University Library



A gateway to Melbourne's research publications

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:
Hu, HW;Wang, JT;Singh, BK;Liu, YR;Chen, YL;Zhang, YJ;He, JZ

Title:
Diversity of herbaceous plants and bacterial communities regulates soil resistome across forest biomes

Date:
2018-09-01

Citation:
Hu, H. W., Wang, J. T., Singh, B. K., Liu, Y. R., Chen, Y. L., Zhang, Y. J. & He, J. Z. (2018). Diversity of herbaceous plants and bacterial communities regulates soil resistome across forest biomes. Environmental Microbiology, 20 (9), pp.3186-3200. https://doi.org/10.1111/1462-2920.14248.

Persistent Link:
https://hdl.handle.net/11343/283889

Author Manuscript

Title page

Diversity of herbaceous plants and bacterial communities regulates soil resistome across forest biomes

Hang-Wei Hu[1,2,6], Jun-Tao Wang[1,6], Brajesh K. Singh[3,4], Yu-Rong Liu[1], Yong-Liang Chen[5], Yu-Jing Zhang[2], Ji-Zheng He[1,2,*]

1. State Key Laboratory of Urban and Regional Ecology, Research Centre for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing 100085, China

2. Faculty of Veterinary and Agricultural Sciences, The University of Melbourne, Parkville, Victoria 3010, Australia

3. Hawkersbury Institute for the Environment, Western Sydney University, Penrith South DC, NSW 2751, Australia

4. Global Centre for Land-Based Innovation, Western Sydney University, Penrith South DC, NSW 2751, Australia

5. State Key Laboratory of Vegetation and Environmental Change, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China

6. These authors contributed equally to this work.

***Author for correspondence**

Ji-Zheng He, Research Centre for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing 100085, China. Email: jzhe@rcees.ac.cn

Running title: Plant communities regulate forest soil resistomes

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of record. Please cite this article as doi:10.1111/1462-2920.14248.

This article is protected by copyright. All rights reserved.

Author Manuscript

Originality-Significance statement

This is the first large-scale profiling of a broad spectrum of antibiotic resistance genes (ARGs) in natural forest biomes, providing important information on how the interactions of aboveground plants and belowground microbes as well as soil properties and climate shape the distribution patterns of antibiotic resistance. Our results suggest that natural forest soils are a significant source of ARGs, and their large-scale distribution is regulated by the diversity of bacteria and herbaceous plants even when simultaneously accounting for multiple environmental drivers. This knowledge is critical in understanding the origins and evolution of natural soil resistome, and may help to design effective framework to predict and manage the spread of environmental ARGs.

Summary

Antibiotic resistance is ancient and prevalent in natural ecosystems and evolved long before the utilization of synthetic antibiotics started, but factors influencing the large-scale distribution patterns of natural antibiotic resistance genes (ARGs) remain largely unknown. Here, a large-scale investigation over 4000 km was performed to profile soil ARGs, plant communities, and bacterial communities from 300 quadrats across five forest biomes with minimal human impact. We detected diverse and abundant ARGs in forests, including over 160 genes conferring resistance to eight major categories of antibiotics. The diversity of ARGs was strongly and positively correlated with the diversity of bacteria, herbaceous plants, and mobile genetic elements (MGEs). The ARG composition was strongly correlated with the taxonomic structure of bacteria and herbs. Consistent with this strong correlation, structural equation modelling demonstrated that the positive effects of bacterial and herb communities on ARG patterns were maintained even when simultaneously accounting for multiple drivers (climate, spatial predictors and edaphic factors). These findings suggest a paradigm that the interactions between aboveground and belowground communities shape the large-scale distribution of soil resistomes, providing new knowledge for tackling the emerging environmental antibiotic resistance.

This article is protected by copyright. All rights reserved.

Author Manuscript

Introduction

Emerging prevalence and dissemination of antibiotic resistance genes (ARGs, collectively known as "resistome") represents a major threat to public health in the 21st century (WHO, 2014; Berendonk *et al.*, 2015). A growing recognition of an environmental component of resistome pools has prompted calls for investigation into the significance of antibiotic resistance in natural settings (Ashbolt *et al.*, 2013). It has been reported that some environmental ARGs have perfect sequence identity to those in human pathogens (Forsberg *et al.*, 2012) and human feces (Nesme *et al.*, 2014), indicative of the possible gene flow between environmental and clinical resistomes (Wright, 2010). Environmental ARGs can be disseminated via physical forces such as leaching to water, atmospheric transport/deposition of dust, dissipation by animals or birds, and delivery of ARGs-carrying products (Allen *et al.*, 2010; McEachran *et al.*, 2015). They are also subject to horizontal gene transfer mediated by mobile genetic elements (MGEs) such as integrons, transposons, and plasmids (Gillings *et al.*, 2015). Therefore, environmental ARGs have been recognized as contaminants of global concern (Pruden *et al.*, 2006), and may constitute an important contributor to clinical antibiotic resistance if acquired by human commensals and pathogens.

