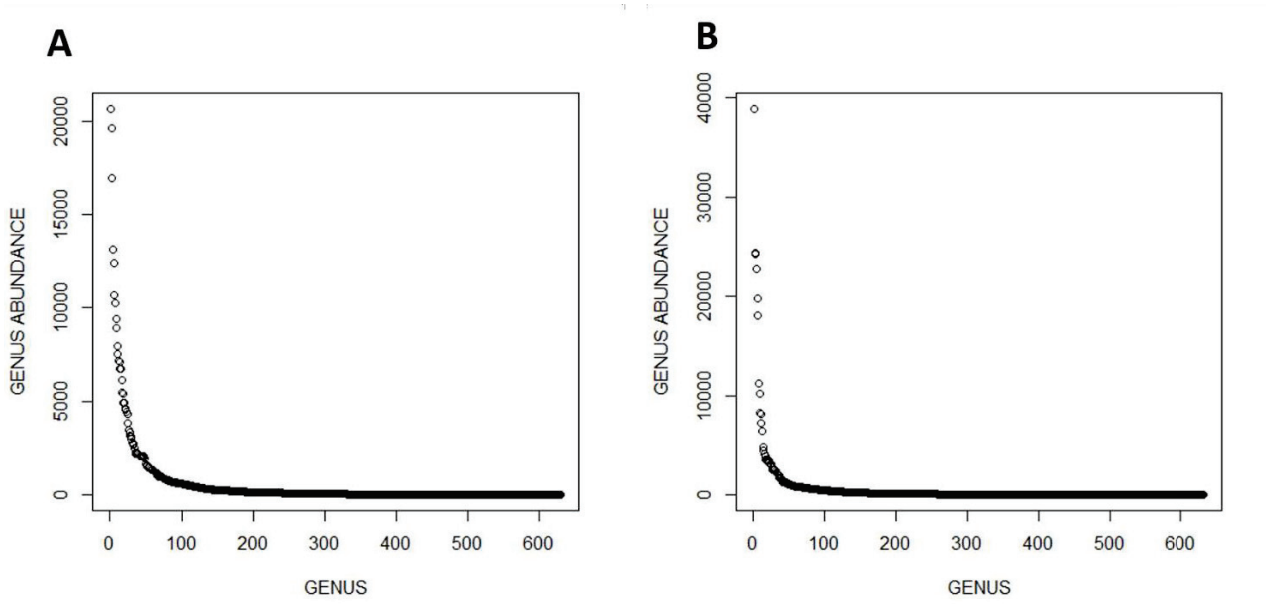


Figure S1. Total abundance of the ensemble of genera annotated using RDP classifier from DNA extracted from (A) the Scottish Agricultural College soil; (B) La Côte de Saint André soil. The abundance of each measured genus was determined and plotted in a decreasing order using R. N=3.

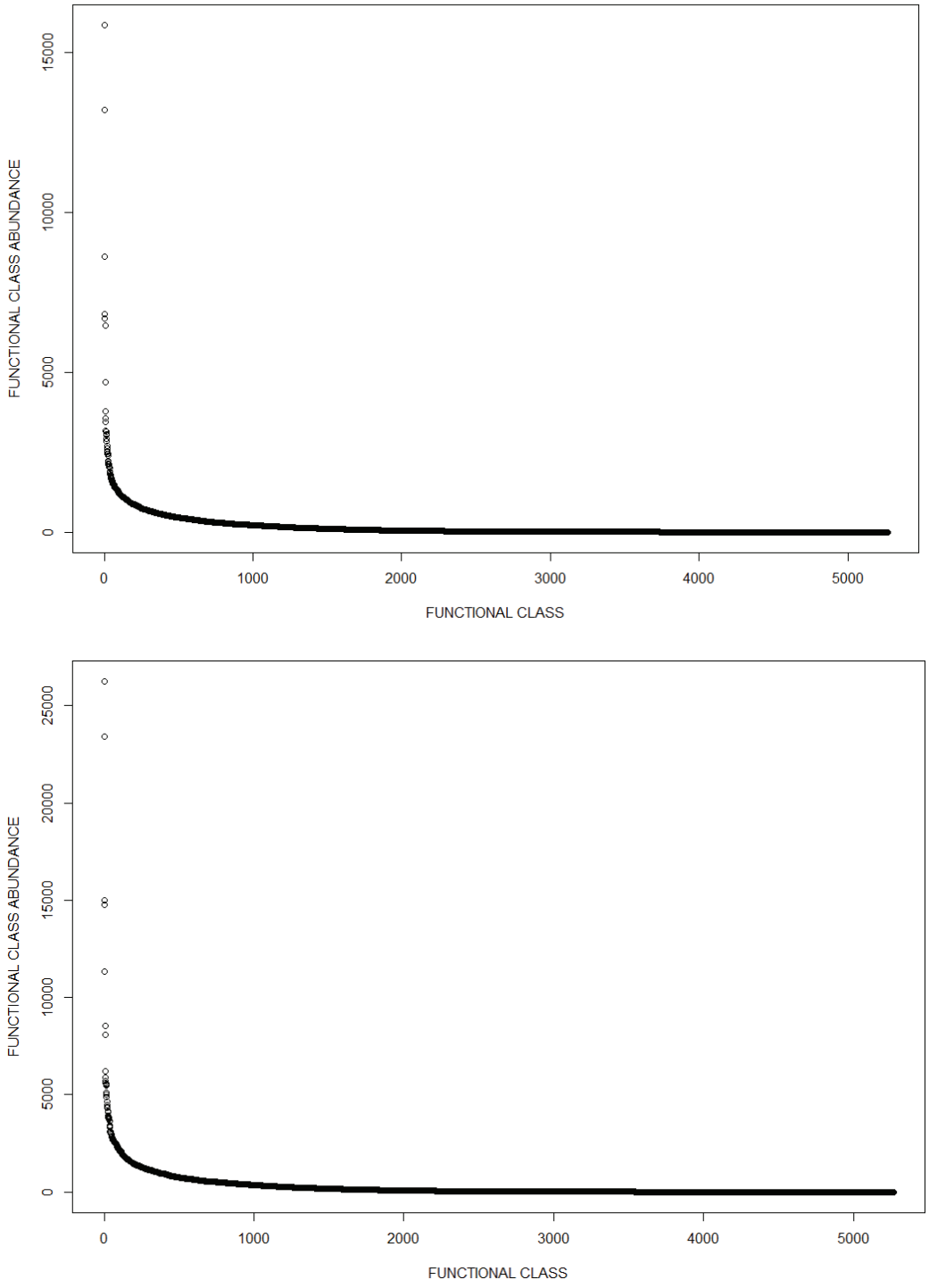


Figure S2. Total abundance of the ensemble of functional classes classified using SEED from DNA extracted from (A) the Scottish Agricultural College soil; (B) La Côte de Saint André soil. The abundance of each measured functional class was determined and plotted in a decreasing order using R. N=3.

SCOTTISH AGRICULTURAL COLLEGE SOIL								
Genus	**Number of sequences**	**Average RA**	**Standard Deviation**	**RA Maxwell 1**	**RA Maxwell 2**	**RA Phenol/ Chloroform**	**RA PowerSoil**	**RA Zymo**
Gp1	17018	5,80	1,46	5,31	7,64	5,47	6,76	3,84
Rhodanobacter	16415	5,60	1,75	4,05	6,46	7,82	6,05	3,61
Gp2	14188	4,84	1,56	3,73	7,21	4,85	5,22	3,18
Saccharibacteria genera incertae sedis	10833	3,69	1,56	4,85	5,37	2,58	4,04	1,64
Bacillus	9385	3,20	1,90	5,58	0,29	2,96	3,53	3,63
Subdivision3_genera incertae sedis	8834	3,01	1,25	3,86	4,58	2,22	2,94	1,46
Gp3	8514	2,90	0,71	2,63	4,13	2,45	2,91	2,40
Gemmatimonas	8012	2,73	2,11	1,38	2,21	5,56	4,16	0,36
WPS-2 genera incertae sedis	7130	2,43	0,39	2,10	2,00	2,81	2,39	2,85
Gaiella	6304	2,15	1,35	1,31	1,35	2,64	1,13	4,31
Conexibacter	5824	1,99	1,64	0,93	0,78	2,11	1,36	4,77
Rudaea	5821	1,99	0,50	1,48	2,53	1,42	2,30	2,20
Arthrobacter	5708	1,95	0,97	1,62	0,64	3,08	1,66	2,73
Rhizomicrobium	5563	1,90	0,37	1,46	2,45	2,03	1,70	1,85
Spartobacteria genera incertae sedis	5446	1,86	0,34	1,48	1,66	1,71	2,26	2,17
Candidatus Solibacter	5050	1,72	0,47	1,54	2,54	1,40	1,43	1,70
Phenylobacterium	4489	1,53	0,71	1,07	2,46	1,12	0,87	2,13
Roseiarcus	4313	1,47	0,93	0,83	1,56	0,92	0,99	3,05
Povalibacter	4030	1,37	0,27	1,00	1,64	1,60	1,20	1,44
Ktedonobacter	3968	1,35	0,89	1,61	0,38	2,55	1,66	0,57
Mucilaginibacter	3811	1,30	0,61	1,54	1,85	1,01	1,72	0,37
Parcubacteria genera incertae sedis	3733	1,27	0,48	1,51	1,86	1,14	1,28	0,57
Rhodoferax	3613	1,23	0,44	1,01	1,69	1,67	1,12	0,67
Nakamurella	3537	1,21	0,86	0,65	0,09	2,14	1,22	1,92
WPS-1_genera incertae sedis	3171	1,08	0,30	0,91	1,57	0,99	1,15	0,78

Table S1. Average relative abundance (RA) and relative abundances per method of the 25 most abundant genera from the rarefied pool of sequences from the Scottish Agricultural College soil. Maxwell 1 and 2: variants of the Maxwell RSC PureFood GMO and Authentication Kit (Promega). PowerSoil: DNeasy PowerSoil Kit (QIAGEN); Zymo: ZymoBIOMICS DNA Mini Kit (Zymo Research). Relative abundances higher than the average plus one standard deviation are highlighted in orange, and those lower than the average minus one standard deviation are highlighted in blue. N=3.

LA COTE DE SAINT ANDRE SOIL								
Genus	**Number of sequences**	**Average RA**	**Standard Deviation**	**RA Maxwell 1**	**RA Maxwell 2**	**RA Phenol/ Chloroform**	**RA PowerSoil**	**RA Zymo**
Spartobacteria_genera_incertae_sedis	20362	11,22	1,79	9,91	13,50	9,37	10,62	12,68
Gp1	11785	6,49	1,94	7,52	8,29	5,15	7,73	3,76
Gaiella	11774	6,49	5,06	3,15	1,72	14,74	6,80	6,02
Gp6	10806	5,95	1,89	6,68	6,96	6,05	7,39	2,68
Subdivision3 genera incertae sedis	10177	5,61	2,26	6,08	8,62	2,30	5,17	5,86
Gp3	9035	4,98	1,02	5,46	6,22	4,20	5,32	3,68
Bacillus	5040	2,78	2,32	3,53	0,30	2,09	1,57	6,38
Gemmatimonas	4987	2,75	1,73	3,50	3,15	2,43	4,65	0,01
Gp2	4807	2,65	1,53	3,57	4,10	1,38	3,55	0,64
Candidatus Solibacter	4164	2,29	0,28	2,21	2,59	1,88	2,53	2,26
WPS-1 genera incertae sedis	3450	1,90	0,53	1,91	2,49	1,70	2,27	1,13
Gp16	3329	1,83	1,26	0,84	0,72	3,75	1,50	2,36
Saccharibacteria genera incertae sedis	2483	1,37	0,50	1,08	1,67	1,01	0,97	2,11
Gp5	2440	1,34	0,66	1,53	2,33	0,63	1,37	0,85
Bradyrhizobium	1981	1,09	0,62	0,90	0,45	1,27	0,78	2,06
Ktedonobacter	1919	1,06	0,67	1,42	0,23	1,90	1,18	0,55
Nitrospira	1879	1,03	0,65	1,41	1,74	0,59	1,29	0,14
Paenibacillus	1872	1,03	0,81	1,24	0,10	0,85	0,69	2,28
Chryseolinea	1820	1,00	0,51	1,23	1,40	0,44	0,48	1,47
Rhizomicrobium	1789	0,99	0,05	1,06	0,99	0,94	0,98	0,95
Candidatus Koribacter	1742	0,96	0,14	1,04	1,09	0,97	0,95	0,74
Parcubacteria genera incertae sedis	1676	0,92	0,40	0,85	1,58	0,62	0,59	0,98
Gp4	1675	0,92	0,52	1,01	1,72	0,66	0,88	0,34
Aridibacter	1645	0,91	0,42	0,87	1,56	0,40	0,76	0,94
Pirellula	1599	0,88	0,42	0,95	1,45	0,75	0,96	0,29

Table S2. Average relative abundance (RA) and relative abundances per method of the 25 most abundant genera from the rarefied pool of sequences from La Côte de Saint André soil. Maxwell 1 and 2: variants of the Maxwell RSC PureFood GMO and Authentication Kit (Promega). PowerSoil: DNeasy PowerSoil Kit (QIAGEN); Zymo: ZymoBIOMICS DNA Mini Kit (Zymo Research). Relative abundances higher than the average plus one standard deviation are highlighted in orange, and those lower than the average minus one standard deviation are highlighted in blue. N=3.

SCOTTISH AGRICULTURAL COLLEGE SOIL							
Functional class	**Number of sequences**	**Average RA**	**Standard Deviation**	**RA Maxwell 1**	**RA Maxwell 2**	**RA Phenol/ Chloroform**	**RA PowerSoil**
5-FCL-like protein	20362	11,22	1,79	9,91	13,50	9,37	10,62
Long-chain-fatty-acid--CoA ligase	11785	6,49	1,94	7,52	8,29	5,15	7,73
3-oxoacyl-[acyl-carrier protein] reductase	11774	6,49	5,06	3,15	1,72	14,74	6,80
TonB-dependent receptor	10806	5,95	1,89	6,68	6,96	6,05	7,39
Cobalt-zinc-cadmium resistance protein CzcA	10177	5,61	2,26	6,08	8,62	2,30	5,17
Adenylate cyclase	9035	4,98	1,02	5,46	6,22	4,20	5,32
COG2363	5040	2,78	2,32	3,53	0,30	2,09	1,57
Butyryl-CoA dehydrogenase	4987	2,75	1,73	3,50	3,15	2,43	4,65
Enoyl-CoA hydratase	4807	2,65	1,53	3,57	4,10	1,38	3,55
Aldehyde dehydrogenase	4164	2,29	0,28	2,21	2,59	1,88	2,53
Acriflavin resistance protein	3450	1,90	0,53	1,91	2,49	1,70	2,27
Aspartate aminotransferase	3329	1,83	1,26	0,84	0,72	3,75	1,50
Beta-lactamase	2483	1,37	0,50	1,08	1,67	1,01	0,97
UDP-glucose 4-epimerase	2440	1,34	0,66	1,53	2,33	0,63	1,37
3-ketoacyl-CoA thiolase	1981	1,09	0,62	0,90	0,45	1,27	0,78
DNA-directed RNA polymerase beta' subunit	1919	1,06	0,67	1,42	0,23	1,90	1,18
Diguanylate cyclase/phosphodiesterase	1879	1,03	0,65	1,41	1,74	0,59	1,29
DNA-directed RNA polymerase beta subunit	1872	1,03	0,81	1,24	0,10	0,85	0,69
DNA polymerase III alpha subunit	1820	1,00	0,51	1,23	1,40	0,44	0,48
Thioredoxin reductase	1789	0,99	0,05	1,06	0,99	0,94	0,98
FIG039061: hypothetical protein related to heme utilization	1742	0,96	0,14	1,04	1,09	0,97	0,95
5-methyltetrahydrofolate--homocysteine methyltransferase	1676	0,92	0,40	0,85	1,58	0,62	0,59
Acetyl-coenzyme A synthetase	1675	0,92	0,52	1,01	1,72	0,66	0,88
D-3-phosphoglycerate dehydrogenase	1645	0,91	0,42	0,87	1,56	0,40	0,76
Carbamoyl-phosphate synthase large chain	1599	0,88	0,42	0,95	1,45	0,75	0,96

Table S3. Average relative abundance (RA) and relative abundances per method of the 25 most abundant functional classes from the rarefied pool of sequences from the Scottish Agricultural College soil. Maxwell 1 and 2: variants of the Maxwell RSC PureFood GMO and Authentication Kit (Promega). PowerSoil: DNeasy PowerSoil Kit (QIAGEN). Relative abundances higher than the average plus one standard deviation are highlighted in orange, and those lower than the average minus one standard deviation are highlighted in blue. N=3.

LA COTE DE SAINT ANDRE SOIL							
Functional class	**Number of sequences**	**Average RA**	**Standard Deviation**	**RA Maxwell 1**	**RA Maxwell 2**	**RA Phenol/ Chloroform**	**RA PowerSoil**
5-FCL-like protein	12027	1,88	0,146	1,82	1,70	2,01	2,00
Long-chain-fatty-acid--CoA ligase	10829	1,69	0,065	1,77	1,62	1,72	1,67
3-oxoacyl-[acyl-carrier protein] reductase	6756	1,06	0,096	1,08	0,92	1,14	1,08
Adenylate cyclase	6685	1,05	0,070	1,03	0,96	1,07	1,13
Cobalt-zinc-cadmium resistance protein CzcA	5687	0,89	0,190	0,96	1,11	0,67	0,82
COG2363	4041	0,63	0,047	0,65	0,69	0,61	0,58
TonB-dependent receptor	3978	0,62	0,158	0,68	0,81	0,44	0,56
Arylsulfatase	2761	0,43	0,021	0,46	0,44	0,41	0,42
Butyryl-CoA dehydrogenase	2753	0,43	0,051	0,41	0,37	0,49	0,46
Beta-lactamase	2666	0,42	0,022	0,42	0,44	0,39	0,42
Acriflavin resistance protein	2666	0,42	0,044	0,44	0,47	0,36	0,41
UDP-glucose 4-epimerase	2550	0,40	0,015	0,38	0,41	0,39	0,41
Enoyl-CoA hydratase	2532	0,40	0,041	0,38	0,35	0,41	0,44
Aldehyde dehydrogenase	2450	0,38	0,039	0,40	0,32	0,41	0,40
Aspartate aminotransferase	2406	0,38	0,019	0,38	0,40	0,37	0,36
DNA-directed RNA polymerase beta' subunit	2360	0,37	0,016	0,36	0,37	0,36	0,39
DNA polymerase III alpha subunit	2290	0,36	0,018	0,34	0,35	0,38	0,36
DNA-directed RNA polymerase beta subunit	2122	0,33	0,014	0,32	0,35	0,34	0,33
Excinuclease ABC subunit A	2079	0,33	0,027	0,35	0,35	0,30	0,30
3-ketoacyl-CoA thiolase	2045	0,32	0,040	0,32	0,27	0,36	0,33
Acetyl-coenzyme A synthetase	2009	0,31	0,038	0,31	0,26	0,35	0,33
D-3-phosphoglycerate dehydrogenase	1988	0,31	0,022	0,29	0,34	0,30	0,31
Diguanylate cyclase/phosphodiesterase	1931	0,30	0,018	0,33	0,30	0,29	0,29
Thioredoxin reductase	1926	0,30	0,020	0,31	0,28	0,30	0,32
Hypothetical protein related to heme utilization	1908	0,30	0,010	0,29	0,31	0,29	0,31

Table S4. Average relative abundance (RA) and relative abundances per method of the 25 most abundant functional classes from the rarefied pool of sequences from La Côte de Saint André soil. Maxwell 1 and 2: variants of the Maxwell RSC PureFood GMO and Authentication Kit (Promega). PowerSoil: DNeasy PowerSoil Kit (QIAGEN). Relative abundances higher than the average plus one standard deviation are highlighted in orange, and those lower than the average minus one standard deviation are highlighted in blue. N=3.

SCOTTISH AGRICULTURAL COLLEGE SOIL			
Method	**Genus**	**Annotated sequences after rarefaction**	**Annotated sequences before rarefaction**
Maxwell 1	Gp22	2	2
	Arenicella	2	2
	Phyllobacterium	2	2
	Falsirhodobacter	2	2
	Hydrogenophaga	1	4
	Geothermomicrobium	2	2
	Curtobacterium	2	3
	Nitrosopumilus	3	4
Maxwell 2	Parapedobacter	2	2
	Chelatococcus	2	2
	Amantichitinum	2	2
	Leucobacter	4	4
Phenol/Chloroform method	Acaricomes	2	2
	Sphaerobacter	3	3
	Salinispira	2	2
DNeasy PowerSoil Kit	Syntrophobacter	2	2
	Chitinibacter	2	2
	Thermomarinilinea	3	3
ZymoBIOMICS DNA Mini Kit	Fulvimonas	2	2
	Nitratireductor	2	2
	Paenochrobactrum	1	2
	Sorangium	2	2
	Desulfomonile	2	2
	Hungatella	2	2
	Staphylococcus	63	95
	Thermocrispum	2	2

Table S5. List of genera detected as unique in both rarefied and non-rarefied pools of sequences from the Scottish Agricultural College soil. Maxwell 1 and 2: variants of the Maxwell RSC PureFood GMO and Authentication Kit (Promega). N=3.

LA COTE DE SAINT ANDRE SOIL			
Method	**Genus**	**Annotated sequences after rarefaction**	**Annotated sequences before rarefaction**
Maxwell 1	Limnobacter	2	2
	Terribacillus	1	2
	GpXIII	2	3
Maxwell 2	Epilithonimonas	1	3
	Pseudenhygromyxa	4	4
	Thioreductor	2	2
Phenol/Chloroform method	Sporomusa	2	2
	Actinocatenispora	3	3
DNeasy PowerSoil Kit	Tolumonas	1	2
	Nereida	1	2
	Smithella	1	2
ZymoBIOMICS DNA Mini Kit	Alistipes	2	2
	Desulfomonile	2	2
	Nitrosomonas	6	6
	Gracilibacillus	6	6
	Staphylococcus	130	130
	Anaerosinus	3	3
	Melioribacter	2	2

Table S6. List of genera detected as unique in both rarefied and non-rarefied pools of sequences from La Côte de Saint André soil. Maxwell 1 and 2: variants of the Maxwell RSC PureFood GMO and Authentication Kit (Promeg). N=3.