Soil is one of the largest environmental reservoir comprising approximately 30% of known ARGs in public repositories (Forsberg *et al.*, 2012; Nesme *et al.*, 2014), and is one of the most complex ecosystems in terms of ecological niches and biodiversity (Nesme and Simonet, 2014). To date, a large body of studies have documented the prevalence of ARGs in environments under anthropogenic perturbations, such as soils with manure amendments (Zhu *et al.*, 2013; Zhang *et al.*, 2017), contaminated with pharmaceutical residues (Zhu *et al.*, 2017) or heavy metals (Hu *et al.*, 2016b, 2017), and irrigated with reclaimed water (Wang *et al.*, 2014; Han *et al.*, 2016). We have, however, limited knowledge of the background levels of antibiotic resistance and the distribution of ARGs in natural settings with little or no human disturbance. In fact, most antibiotics are natural compounds produced by different taxa of bacteria, fungi or plants (D'Costa *et al.*, 2006; Chen *et al.*, 2013), and the presence of natural antibiotics may have exerted consistent selection pressure on microorganisms since long before the introduction of human-made antibiotics (Allen *et al.*, 2010; Wright and Poinar, 2012). Therefore, it is not surprising that ARGs were identified from ancient permafrost sediment (D'Costa *et al.*, 2011), Alaskan permafrost soil (Allen *et al.*, 2009), glaciers (Segawa *et al.*, 2013), Antarctic soils (Wang *et al.*, 2016a), underground cave (Pawlowski *et al.*, 2016), and deep ocean sediment (Chen *et al.*, 2013). However, we lack a general

Wiley-Blackwell and Society for Applied Microbiology
This article is protected by copyright. All rights reserved.

Author Manuscript

understanding of the key determinants shaping the diversity and composition of ARGs in natural ecosystems. These knowledge gaps have precluded the inclusion of natural resistomes in management strategies and risk assessment frameworks (Ashbolt *et al.*, 2013).

Forest ecosystems occupy approximately 30% of Earth's land surface (Bonan, 2008) representing important habitats for trees, herbs, and microbes (Baldrian, 2017). Plants can influence the structure of soil microbial communities through plant litters and root exudation (Prober *et al.*, 2015; Wang *et al.*, 2016b), which might indirectly shape soil resistome composition via affecting the taxonomic and phylogenetic structure of ARGs-bearing microbes (Forsberg *et al.*, 2014). On the other hand, forest soils contain antibiotics produced by soil microbes, and potentially toxic aromatic compounds derived from degradation processes, as well as specialized antibiotics (e.g., salicylic acid, phytoalexins and flavonoids) and signalling molecules from plant exudates (Alonso *et al.*, 2001; Yergeau *et al.*, 2014), which are reported to increase expression of antibiotic resistance in the plant rhizosphere (Yergeau *et al.*, 2014). In addition, forest soils are characterized by nutrient scarcity, and represent a highly competitive environment for microbes to survive, interact, and compete for space and nutrients, hence it is expected to observe a wide occurrence of ARGs-bearing microbes (producing low concentrations of antibiotics as signalling molecules to modulate the interactions within microbial communities; or producing high concentrations of antibiotics as a weapon to counter other microbial taxa probably in some micro-environments) in forest soils (Martinez, 2008). Despite forest soils being hypothesized as an important reservoir of ARGs, the large-scale distribution patterns of ARGs and how their compositions are regulated by environmental factors and plant and bacterial communities remain less understood.