Table S7a. List of functions detected as unique in both rarefied and non-rarefied pools of sequences from the Scottish Agricultural College soil. N=3.

SCOTTISH AGRICULTURAL COLLEGE SOIL			
Method	Function	Annotated sequences after rarefaction	Annotated sequences before rarefaction
Maxwell 1	Transcriptional antiterminator of lichenan operon, BglG family	1	2
	Pectate lyase precursor (EC 4.2.2.2)	2	2
	Positive regulator of L-idonate catabolism	2	2
	Threonine kinase in B12 biosynthesis	2	2
	Amidohydrolase YlmB, involved in salvage of thiamin pyrimidine moiety	2	2
	Putative 2Fe-2S ferredoxin CbiW involved in B12 biosynthesis	2	4
	MutS domain protein, family 6	1	2
	Toluene-4-monooxygenase, subunit TmoC	2	2
	Cytochrome c-type heme lyase subunit nrfE, nitrite reductase complex assembly	1	2
	LSU ribosomal protein L35e (L29p)	1	2
	Electron transport complex protein RnfE	2	2
Maxwell 2	D-erythrulose 4-phosphate dehydrogenase EryC (EC 1.1.1.-)	2	2
	L-rhamnonate dehydratase (EC 4.2.1.90)	2	2
	ABC transporter, predicted N-acetylneuraminate transport system permease protein 1	2	2
	Phycocyanobilin lyase alpha subunit	2	2
	Coenzyme PQQ synthesis protein A	2	2
	Flap structure-specific endonuclease (EC 3.-.-.-)	2	2
	Cryptochrome	2	2
	enoyl-CoA hydratase, R-specific	2	2
	Nitric oxide reductase activation protein NorE	2	2
	Formate hydrogenlyase subunit 2	2	2
	FIG098787: protein involved in DMSP breakdown	2	2

	IncF plasmid conjugative transfer pilus assembly protein TraL	2	2
Phenol/ Chloroform method	Indolepyruvate oxidoreductase subunit IorB (EC 1.2.7.8)	1	2
	3,7-dideoxy-D-threo-hepto-2, 6-diulosonate synthase (EC 1.4.1.24)	1	2
	Beta-glucoside bgl operon antiterminator, BglG family	1	2
	Predicted L-arabinose ABC transport system, periplasmic arabinose-binding protein	1	2
	Glycine reductase component B alpha subunit (EC 1.21.4.2)	1	2
	putative isobutyryl-CoA mutase, chain B	2	2
	Adrenodoxin, mitochondrial ferredoxin	2	2
	Chlorophyll a synthase ChlG (EC 2.5.1.62)	1	2
	probable iron binding protein for iron-sulfur cluster assembly	1	2
	Acyl-CoA dehydrogenase (EC 1.3.8.1), Mycobacterial subgroup FadE32	1	2
	Fimbrial assembly protein FimB	1	2
	benABC operon transcriptional activator BenR	1	2
	Hydroxyaromatic non-oxidative decarboxylase protein B (EC 4.1.1.-)	1	2
	p-cumate dioxygenase small subunit (CmtAc)	1	2
	Carboxy-terminal intein-mediated trans-splice	1	2
	VapB5 protein (antitoxin to VapC5)	1	2
	quinol oxidase polypeptide I QoxB (EC:1.9.3.-)	1	2
	Predicted transcriptional regulator of 4-carboxymuconolactone decarboxylase, Rrf2 family	1	2
	bacteriophytochrome heme oxygenase BphO	1	2
	Fe2+/Zn2+ uptake regulation proteins	2	2
	Multidrug resistance protein [function not yet clear]	1	2
DNeasy® PowerSoil® Kit	Succinylglutamic semialdehyde dehydrogenase (EC 1.2.1.71)	1	13
	Amino acid-binding ACT	1	4
	Quinate/shikimate dehydrogenase [Pyrroloquinoline-quinone] (EC 1.1.99.25)	1	10
	Hca operon (3-phenylpropionic acid catabolism) transcriptional activator HcaR	1	3
	Saccharopine dehydrogenase (EC 1.5.1.9)	1	8

Transcriptional activator MetR	1	14
2-hydroxy-3-keto-5-methylthiopentenyl-1-phosphate phosphatase related protein	1	11
Gamma-glutamyl-putrescine oxidase (EC1.4.3.-)	1	2
Serine hydroxymethyltransferase 1 (EC 2.1.2.1)	1	3
Acetoin catabolism protein X	1	10
Transcriptional regulator of succinyl CoA synthetase operon	1	4
Intracellular alpha-amylase (EC 3.2.1.1)	1	3
Uronate dehydrogenase (EC 1.1.1.203)	1	6
2,5-diketo-D-gluconate reductase A (EC 1.1.1.274)	2	7
High-affinity gluconate transporter GntT	1	2
D-glycerate 3-kinase (EC 2.7.1.31), plant type	1	5
Transcriptional regulatory protein UhpA	1	3
Possible GPH family transporter (TC 2.A.2) for arabinosides	2	9
Predicted rhamnose oligosaccharide ABC transport system, permease component 2	2	9
Predicted rhamnose oligosaccharide ABC transporter, permease component 1	1	5
Rhamnulokinase RhaK in alpha-proteobacteria (EC 2.7.1.5)	2	15
Malonate decarboxylase beta subunit	1	10
Malonate decarboxylase delta subunit	1	4
Maltose regulon modulator	1	6
Maltose-6'-phosphate glucosidase (EC 3.2.1.122)	1	13
Endo-1,4-beta-mannosidase	1	5
Endo-mannanase (EC 3.2.1.78)	1	5
Formylmethanofuran dehydrogenase (tungsten) subunit B (EC 1.2.99.5)	1	2
Methanol:corrinoid methyltransferase	1	4
Two-component sensor histidine kinase, malate (EC 2.7.3.-)	1	2
Xylose activator XylR (AraC family)	1	3
Cell division topological specificity factor MinE	1	16
RNA polymerase sigma-H factor AlgT	1	17
Mannosyltransferase WboB	1	2

Capsular polysaccharide export system protein KpsS	1	7
Colanic acid biosynthesis glycosyl transferase WcaE	1	8
Legionaminic acid cytidylyltransferase (EC 2.7.7.43)	1	3
Long-chain fatty-acid-AMP ligase, Mycobacterial subgroup FadD26	1	2
Phenolpthiocerol synthesis type-I polyketide synthase PpsB (EC 2.3.1.41)	1	13
Cyclopropane-fatty-acyl-phospholipid synthase 2, CmaA2 (EC 2.1.1.79)	1	2
Xanthan biosynthesis chain length determinant protein GumC	1	3
Thiamin-phosphate synthase ThiN (EC 2.5.1.3)	1	9
CblZ, a non-orthologous displasment for Alpha-ribazole-5'-phosphate phosphatase	1	3
Cobalamin biosynthesis protein CbiG	1	8
Protoporphyrin IX Mg-chelatase subunit H (EC 6.6.1.1)	1	16
Sulfopyruvate decarboxylase - alpha subunit (EC 4.1.1.79)	1	8
FIG000605: protein co-occurring with transport systems (COG1739)	1	23
Dihydroneopterin triphosphate pyrophosphohydolase	1	2
Possible glutaminyl transferase clustered with tetrahydromethanopterin biosynthesis genes	1	10
Regulatory protein (induces abgABT, used to catabolize p-aminobenzoyl-glutamate)	1	2
Substrate-specific component QueT (COG4708) of predicted queuosine-regulated ECF transporter	1	4
CRISPR-associated helicase Cas3, protein	1	12
CRISPR-associated protein Cas2	1	13
CRISPR-associated protein Csb1	1	18
DinG family ATP-dependent helicase YpvA	1	5
Archaeal DNA polymerase I (EC 2.7.7.7)	2	28
Replication factor A (ssDNA-binding protein)	2	13
Putative deoxyribonuclease similar to YcfH, type 4	1	14
Spore cortex-lytic enzyme, N-acetylglucosaminidase SleL (EC 3.2.1.-)	2	10
Spore germination protein GerKC	1	10
Stage IV sporulation protein B	1	6

Stage V sporulation protein AF (SpoVAF)	1	9
Stage V sporulation protein required for dehydratation of the spore core and assembly of the coat (SpoVS)	1	8
Hydroxyneurosporene synthase	1	4
Acyl-CoA dehydrogenase (EC 1.3.8.1), Mycobacterial subgroup FadE14	1	2
Acyl-CoA dehydrogenase (EC 1.3.8.1), Mycobacterial subgroup FadE21	1	4
Acyl-CoA dehydrogenase (EC 1.3.8.1), Mycobacterial subgroup FadE24	2	8
Pyoverdine sidechain non-ribosomal peptide synthetase PvdD	1	9
Substrate-specific component NikM of nickel ECF transporter	3	6
MSHA biogenesis protein MshE	2	17
Inner membrane protein forms channel for type IV secretion of T-DNA complex (VirB8)	1	13
Na(+) H(+) antiporter subunit A (TC 2.A.63.1.2)	1	3
Manganese uptake regulation protein MUR	1	4
Phosphoprotein phosphatase PppA	1	3
HCOMODA decarboxylase	1	2
Transcriptional activator protein LuxR	1	8
conserved protein with predicted RNA binding PUA domain	1	3
Nucleotidase YfbR, HD superfamily	2	10
FIG007491: hypothetical protein YeeN	1	2
Nitric oxide reductase activation protein NorQ	1	15
Nitrous oxide reductase maturation periplasmic protein NosX	2	16
Nitrite transporter NirC	1	3
Nitrogenase FeMo-cofactor carrier protein NifX	1	2
Pseudouridine 5'-phosphate glycosidase	2	19
Phage major capsid protein	1	3
Integron integrase IntI2	1	3
Alkylphosphonate utilization operon protein PhnA	1	9
Phosphonoacetaldehyde hydrolase (EC 3.11.1.1)	1	17
Maleylacetate reductase (EC 1.3.1.32)	2	9
FIG001553: Hydrolase, HAD subfamily IIIA	1	2

N-linked glycosylation glycosyltransferase PglG	1	2
Chaperone-modulator protein CbpM	1	2
Quinohemoprotein amine dehydrogenase alpha subunit (EC 1.4.99.-)	1	4
Similar to ribosomal large subunit pseudouridine synthase D, Bacillus subtilis YjbO type	1	6
LSU ribosomal protein L23Ae (L23p)	1	6
LSU ribosomal protein L5e (L18p)	1	11
LSU ribosomal protein L7Ae	1	10
LSU ribosomal protein P0 (L10p)	1	22
SSU ribosomal protein S15Ae (S8p)	1	16
SSU ribosomal protein S16e (S9p)	2	11
SSU ribosomal protein S28e	1	5
SSU ribosomal protein S3e (S3p)	1	12
SSU ribosomal protein S4e	3	18
SCIFF radical SAM maturase	1	2
Ribosome protection-type tetracycline resistance related proteins, group 2	1	13
Threonyl-tRNA synthetase-related protein	2	14
Programmed cell death antitoxin MazE	2	11
Programmed cell death toxin PemK	1	5
Adenylylsulfate reductase beta-subunit (EC 1.8.99.2)	1	12
Coenzyme F420 hydrogenase beta subunit (FrcB) (EC 1.12.98.1)	1	5
tRNA-intron endonuclease (EC 3.1.27.9)	1	16
Isoflavone reductase homolog P3 (EC 1.3.1.-)	1	2
Probable monothiol glutaredoxin GrlA	1	2
Nudix dNTPase DR1776 (EC 3.6.1.-)	1	3
cytochrome c subunit of flavocytochrome c sulfide dehydrogenase	1	2
Neuraminidase NanP	1	25
Zn(II) and Co(II) transmembrane diffusion facilitator	1	2
ESAT-6-like protein EsxL	1	3

Table S7b. List of functions detected as unique in both rarefied and non-rarefied pools of sequences from La Côte de Saint André soil. N=3.

	LA COTE DE SAINT ANDRE SOIL		
Method	Function	Annotated sequences after rarefaction	Annotated sequences before rarefaction
Maxwell 1	4-amino, 4-deoxychorismate mutase (EC 5.4.99.-)	2	**2**
	Ketoisovalerate oxidoreductase subunit VorD (EC 1.2.7.7)	2	**2**
	PTS system, cellobiose-specific IIB component (EC 2.7.1.69)	2	**2**
	PTS system, diacetylchitobiose-specific IIC component (EC 2.7.1.69)	2	**2**
	Hypothetical oxidoreductase YqhD (EC 1.1.-.-)	2	**2**
	Glycoside hydrolase family 2, sugar binding	2	**2**
	Xylose oligosaccharides ABC transporter, ATP-binding protein 1	2	**2**
	D-xylose proton-symporter XylT	2	**2**
	Colanic acid biosysnthesis glycosyl transferase WcaI	2	**2**
	ADP-heptose--lipooligosaccharide heptosyltransferase, putative	2	**2**
	Substrate-specific component NiaX of predicted niacin ECF transporter	2	**2**
	CRISPR-associated protein Csc2	2	**2**
	SinR, regulator of post-exponential-phase responses genes (competence and sporulation)	2	**2**
	Spore germination protein GerHB/GerIB	2	**2**
	Spore germination protein GerLA	2	**2**
	Spore germination protein GerLC	2	**2**
	Stage V sporulation protein B	2	**2**
	KapD, inhibitor of KinA pathway to sporulation	2	**2**
	Sporulation sigma-E factor processing peptidase (SpoIIGA)	2	**2**
	Acyl-CoA dehydrogenase (EC 1.3.8.1), Mycobacterial subgroup FadE14	2	**2**
	Dehydrosqualene desaturase (EC 1.14.99.-) (Diapophytoene desaturase) (4,4'-diapophytoene desaturase)	2	**2**

	Hypothetical protein GBAA1986 associated with anthrachelin biosynthesis, isomerase-like TIM barrel domain	2	**2**
	DotA protein	2	**2**
	DotD	2	**2**
	5'-methylthioadenosine/S-adenosylhomocysteine nucleosidase related protein BA2564	2	**2**
	Periplasmic nitrate reductase component NapE	2	**2**
	Nitrogenase (molybdenum-iron) reductase and maturation protein NifH	1	**2**
	Possible malonate-semialdehyde dehydrogenase in pyrimidine catabolism	2	**2**
	Predicted flavin reductase RutF in novel pyrimidine catabolism pathway	3	**3**
	Hypoxanthine/guanine permease PbuG	2	**2**
	Glycine/sarcosine/betaine reductase component C chain 2	2	**2**
	V-type ATP synthase subunit K (EC 3.6.3.14)	2	**2**
	AA3-600 quinol oxidase subunit IIII	2	**2**
	Hypothetical ATP-binding protein, containing DUF265 domain	2	**2**
	ADP-ribose pyrophosphatase, mitochondrial precursor (EC 3.6.1.13)	2	**2**
	bacteriophytochrome heme oxygenase BphO	2	**2**
	Phage shock protein E	2	**2**
	SgrR, sugar-phosphate stress, transcriptional activator of SgrS small RNA	2	**2**
	Fosfomycin resistance protein FosA	2	**2**
Maxwell 2	CsoS1D-associated protein 3	2	**2**
	Bicarbonate transporter, bicarbonate binding protein	2	**2**
	ribulose 1,5-bisphosphate carboxylase/oxygenase activase	1	**2**
	Beta-glucoside ABC transport system, sugar-binding protein	2	**2**
	Glucose ABC transporter, permease component	1	**2**
	D-erythrulose 4-phosphate dehydrogenase EryC (EC 1.1.1.-)	1	**2**
	Glycogen biosynthesis protein GlgD, glucose-1-phosphate adenylyltransferase family	2	**2**
	Hexose phosphate uptake regulatory protein UhpC	3	**5**
	Malonate transporter, MadL subunit	2	**3**
	Methyl coenzyme M reductase alpha subunit (EC 2.8.4.1)	2	**2**
	Methyl coenzyme M reductase beta subunit (EC 2.8.4.1)	1	**2**

polyhydroxyalkanoate synthesis repressor PhaR	1	**2**
Signal recognition particle associated protein	2	**3**
GDP-mannose 4,6 dehydratase Gmd (EC 4.2.1.47)	3	**3**
DltD protein	1	**2**
Mg protoporphyrin IX monomethyl ester oxidative cyclase (aerobic) (EC 1.14.13.81)	2	**2**
Phosphopantetheine adenylyltransferase, type II eukaryotic (EC 2.7.7.3)	2	**2**
Cysteate synthase	2	**2**
PUA-PAPS reductase like fusion	1	**2**
Cytochrome cd1 nitrite reductase (EC:1.7.2.1)	2	**2**
Riboflavin synthase archaeal (EC 2.5.1.9)	2	**2**
Acyclic terpenes utilization regulator AtuR, TetR family	2	**2**
Hemin-binding lipoprotein HbpA	2	**2**
Pyridoxamine 5'-phosphate oxidase-related putative heme iron utilization protein	1	**2**
Pyochelin biosynthetic protein PchC, predicted thioesterase	2	**2**
Outer membrane pyoverdine eflux protein	1	**2**
IcmJ (DotN) protein	1	**2**
MSHA pilin protein MshA	2	**2**
Riboflavin transporter RibN	2	**2**
P-hydroxylaminobenzoate lyase	2	**2**
putative ring-hydroxylating dioxygenase small subunit	1	**2**
Hydroxyaromatic non-oxidative decarboxylase protein B (EC 4.1.1.-)	2	**3**
regulator protein,SalR	1	**2**
NADH dehydrogenase [ubiquinone] iron-sulfur protein 1 (EC 1.6.5.3) (EC 1.6.99.3)	2	**2**
Ferredoxin-type protein NapG (periplasmic nitrate reductase)	1	**2**
Phage minor capsid protein	3	**3**
hypothetical protein clusted with conjugative transposons, BF0131	1	**2**
Photosystem II stability/assembly factor HCF136/Ycf48	2	**3**
pH adaptation potassium efflux system d	2	**2**
pH adaption potassium efflux system protein PhaD	1	**2**
Chaperone-modulator protein CbpM	2	**2**
V-type ATP synthase subunit D (EC 3.6.3.14)	3	**3**

	NADH-reducing hydrogenase maturation factor	5	**5**
	[Fe] hydrogenase, HymB subunit, putative	1	**2**
	Sodium-transporting N-ATPase gamma subunit	1	**2**
	Sulfite reduction-associated complex DsrMKJOP iron-sulfur protein DsrO (=HmeA)	2	**2**
	DNA-directed RNA polymerases I, II, and III 8.3 kDa polypeptide (EC 2.7.7.6)	1	**2**
	Predicted RNA-binding protein COG1491	2	**2**
	ADP-ribose 1-phosphate phophatase	1	**2**
	Probable glutathione S-transferase (EC 2.5.1.18), YfcG homolog	2	**2**
	Superoxide reductase (EC 1.15.1.2)	2	**3**
	IncF plasmid conjugative transfer protein TraN	2	**2**
	Ynd	2	**2**
	ABC transporter, ATP-binding protein	1	**3**
Phenol / Chloroform method	Arginine ABC transporter, permease protein ArtQ	1	**3**
	Arginine/ornithine ABC transporter, ATP-binding protein AotP	2	**3**
	Arginine/ornithine ABC transporter, periplasmic arginine/ornithine binding protein	1	**2**
	Gamma-glutamyltranspeptidase PgsD/CapD (EC 2.3.2.2), catalyses PGA anchorage to peptidoglycan	2	**2**
	Putative GntR-family regulatory protein and aminotransferase near polyamine transporter	1	**3**
	S-adenosylmethionine decarboxylase proenzyme (EC 4.1.1.50), prokaryotic class 1A	1	**2**
	P450 cytochrome, component of G-protein-coupled receptor (GPCR) system	1	**3**
	Urea channel UreI	2	**6**
	Arabinose sensor protein	1	**4**
	Predicted cellobiose ABC transport system, permease protein 1	2	**2**
	Tagatose-6-phosphate kinase GatZ (EC 2.7.1.144)	1	**2**
	Putative beta-glucuronidase	1	**4**
	Malonate decarboxylase delta subunit	1	**3**
	Pullulanase (EC 3.2.1.41)	1	**3**
	Beta-mannan and fructose induced hypothetical ABC transporter, permease protein 1	1	**2**

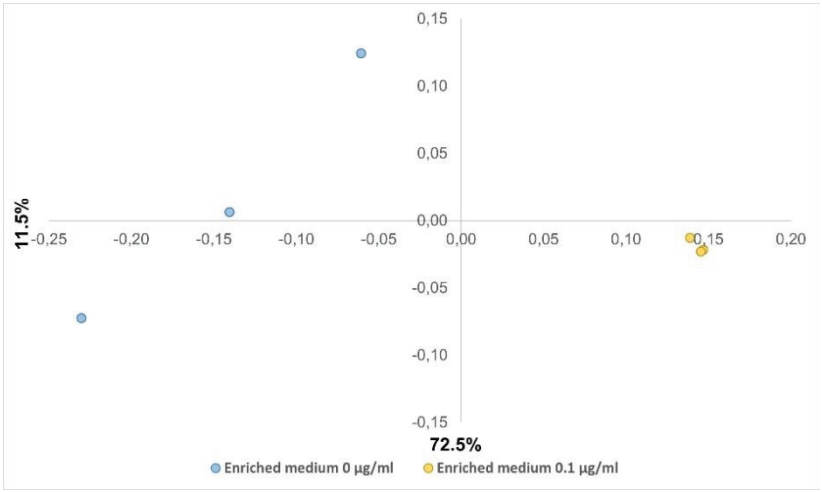
	Predicted glycosylase, COG2152	1	**3**
	Predicted exported glycosyl hydrolase family 31 protein	2	**2**
	Trehalose 6-phosphate phosphorylase (EC 2.4.1.216)	1	**3**
	LysR family transcriptional regulator YhjC	1	**2**
	UDP-N-acetylglucosamine 4,6-dehydratase (inverting) (EC 4.2.1.115)	1	**4**
	Gamma-D-Glutamyl-meso-Diaminopimelate Amidase	1	**2**
	Capsular polysaccharide synthesis enzyme Cap8C	2	**4**
	CDP-glycerol: N-acetyl-beta-D-mannosaminyl-1,4-N-acetyl-D-glucosaminyldiphosphoundecaprenyl glycerophosphotransferase	1	**2**
	Monogalactosyldiacylglycerol synthase (EC 2.4.1.46)	1	**2**
	Duf71: Diphthamide synthesis Dph6	2	**2**
	Lipoate synthase, cyanobacterial paralog	2	**2**
	HMP-PP hydrolase (pyridoxal phosphatase) Cof, detected in genetic screen for thiamin metabolic genes (PMID:15292217)	3	**3**
	duf556 family protein	2	**4**
	Nitric oxide synthase oxygenase (EC 1.-.-.-)	2	**6**
	L-cystine uptake protein TcyP	1	**2**
	DNA sulfur modification protein DndE	1	**2**
	ATP-dependent helicase DinG/Rad3	2	**2**
	DNA topoisomerase III, TraE-type (EC 5.99.1.2)	1	**2**
	Outer spore coat protein E	2	**2**
	Protein GerPA, required for proper assembly of spore coat, mutations lead to super-dormant spore	1	**2**
	Stage II sporulation serine phosphatase for sigma-F activation (SpoIIE)	1	**2**
	RNA polymerase sporulation specific sigma factor SigK	1	**2**
	Stage II sporulation protein M (SpoIIM)	1	**2**
	Stage V sporulation protein AD (SpoVAD)	3	**4**
	CrtV-methyltransferase-like protein	1	**3**
	Fatty acyl-coenzyme A elongase	1	**2**
	Trans-2,3-enoyl-cozyme A reductase of elongase (EC 1.-.-.-)	1	**2**
	Acyl-CoA dehydrogenase (EC 1.3.8.1), Mycobacterial subgroup FadE4	1	**3**