In this study, we used high-throughput quantitative PCR (HT-qPCR) techniques to profile 285 ARG subtypes encoding resistance to eight major classes of antibiotics from 300 soil samples collected across five forest biomes in China. This large-scale investigation through establishing 300 quadrats over 4000 km from boreal forest to tropic rainforest allows simultaneous documentation of plant communities (including trees and herbs), soil microorganisms, soil resistome, and a wide range of climatic variables and soil properties. We hypothesized that (i) the natural forest ecosystem is a significant reservoir of ARGs, due to the antimicrobial molecules produced by the soil microorganism and plant rhizosphere; (ii) the changes in ARG patterns are closely associated with the changes in the bacterial communities, which is assumed as an important determinant of ARGs in natural settings; and (iii) the impacts of plant communities on the diversity and structure of soil microbial

Wiley-Blackwell and Society for Applied Microbiology
This article is protected by copyright. All rights reserved.

communities will ultimately have consequences on the soil ARG profiles. Our study provides important information on the background level of ARGs and the key drivers shaping the large-scale biogeographic distribution of antibiotic resistance in natural forest biomes. This knowledge is critical in understanding the origins and evolution of natural soil resistome, and may help to design effective framework to predict and manage the spread of environmental ARGs.

Results

Diversity and relative abundance of ARGs and MGEs

The HT-qPCR array detected a total of 160 ARGs and 9 MGEs (including 2 integrase genes and 7 transposase genes) in the 300 soil samples collected from five forest biomes (Fig. 1a). The average number of detected ARGs considerably changed from 95.5 (± 20.1) in temperate mixed coniferous forest to 73.8 (± 16.1) in subtropical forest (Fig. 1b). The diversity of ARGs in temperate mixed coniferous forest and temperate deciduous forest (94.2 ± 20.7) was significantly higher than that in boreal forest (85.7 ± 18.2) and subtropical forest ($P < 0.05$). The numbers of detected MGEs showed a similar pattern to that of ARGs (Fig. 1c), with the lowest found in subtropical forest (3.3 ± 0.8) and the highest in temperate mixed coniferous forest (4.2 ± 1.3). The diversity of MGEs in temperate mixed coniferous forest and temperate deciduous forest (4.1 ± 0.8) was significantly higher than that in subtropical forest ($P < 0.05$). The detected ARGs could potentially confer resistance to eight major classes of antibiotics, with the most frequently detected ARGs being multidrug and β-lactam resistance (Fig. S1a), accounting for 22.4% and 21.8%, respectively, of the total ARGs. Other frequently encountered ARG types included aminoglycoside (13.5%), MLSB (14.1%), tetracycline (11.2%), and vancomycin (8.2%) resistance, depending on the representation of primer sets on the array. The detected ARGs included all major resistance mechanisms: antibiotic inactivation (42.9%), multidrug efflux pumps (33.5%), and cellular protection (20.0%) (Fig. S1b). Venn diagram showed that a total of 59 ARGs were shared among the five forest ecosystems (Fig. S2a), and these commonly shared ARGs were dominated by the multidrug and β-lactam resistance genes (Fig. S2b).

We further calculated the relative abundance of ARGs and MGEs by normalizing against the bacterial 16S rRNA gene. Similar to the pattern of the ARG diversity, the relative abundance of ARGs highly varied across forest biomes, with the most abundant ARGs observed in temperate mixed coniferous forest, and the lowest in subtropical forest ($P < 0.05$)

Wiley-Blackwell and Society for Applied Microbiology
This article is protected by copyright. All rights reserved.

Author Manuscript

Author Manuscript

(Fig. 1d). The relative abundance of ARGs was significantly higher in boreal and temperate (mixed coniferous and deciduous) forests than subtropical and tropical forests ($P < 0.05$). The relative abundance of MGEs significantly shifted across forest biomes (Fig. 1e), and was significantly lower in subtropical and tropical forest than in other three forest types ($P < 0.05$). Spearman's correlation analysis revealed significantly positive relationships between the relative abundances of total MGEs with total ARGs and individual ARG types (Table S3). The relative abundances of specific MGEs, such as *intI*, *intI1*, and *tnpA-02* genes were also significantly and positively correlated with those of total ARGs and individual ARG types.

Gene-specific analyses of ARGs and MGEs subtypes

The patterns of the relative abundance of ARG and MGE subtypes across forest biomes were explored in heat maps based on the HT-qPCR array results (Fig. 2a). Soil samples (columns in the heat map) could be separated into several clusters: samples from subtropical and tropical forest tended to cluster together; and samples from the other three forest biomes were distributed into three separate major clusters with each containing samples from all three forest biomes. This finding was further corroborated by non-metric multidimensional scaling (NMDS) ordinations based on the Bray-Curtis distances of ARGs (Fig. 2b), which demonstrated that subtropical and tropical forests harboured significantly different ARG compositions from temperate and boreal forests (PerMANOVA, $P < 0.05$). The profiles of individual ARG subtypes (rows in the heat map) could be classified into three distinct patterns: (A) These ARG subtypes are prevalent with high relative abundance in the majority of forest soils and tended to be less abundant in subtropical and tropical forest; (B) Compared to pattern A, these ARGs are less abundant in forest soils, and tend to be decreased in subtropical and tropical forests; (C) These ARG subtypes are in low abundance or below the detection limit in most forest soils.