	C-1'-hydroxylase CruF	2	**3**
	Triacylglycerol lipase (EC 3.1.1.3)	1	**3**
	Heme ABC transporter, cell surface heme and hemoprotein receptor HmuT	1	**3**
	Petrobactin ABC transporter, periplasmic binding protein	1	**2**
	Acyl-homoserine lactone acylase PvdQ (EC 3.5.1.-), quorum-quenching	2	**2**
	Hypothetical protein PvdX	1	**2**
	PvdE, pyoverdine ABC export system, fused ATPase and permease components	1	**6**
	Transmembrane component of energizing module of ECF transporters in Mycobacteria	1	**3**
	Transmembrane component STY3231 of energizing module of queuosine-regulated ECF transporter	1	**2**
	10 kDa culture filtrate antigen CFP-10 (EsxB)	1	**3**
	Lmo0066 homolog within ESAT-6 gene cluster, similarity to ADP-ribosylating toxins	2	**5**
	Membrane protein EccE1, component of Type VII secretion system ESX-1	1	**3**
	Putative riboflavin transporter RibN homologue	1	**2**
	1,2-dihydroxynaphthalene dioxygenase	2	**4**
	3,4-dihydroxyphenylacetate 2,3-dioxygenase	1	**4**
	Gallate permease	2	**3**
	3-oxoadipate enol-lactone hydrolase/4-carboxymuconolactone decarboxylase	1	**2**
	Flagella-related protein FlaI	1	**3**
	Protein-glutamine gamma-glutamyltransferase (EC 2.3.2.13)	1	**2**
	Flagellar biosynthesis protein FliT	1	**3**
	Flagellar hook subunit protein	1	**2**
	Transcriptional activator similar to cyn operon regulator	3	**3**
	Ribose-1,5-bisphosphate isomerase	1	**2**
	Guanine-hypoxanthine permease	2	**2**
	T4-like phage protein with ferrochetalase domain	1	**2**
	Integron integrase IntI2	1	**2**
	Chlorosome protein I, 2Fe-2S ferredoxin	2	**3**
	Ferredoxin-NADP(+) reductase (EC 1.18.1.2)	1	**2**
	Photosynthetic reaction center L subunit	1	**2**

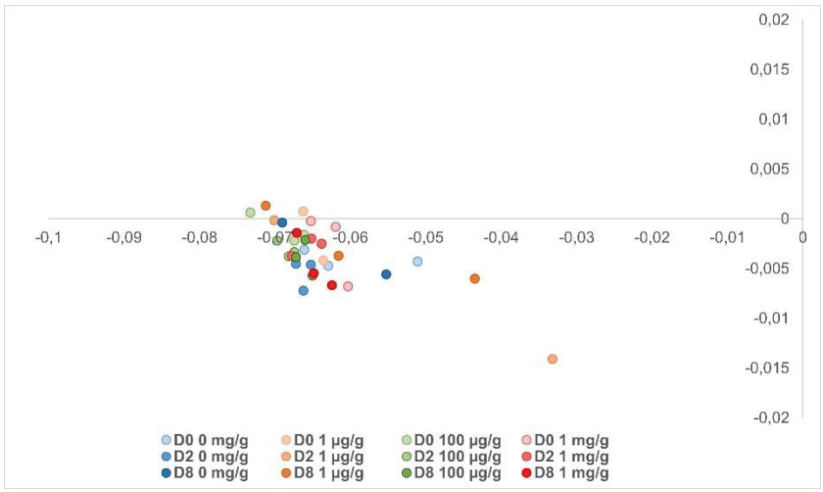
	FIG145533: Methyltransferase (EC 2.1.1.-)	2	**2**
	proteasome regulatory subunit Rpn12	1	**2**
	proteasome regulatory subunit Rpt1	2	**4**
	proteasome subunit alpha5 (EC 3.4.25.1)	2	**3**
	LSU ribosomal protein L40e	2	**4**
	SSU ribosomal protein S29e (S14p)	1	**2**
	SCIFF radical SAM maturase	2	**3**
	Eukaryotic translation initiation factor 2 beta subunit	3	**4**
	Autoinducer 2 (AI-2) ABC transport system, membrane channel protein LsrC	1	**3**
	LysR family transcriptional regulator PA2877	1	**2**
	VapC4 antibacterial toxin protein	1	**3**
	Putative analog of CcoH, COG3198	1	**3**
	Energy-conserving hydrogenase (ferredoxin), subunit A	1	**2**
	Hypothetical protein YrhD	1	**2**
	Archaeal transcription factor E	1	**2**
	RNA polymerase principal sigma factor HrdD	1	**5**
	NRPS module 2 PG-PG-Asn	1	**4**
	Phytochelatin synthase (EC 2.3.2.15), eukaryotic type	2	**3**
	Osmoprotectant ABC transporter inner membrane protein YehW	2	**4**
	Transcriptional regulator, Crp/Fnr family	1	**2**
	Malonate-semialdehyde dehydrogenase (EC 1.2.1.18)	1	**3**
	Mercuric transport protein, MerT	2	**2**
	Transmembrane transport protein MmpL1	1	**4**
DNeasy ®	threonine dehydrogenase	3	**3**
PowerSoil® Kit	NADH peroxidase (EC 1.11.1.1)	1	**2**
	Transcriptional repressor of the myo-inositol catabolic operon DeoR family	2	**2**
	Arabinose operon regulatory protein	2	**2**
	Predicted arabinose ABC transporter, permease protein 2	2	**2**
	L-rhamnose operon transcriptional activator RhaR	1	**2**

Rhamnulose-1-phosphate aldolase (EC 4.1.2.19)	1	**3**
NTD biosynthesis operon putative oxidoreductase NtdC (EC 1.-.-.-)	1	**2**
GDSL-like lipase/acylhydrolase	1	**2**
Predicted trehalose ABC transporter, permease component 1, Archaea-type	1	**3**
Methoxy mycolic acid synthase MmaA2 (EC 2.1.1.79)	1	**5**
Putative sugar nucleotidyltransferase	2	**2**
2-oxoglutarate carboxylase, small subunit (EC 6.4.1.7)	1	**2**
2-vinyl bacteriochlorophyllide hydratase BchF (EC 4.2.1.-)	1	**2**
Light-independent protochlorophyllide reductase iron-sulfur ATP-binding protein ChlL (EC 1.18.-.-)	1	**2**
Sepiapterin reductase (EC 1.1.1.153)	2	**4**
Erythronate-4-phosphate dehydrogenase (EC 1.1.1.290)	2	**3**
DNA recombination-dependent growth factor C	2	**4**
Spore germination protein GerHA/GerIA	1	**2**
Spore germination protein YpeB	1	**4**
Stage III sporulation protein AB	1	**2**
Stage III sporulation protein AE	2	**2**
(S)-3-O-geranylgeranylglyceryl phosphate synthase	1	**2**
Long-chain fatty-acid-CoA ligase (EC 6.2.1.3), Mycobacterial subgroup FadD36	1	**3**
Acyl-CoA dehydrogenase (EC 1.3.8.1), Mycobacterial subgroup FadE21	1	**3**
Hemoglobin-dependent two component system, sensory histidine kinase HrrS	2	**2**
Duplicated ATPase component of energizing module of riboflavin ECF transporter	2	**2**
Putative ESAT-secreted protein, BA2188 homolog	1	**2**
General secretion pathway protein C	1	**2**
Autoinducer synthesis protein LuxI	3	**4**
Hypothetical protein, slr1506/slr1944 homolog	2	**2**
GldB	1	**2**
Motility accessory factor	2	**4**
Flagellar transcriptional activator FlhC	2	**5**
Phage head-to-tail joining protein	1	**2**
Conjugative transposon protein TraF	1	**2**

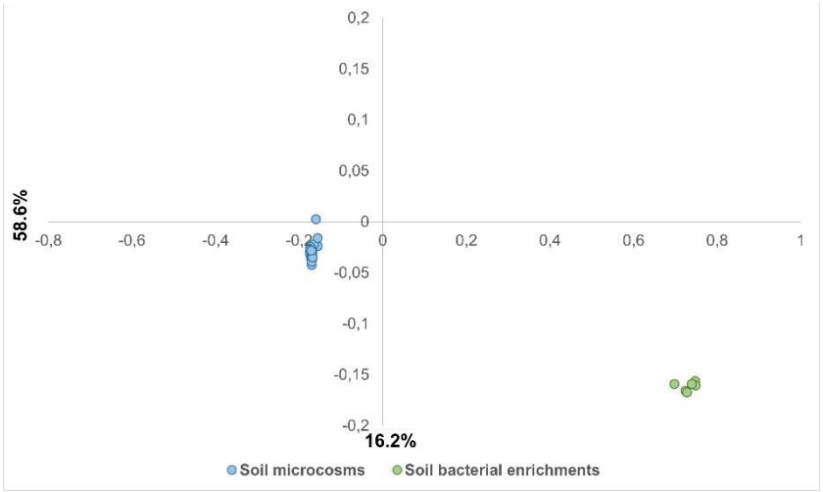
Possible potassium-efflux system protein	1	**3**
proteasome regulatory subunit Rpt3	1	**2**
Ribonuclease M5 (EC 3.1.26.8)	1	**2**
SSU ribosomal protein S27e	2	**4**
SSU ribosomal protein S35mt, mitochondrial	2	**2**
Holin-like protein CidA	1	**2**
Glycosyltransferase SypP	1	**2**
VapC11 antibacterial toxin protein	2	**2**
CO dehydrogenase/acetyl-CoA synthase, CO dehydrogenase subunit (EC 1.2.99.2)	1	**2**
Energy-conserving hydrogenase (ferredoxin), subunit E	2	**3**
Sodium-transporting N-ATPase C subunit	2	**2**
Na(+)-translocating NADH-quinone reductase subunit C (EC 1.6.5.-)	3	**4**
Na(+)-translocating NADH-quinone reductase subunit D (EC 1.6.5.-)	2	**4**
Anaerobic arsenite oxidase ArxA	2	**2**
Predicted transcriptional regulator of NADH dehydrogenase, Rrf2 family	2	**3**
Cyanobacteria-specific RpoD-like sigma factor, type-12	3	**3**
Inhibitor of vertebrate lysozyme precursor	1	**2**
IncF plasmid conjugative transfer protein TraD	2	**2**
IncQ plasmid conjugative transfer DNA nicking endonuclease TraR (pTi VirD2 homolog)	1	**2**
IncQ plasmid conjugative transfer protein TraG	2	**2**
Minor pilin of type IV secretion complex (VirB5)	2	**4**



(a)



(b)



(c)

Figure S1. (a) Bacterial community composition PCoA of: (a) soil microcosms and bacterial enrichments; (b) soil microcosms; (c) soil bacterial enrichments.

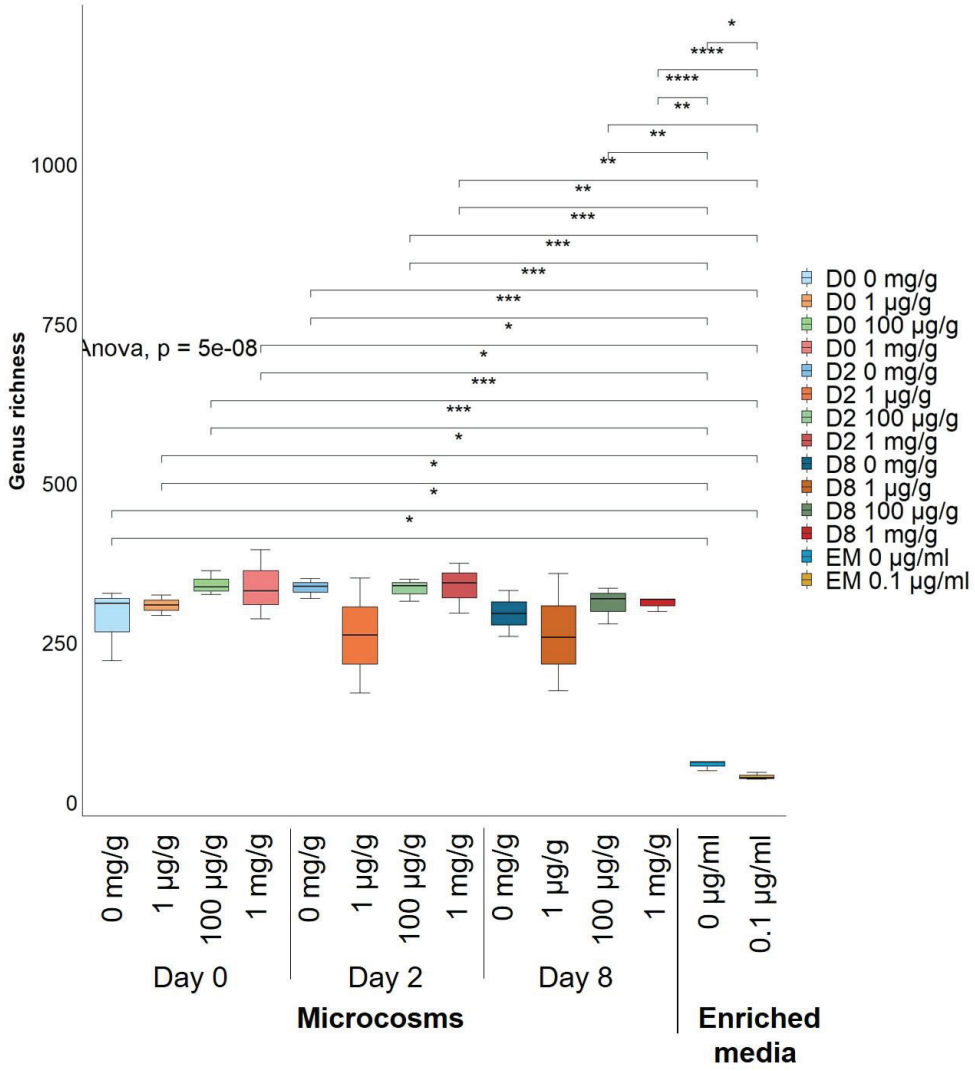


Figure S2. Bacterial richness measured in soil microcosms and enriched media polluted at different gentamicin concentrations. *p-value ≤ 0.05; **p-value ≤ 0.01; ***p-value ≤ 0.001; ***p-value ≤ 0.0001. n=3.

		Soil bacterial enrichments	Soil microcosms
Co-assembly of metagenomic reads (MEGAHIT)	Input	2908668 reads (9 samples)	9556334 reads (36 samples)
	Assembled contigs	4155	137
	Contig length	1000 – 27273 bp	1000 – 2977 bp
	Average length	2238 bp	1179 bp
	N50	2417 bp	1147 bp
Mapping (Bowtie2)	Maximum alignment rate	85.47%	0.08%
Binning	Bins wih >50% completion and <10% redundancy	2 (*Bacillus* and *Lysinibacillus)*	0
ARG screening on anvi'o profiles	Hits with >60%id and >33 aa	38 hits	0 hits

Table S1. Assembly and ARG screening on metagenomic reads obtained from soil bacterial enrichments (non-polluted or polluted with gentamicin at 100 ng/ml or 12 µg ml) and from soil microcosms (non-polluted or polluted with gentamicin at 1 µg/g, 100 µg/g or 1 ml/g).

	Day 0				Day 2				Day 8				Increased RA over time		
ARG	**0 µg/g**	**1 µg/g**	**100 µg/g**	**1 mg/g**	**0 µg/g**	**1 µg/g**	**100 µg/g**	**1 mg/g**	**0 µg/g**	**1 µg/g**	**100 µg/g**	**1 mg/g**	**1 µg/g**	**100 µg/g**	**1 mg/g**
ceoB	1,04E-05	0,00E+00	1,28E-05	7,53E-06	8,16E-07	6,48E-06	3,68E-06	7,84E-06	1,78E-06	1,83E-05	1,38E-05	1,30E-05			
acrD	0,00E+00	5,01E-06	0,00E+00	0,00E+00	2,75E-06	1,14E-06	2,16E-06	0,00E+00	0,00E+00	2,86E-06	8,79E-06	3,76E-06			
smeB	0,00E+00	2,96E-06	0,00E+00	2,80E-06	1,63E-06	0,00E+00	1,08E-06	0,00E+00	0,00E+00	2,86E-06	7,86E-06	1,44E-06			
amrB	3,91E-06	0,00E+00	0,00E+00	2,80E-06	0,00E+00	0,00E+00	1,52E-06	8,21E-07	0,00E+00	2,86E-06	4,19E-06	2,32E-06			
baeR	0,00E+00	0,00E+00	1,78E-06	1,40E-06	0,00E+00	1,14E-06	1,08E-06	0,00E+00	0,00E+00	2,86E-06	9,31E-07	2,32E-06			

Table S2a. Genes related to aminoglycoside resistance that increased their average relative abundance over time under gentamicin pollution at high concentrations (Pearson coefficient between time and gentamicin concentration ≥0.9). Obtained from blasting non-assembled reads against the CARD database, filtering at ≥60% identity and ≥33 aminoacids and selecting the best hit. Leave-One-Out cross-validation was applied to each gene on the table: no significant increase over time in gentamicin-polluted soils was detected for any of these genes. *ceoB* and *smeB:* multidrug efflux pumps that include aminoglycoside efflux. *acrD* and *amrB:* aminoglycoside efflux pumps. *baeR:* promoter of *acrD* expression. N=3.

ARG	Day 0				Day 2				Day 8				Increased RA over time		
	0 µg/g	1 µg/g	100 µg/g	1 mg/g	0 µg/g	1 µg/g	100 µg/g	1 mg/g	0 µg/g	1 µg/g	100 µg/g	1 mg/g	1 µg/g	100 µg/g	1 mg/g
muxB	3,25E-05	4,20E-05	2,50E-05	4,85E-05	2,07E-05	3,50E-05	3,27E-05	3,28E-05	2,02E-05	3,34E-05	4,05E-05	2,55E-05			
mexF	2,08E-05	4,50E-06	9,46E-06	1,96E-05	1,27E-05	1,18E-05	6,27E-06	1,25E-05	3,56E-06	2,01E-05	2,21E-05	1,96E-05			
msbA	2,08E-05	9,09E-06	1,04E-05	2,40E-05	4,38E-06	9,93E-06	9,46E-06	8,74E-06	2,05E-05	1,02E-05	2,19E-05	2,28E-05			
mexQ	3,91E-06	0,00E+00	9,36E-06	8,58E-06	6,43E-06	1,91E-06	5,40E-06	1,62E-06	0,00E+00	2,11E-05	1,35E-05	9,18E-06			
mdsB	8,53E-06	5,01E-06	0,00E+00	1,23E-05	1,63E-06	8,76E-06	2,60E-06	4,75E-06	6,92E-06	1,45E-05	1,90E-05	9,18E-06			
oqxB	3,91E-06	3,06E-06	1,78E-06	1,26E-05	5,32E-06	9,90E-06	5,83E-06	5,36E-06	5,13E-06	1,08E-05	1,09E-05	9,81E-06			
mexW	0,00E+00	6,03E-06	6,15E-06	1,40E-06	0,00E+00	9,16E-06	6,86E-06	4,69E-06	1,77E-05	1,23E-05	1,12E-05	3,10E-06			
acrB	0,00E+00	0,00E+00	0,00E+00	7,02E-06	1,93E-06	4,20E-06	1,08E-06	4,56E-06	0,00E+00	2,86E-06	8,59E-06	2,32E-06			
acrF	0,00E+00	0,00E+00	0,00E+00	6,13E-06	1,63E-06	4,20E-06	1,08E-06	1,60E-06	0,00E+00	4,30E-06	5,12E-06	1,55E-06			
mtrA	3,91E-06	0,00E+00	0,00E+00	3,33E-06	8,16E-07	3,42E-06	0,00E+00	1,62E-06	5,13E-06	5,12E-06	2,79E-06	6,19E-06			
rosA	5,21E-06	1,53E-06	1,78E-06	0,00E+00	0,00E+00	5,71E-06	0,00E+00	1,62E-06	0,00E+00	4,51E-06	2,79E-06	2,32E-06			

Table S2b. Genes related to multidrug efflux that increased their average relative abundance over time under gentamicin pollution at high concentrations (Pearson coefficient between time and gentamicin concentration ≥0.9). Obtained from blasting non-assembled reads against the CARD database, filtering at ≥60% identity and ≥33 aminoacids and selecting the best hit. Leave-One-Out cross-validation was applied to each gene on the table: no significant increase over time in gentamicin-polluted soils was detected for any of these genes. N=3.