Among the 160 ARGs and 9 MGEs detected, 32 ARGs and 2 MGEs were found in more than 98% of the forest soil samples (Table S4). These ubiquitous genes contained 1 aminoglycoside resistance gene, 9 β-lactam resistance genes, 4 macrolide-lincosamide-streptogramin B (MLSB) resistance genes, 5 tetracycline resistance genes, 5 vancomycin resistance genes, and 8 multidrug resistance genes. Some widely-distributed ARG and MGE subtypes included *intl1*, *aacC*, *mexF*, *blaSFO*, *cphA-01*, *fox5*, *mphA-01*, *oleC*, *tetR-02*, *vanC-03*, *acrA-05*, *oprD*, and *oprl* genes. The *mexF* gene (multidrug resistance) was the most abundant resistance gene, up to 20.2% as abundant as the 16S rRNA gene. The *intl1* gene was the most abundant MGE, ranging from 0.02-14.4%.

Wiley-Blackwell and Society for Applied Microbiology

This article is protected by copyright. All rights reserved.

Author Manuscript

Soil bacterial, herb and tree communities

Bacterial communities were determined by targeting the 16S rRNA gene on the Miseq platform, which yielded a total of 15,421,630 high-quality sequences from the 300 soils, with an average of 51,405 sequences per sample. These sequences could be clustered to 40,399 OTUs at 97% sequence identity, with the number of OTUs per sample ranging from 918 to 7110. Bacterial communities were dominated by Proteobacteria, Acidobacteria, Actionobacteria, Verrucomicrobia, Planctomycetes, and Chloroflexi at the phylum level (Fig. S3a).

Herb communities were comprised of three major phyla Magnoliophyta, Pteridophyta, and Lycopodiophyta, and their relative abundance highly varied across forest biomes, with Pteridophyta predominant in subtropical forest and Lycopodiophyta found only in tropical forest (Fig. S3b). Tree communities showed substantial differences across forest biomes (Fig. S3c). For example, the Ericales and Fabales families belonging to the Angiospermae phylum were only found in subtropical and tropical forests, while the Rosales family was more dominant in boreal and temperate mixed coniferous forests.

Soil bacterial diversity was calculated for each sample using the numbers of OTUs at 97% sequence similarity. The changes in the diversity of bacteria and herbs across forest biomes were very similar, with the lowest level of diversity recorded in subtropical forest (Figs. S3d and S3e), which mimicked the diversity patterns of ARGs (Fig. 1b). Tree diversity significantly increased from boreal forest (2.08 ± 1.07) to tropical forest (34.0 ± 12.2) (Fig. S3f).

Relationships between ARGs with MGEs, plants, bacteria, and environmental factors

Regression analysis was conducted to explore the relationships between the diversity of ARGs with the diversity of MGEs, bacteria and plants, soil properties and climatic factors. The ARG diversity had strongly positive linear relationships with the diversity of bacteria (Fig. 3a), herbs (Fig. 3b), and MGEs (Fig. 3c) (all $P < 0.001$). Soil pH was the most important edaphic factor that was significantly correlated with the ARG diversity, with the highest diversity predicted around pH 7 (Fig. 3d). The relationships between the ARG diversity with tree diversity, spatial factors, soil properties, and climatic factors were not significant, and only the correlations with annual mean precipitation (AMP) and available potassium were statistically significant (Fig. S4).

Wiley-Blackwell and Society for Applied Microbiology

This article is protected by copyright. All rights reserved.

Author Manuscript

Mantel tests were further performed to examine the correlations between the ARG composition with the compositions of MGEs, bacteria and plants as well as environmental factors (Table S5). Similarly, we found that MGEs, bacteria, herbs, and soil pH were the most important factors that had significantly positive correlations with the ARG composition. Soil properties including soil organic carbon (SOC), total nitrogen (TN), and moisture content, climatic factors including annual mean temperature (AMT) and AMP, and spatial factor including latitude were also drivers that may significantly influence the patterns of ARGs in forests. This result was further supported by the CCA test which revealed that soil pH, AMT, AMP, latitude, and moisture content were important drivers shaping the ARG patterns (Fig. S5).