ARG	Day 0				Day 2				Day 8				Increased RA over time		
	0 µg/g	1 µg/g	100 µg/g	1 mg/g	0 µg/g	1 µg/g	100 µg/g	1 mg/g	0 µg/g	1 µg/g	100 µg/g	1 mg/g	1 µg/g	100 µg/g	1 mg/g
tetB(60)	0,00E+00	0,00E+00	1,78E-06	1,26E-05	2,75E-06	3,05E-06	1,08E-06	3,11E-06	0,00E+00	8,59E-06	4,19E-06	8,15E-06			
tetA(58)	3,91E-06	5,93E-06	1,78E-06	1,40E-06	0,00E+00	6,48E-06	3,97E-06	4,58E-06	0,00E+00	7,21E-06	6,38E-06	6,08E-06		-	
otrC	5,21E-06	0,00E+00	0,00E+00	0,00E+00	8,16E-07	4,20E-06	1,08E-06	8,21E-07	0,00E+00	1,65E-06	1,40E-06	3,76E-06			
otr(A)	0,00E+00	0,00E+00	2,59E-06	5,61E-06	0,00E+00	1,91E-06	3,97E-06	8,21E-07	0,00E+00	3,69E-06	1,77E-06	2,18E-06			
adeB	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	1,91E-06	0,00E+00	0,00E+00	0,00E+00	2,86E-06	6,63E-06	0,00E+00			

Table S2c. Genes related to tetracycline efflux that increased their average relative abundance over time under gentamicin pollution at high concentrations (Pearson coefficient between time and gentamicin concentration ≥0.9). Obtained from blasting non-assembled reads against the CARD database, filtering at ≥60% identity and ≥33 aminoacids and selecting the best hit. Leave-One-Out cross-validation was applied to each gene on the table: no significant increase over time in gentamicin-polluted soils was detected for any of these genes. N=3.

ARG	Day 0				Day 2				Day 8				Increased RA over time		
	0 µg/g	1 µg/g	100 µg/g	1 mg/g	0 µg/g	1 µg/g	100 µg/g	1 mg/g	0 µg/g	1 µg/g	100 µg/g	1 mg/g	1 µg/g	100 µg/g	1 mg/g
Streptomyces	0,00E+00	2,28E-05	3,56E-06	7,18E-06	1,63E-06	6,48E-06	5,05E-06	7,16E-06	0,00E+00	1,11E-05	7,86E-06	7,52E-06			
rphA	3,91E-06	0,00E+00	4,27E-06	1,93E-06	2,75E-06	1,14E-06	5,05E-06	7,98E-06	5,13E-06	6,61E-06	6,38E-06	4,95E-06			
Bifidobacteria	0,00E+00	1,53E-06	5,33E-06	3,33E-06	8,16E-07	1,14E-06	0,00E+00	3,89E-06	0,00E+00	3,11E-06	7,31E-06	4,32E-06			
mupA	0,00E+00	0,00E+00	0,00E+00	1,93E-06	1,93E-06	1,14E-06	0,00E+00	8,02E-07	0,00E+00	2,89E-06	3,00E-06	3,76E-06			
bcrA	0,00E+00	5,01E-06	0,00E+00	0,00E+00	8,16E-07	1,14E-06	1,08E-06	8,02E-07	0,00E+00	2,86E-06	1,40E-06	2,18E-06			

Table S2d. Genes related to other resistance mechanisms that increased their average relative abundance over time under gentamicin pollution at high concentrations (Pearson coefficient between time and gentamicin concentration ≥0.9). Obtained from blasting non-assembled reads against the CARD database, filtering at ≥60% identity and ≥33 aminoacids and selecting the best hit. Leave-One-Out cross-validation was applied to each gene on the table: no significant increase over time in gentamicin-polluted soils was detected for any of these genes. *Streptomyces:* mutant conferring resistance to aminocoumarin. *rphA:* rifampicin enzymatic inactivation. *BIfidobacteria:* intrinsically resistant form of *ileS* (isoleucyl-tRNA synthetase) in *Bifidobacteria* that confers resistance to mupirocin. *mupA:* alternative *ileS* in *Staphylococcus* conferring resistance to mupirocin. *bcrA:* ABC transporter conferring resistance to bacitracin. N=3.

Gentamicin concentration	Bioavailable fraction
0 ng/g	**<QL**
1 µg/g	**<QL**
100 µg/g	**<QL**
1 mg/g	**<QL**

Table S3. Gentamicin concentrations in the available fraction of soils polluted at 1 µg/g, 100 µg/g and 1 mg/g. Standards were made by adding gentamicin to La Côte de Saint André soil and extracting the bioavailable fraction of gentamicin into water. Water was measured by HPLC-MS/MS to measure the concentration of gentamicin that is not adsorbed onto soil and estimate the concentration of gentamicin adsorbed onto soil. No trace of gentamicin was detected in water, suggesting that gentamicin was completely adsorbed onto soil. <QL= below quantification limit.

Chapter IV. Gentamicin pollution at sub-inhibitory concentrations induces a response in the environmental microbiome and resistome

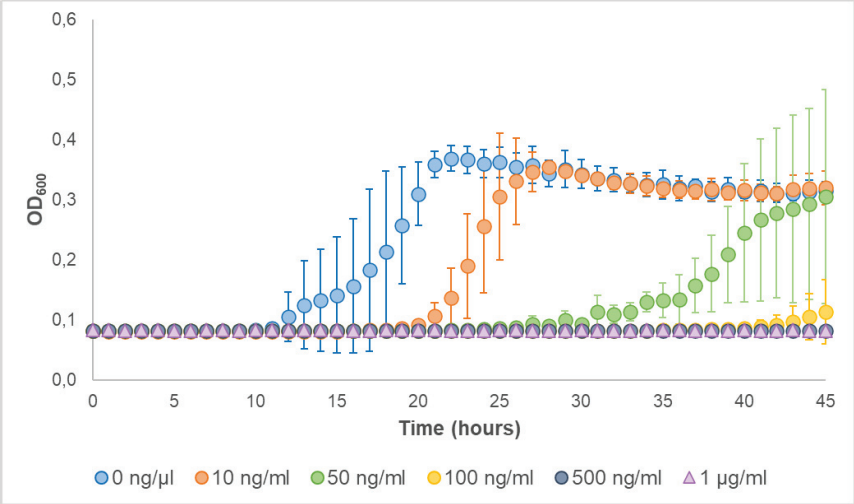


Figure S1. Average OD_{600} of culturable water bacteria incubated in R2A medium with different gentamicin concentrations for 24 hours. 150 µl of culture were transferred to a plate and incubated at 29°C for 45 hours, measuring the OD_{600} every hour. N=3.

[GM]\Time	0 hours	24 hours	48 hours
800 ng/ml	831.7 ± 85.7	703.2 ± 39.8	757 ± 68.8
50 ng/ml	52.54 ± 8.13	35.38 ± 3.95	32.13 ± 3.02
10 ng/ml	15.89 ± 5.26	12.80 ± 0.93	9.35 ± 1.96
0 ng/ml	<QL*	<QL	<QL

Table S1. Gentamicin concentrations in water microcosms measured by HPLC-MS/MS. QL: quantification limit. N=3.

Co-assembly of Illumina metagenomic reads (MEGAHIT)	Input	13,282,170 reads (36 samples)
	Assembled contigs	14983 contigs
	Contig length	1000 - 34750 bp
	Average length	1806 bp
	N50	1697 bp
Mapping (Bowtie2)	Alignment rate range	2.67% - 33.80%

Table S2. Assembly of Illumina metagenomic reads obtained from non-polluted water microcosms and microcosms polluted with gentamicin at 10, 50 or 800 ng/ml.

Co-assembly of Illumina and Nanopore metagenomic reads (Unicycler)	Short-read assembly (SPAdes)	Short reads: 16930 contigs = 13157648 (3 samples)
	Long read input	8000 reads (34,385,859 bp)
	Assembly of long reads to short-read contigs (miniasm)	Fully aligned reads: 392 Partially aligned reads: 1782 Unaligned reads: 5826 Total bases aligned: 9293553 bp Mean alignment identity: 81.1% Splits: 2195
Mapping (Bowtie2)	Alignment rate range	18.55% - 33.21%

Table S3. Hybrid assembly of Oxford Nanopore and Illumina metagenomic reads microcosms polluted with gentamicin at 50 ng/ml after 2-days exposure.

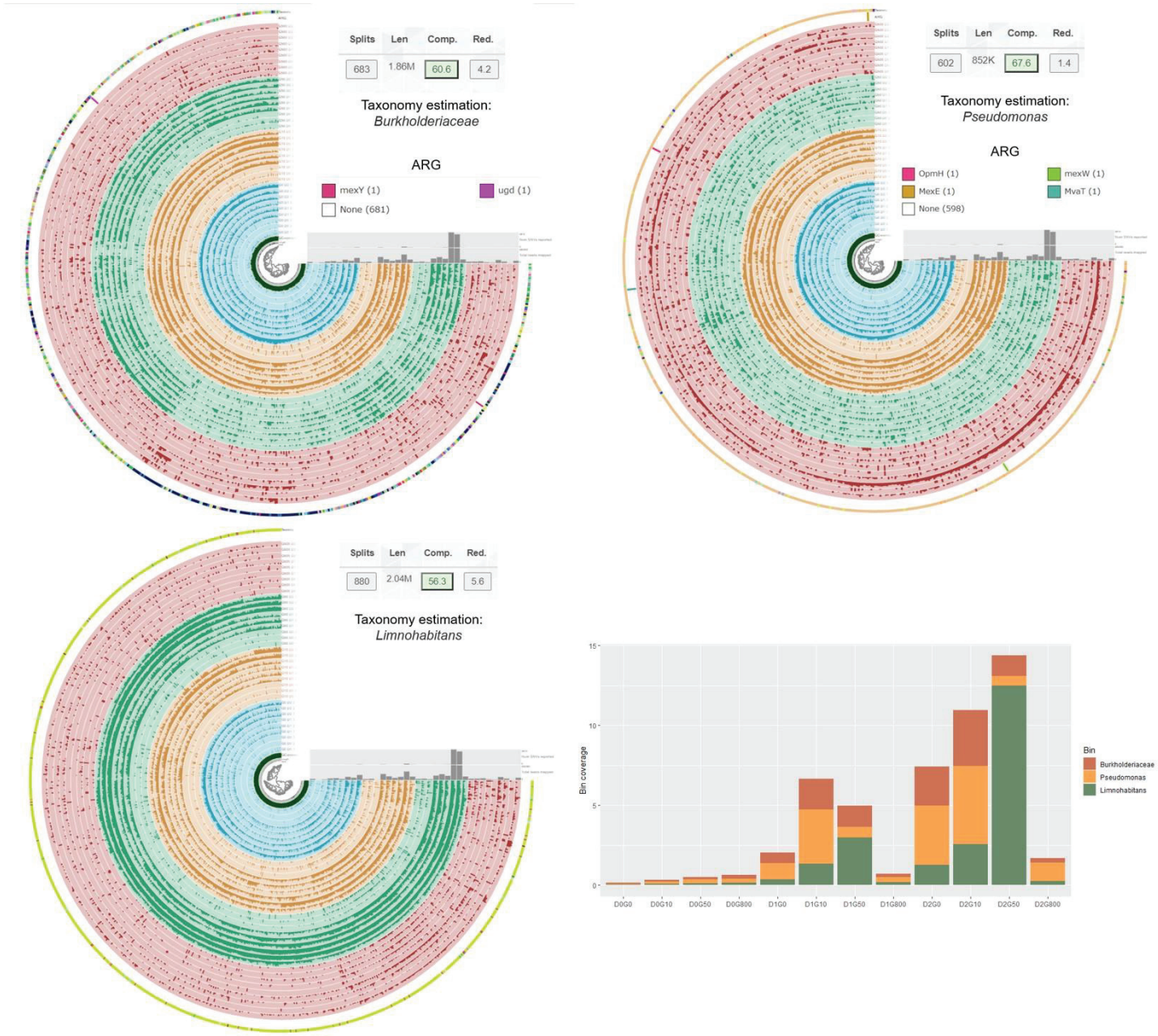


Figure S2. Bins with more than 50% completion and less than 10% redundancy co-assembled from Illumina metagenomic reads sequenced from water microcosms polluted at 0(blue), 10 (orange), 50 (green) and 800 (red) ng of gentamicin per ml of water. Bottom left: bin coverage in metagenomic sequences from water microcosms.

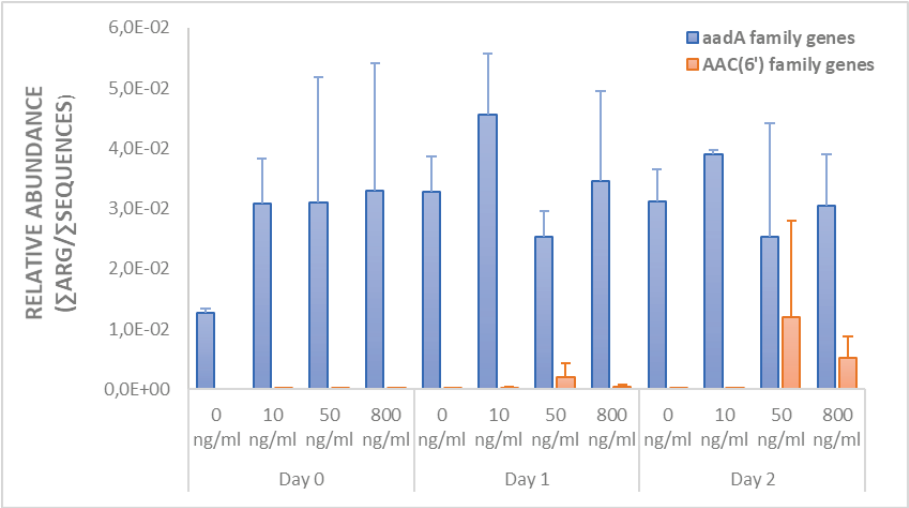


Figure S3. Relative abundance of the aminoglycoside resistance genes from the aadA family (blue) and the AAC(6') family (orange) in class 1 integrons. Chromosomal class 1 integrons were amplified using MRG284/MRG285 primers and sequenced. Aminoglycoside resistance genes were identified by blasting integron reads against the CARD database using Diamond and filtered at a minimum identity of 90%, a minimum length of 50 aminoacids and an e-value of 10^{-10}. The best hit was taken. Reads from genes belonging to the aadA aminoglycoside nucleotidyltransferase family or the AAC aminoglycoside acetyltransferase family, respectively, were grouped together. Then, their abundance was normalized by sequencing depth. N=3.

Primer pair	Sequence	Target genes	Amplicon size (bp)
***aac(6')* family genes**	F: 5'-CATGACCTTGCGATGCTCTATG-3'	AAC(6')-Ib10, AAC(6')-Ib11, AAC(6')-Ib3, AAC(6')-Ib4, AAC(6')-Ib7, AAC(6')-Ib8, AAC(6')-Ib9, AAC(6')-Ib', AAC(6')-Ib-cr, AAC(6')-Ib, ANT(3")-II/AAC(6')-IId, AAC(3)-Ib/AAC(6')-IIb", AAC(6')-30/AAC(6')-Ib', AAC(6')-Ib-Hangzhou, AAC(6')-Ib-Suzhou	201
	R: 5'-TCCAAGAGCAACGTACGACTG-3'		
***aadA* family genes**	F: 5'-GGCCTGAAGCCATACAGTGA-3'	aadA3, aadA8, aadA21, aadA15, aadA24, aadA12, aadA17, aadA25, aadA23, aadA, aadA11, aadA10, aadA6/aadA10, aadA6, aadA13, aadA22, aadA2	234
	R: 5'-AAGAATGTCCTTACGCTGCCA-3'		

Table S4. Details of qPCR primers targeting aminoglycoside resistance genes designed in this study.

		aac(6') family genes				*aadA* family genes			
		Gene copies/16S rRNA gene copies		Transcript copies/16S rRNA transcript copies		Gene copies/16S rRNA gene copies		Transcript copies/16S rRNA transcript copies	
		Average	%SD	Average	%SD	Average	%SD	Average	%SD
Day 0	0 ng/ml	3,6E-05	10	3,1E-07	52	2,3E-04	41	6,6E-07	173
	10 ng/ml	9,9E-06	104	1,5E-07	173	3,2E-04	38	1,2E-07	173
	50 ng/ml	4,5E-05	89	2,4E-08	141	6,3E-04	92	0,0E+00	NA
	800 ng/ml	3,8E-05	37	2,6E-07	173	2,6E-04	26	3,1E-07	102
Day 1	0 ng/ml	9,5E-05	132	2,6E-08	173	8,3E-04	42	2,3E-05	170
	10 ng/ml	5,0E-05	120	0,0E+00	NA	1,0E-03	33	1,9E-06	70
	50 ng/ml	2,2E-04	44	8,2E-07	102	6,9E-04	6	4,2E-08	173
	800 ng/ml	2,5E-04	109	8,8E-07	154	6,7E-04	52	0,0E+00	NA
Day 2	0 ng/ml	4,7E-05	27	7,7E-08	79	8,7E-04	34	9,8E-07	93
	10 ng/ml	5,5E-05	16	6,4E-08	92	1,4E-03	38	7,5E-07	22
	50 ng/ml	1,9E-03	75	4,8E-06	36	1,4E-03	65	1,1E-06	87
	800 ng/ml	7,6E-02	97	2,6E-04	75	8,0E-03	60	5,8E-06	95

Table S5. Average relative abundance and percentage of standard deviation of the relative abundance of gene copies and transcripts from the *aac(6')* and the *aadA* families. Total abundances of aminoglycoside resistance genes and transcripts were assessed by qPCR: 2 µl of DNA or cDNA were amplified using the GoTaq qPCR Master Mix and 341F/534R primers. Then, abundances were normalized by the number of copies of the 16S rRNA gene and transcript, respectively. AAC(6') family gene primers: qPCR efficiency=0.99, qPCR linearity=0.997. aadA family gene primers: qPCR efficiency=1, qPCR linearity=0.996. N=3.

Sample	Latitude	Longitude	Elevation (m)	Filtered volume (ml)	Sample	Latitude	Longitude	Elevation (m)	Filtered volume (ml)
NP01	50.78	15.59	1177	2000	S06	50.75	15.70	1190	1500
NP02	50.78	15.59	1161	2000	S07	50.75	15.70	1195	1500
NP03	50.78	15.59	1181	1500	S08	50.75	15.70	1202	1500
NP04	50.78	15.59	1187	2000	S09	50.75	15.70	1198	1400
NP05	50.78	15.59	1168	2000	S10	50.75	15.70	1209	1350
NP06	50.78	15.59	1173	2000	SF01	50.75	15.71	1290	500
NP07	50.78	15.59	1185	2000	SF02	50.75	15.71	1247	900
NP08	50.78	15.58	1209	2000	SF03	50.75	15.71	1250	500
NP09	50.78	15.59	1164	2000	SF04	50.75	15.70	1233	1500
NP10	50.78	15.59	1146	2000	SF05	50.75	15.70	1184	1500
NP11	50.78	15.59	1131	1500	SF06	50.75	15.70	1216	1500
NP12	50.78	15.59	817	2000	SF07	50.75	15.70	1184	500
NP13	50.78	15.59	812	2000	SF08	50.75	15.70	1179	500
NP14	50.78	15.59	811	2000	SF09	50.74	15.71	1410	1500
NP15	50.78	15.59	838	2000	SF10	50.74	15.70	1386	1500
NP16	50.79	15.59	807	2000	LB01	50.73	15.70	1413	1500
NP17	50.79	15.59	1092	2000	LB02	50.74	15.70	1416	1500
NP18	50.79	15.59	1097	2000	LB03	50.74	15.70	1424	1500
NP19	50.79	15.59	1093	1900	LB04	50.74	15.70	1424	1500
NP20	50.79	15.59	1095	1500	LB05	50.74	15.70	1428	1500
S01	50.75	15.71	1253	1500	LB06	50.74	15.70	1416	1500
S02	50.75	15.71	1248	1500	LB07	50.74	15.70	1425	1500
S03	50.75	15.71	1239	1500	LB08	50.74	15.71	1430	1500
S04	50.75	15.71	1233	1500	LB09	50.74	15.71	1410	1500
S05	50.75	15.70	1215	1500	LB10	50.74	15.71	1395	1500

Table S1. Latitude, longitude, elevation and filtered volume of each sample included in this study.

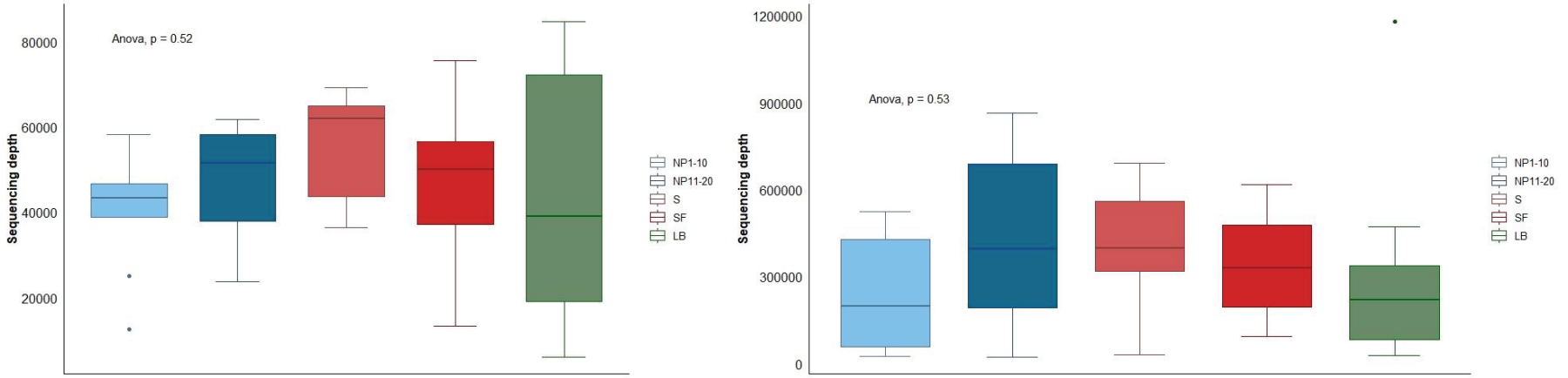


Figure S1. Depths obtained from the sequencing of the V3-V4 region of the 16S rRNA gene (left) and snow metagenomes (right). Results were submitted to an ANOVA test, and significant differences between sites were determined by pairwise t-tests. No significant differences were found between sites. n=10.