Co-occurrence patterns of ARGs, MGEs, bacteria, and herbs

The links between ARGs, MGEs, and bacterial and herb communities were further explored using the local similarity analyses, which were visualized as association networks based on strong ($\rho > 0.8$) and significant ($P < 0.05$) correlations. Significant co-occurring correlations were observed in the ARGs-MGEs-bacteria-herb networks, suggesting their strong interactions in forests (Fig. 4a). Interestingly, the *intI1* and *intI* genes (belonging to MGEs) had the most intensive connections with ARGs (Fig. 4a), which can potentially confer resistance to multiple classes of antibiotics (Fig. 4b). The co-occurrence pattern between ARGs and MGEs was further assessed in a network including only ARG/MGE subtypes (Fig. S6). The entire ARGs-MGEs network was clearly separated into four clusters, with the Module III composed of the *intI1* and *intI* genes that had positive relationships with all the major classes of ARGs (Fig. S6). The Gleicheniaceae, Poaceae, Zingberaceae, Cyperaceae, Apiaceae, and Woodsiaceae families were the key herb taxa that had significant correlations with bacteria and ARGs (Fig. 4).

The *blaSFO* (β-lactam resistance), *cmlA1-02* (FCA resistance), *oleC* (MLSB resistance), and *emrD* (sulfonamide resistance) genes were highly connected with a variety of bacterial taxa (Fig. 4a), which were mainly affiliated with Actinobacteria, Proteobacteria, Chloroflexi, Acidobacteria (Fig. S7), implying that these bacterial phyla are possible ARG hosts. Single bacterial taxa significantly correlated with multiple ARG subtypes, suggestive of a bacterial host to harbour multiple ARGs (Fig. S7). Procrustes analysis further demonstrated significant correlations between the bacterial community compositions and ARG contents with highly significant goodness-of-fit measures (Fig. 2c).

Wiley-Blackwell and Society for Applied Microbiology

This article is protected by copyright. All rights reserved.

Author Manuscript

Structural equation modelling accounting for multiple drivers

We used Random Forest modelling to identify the major factors predicting the ARG diversity, and structural equation model (SEM) to test whether the relationships between the diversity of plants, bacteria and ARGs is maintained when accounting for multiple drivers simultaneously. Random Forest analysis indicated that microbial diversity was more important than other environmental factors in predicting the ARG patterns (Fig. S8). Plant diversity (including tree and herb diversity) was as important as most environmental drivers. Based on the known relationships among the drivers of ARG patterns, we constructed an *a priori* model (Fig. S9) to determine the direct and indirect effects of soil properties (including soil pH), climatic conditions (including AMT), biome types, the diversity of bacteria, trees and herbs, and MGEs, on ARG patterns. Our final SEM explained 67% of the variance in the ARG diversity across forest biomes (Fig. 5a). Consistent with the correlation analyses (Fig. 3), microbial diversity and herb diversity showed positive effects on the ARG diversity even after accounting for the simultaneous direct and indirect effects of multiple variables (Fig. 5b). No significant and direct effects of plant diversity on ARG patterns were found, but herb diversity had indirect effects on ARG patterns through its positive impacts on microbial diversity (Fig. 5a). Microbial diversity exerted direct positive effects on ARG patterns, and also indirectly affected ARGs through its positive effects on MGEs (Fig. 5a). Not surprisingly, biome types, climate, spatial factors and edaphic factors showed strong direct and indirect effects on ARGs as indicated by the standardized total effects from SEM (Fig. 5b), assuming that these factors have been identified to be critical in shaping the distribution pattern of soil microbes. Collectively, these results demonstrate that microbial and herb diversity play critical roles in shaping the ARG patterns, which are also directly and indirectly affected by other drivers in forests.

Discussion

Ubiquitous ARGs in forest biomes

A key finding of this study is that more than 160 ARGs (encoding resistance to eight major classes of antibiotics) were detected from the 300 forest soils with minimal anthropogenic disturbance. In line with previous screening of ARGs in polluted estuaries (Zhu *et al.*, 2017), urban environments (Xiang *et al.*, 2018), agricultural soils (Hu *et al.*, 2016a, 2016b; Chen *et al.*, 2017) and Chinese swine farms (Zhu *et al.*, 2013), genes conferring resistance to multidrug, β-lactam, aminoglycoside, MLSB, tetracycline, and

Wiley-Blackwell and Society for Applied Microbiology
This article is protected by copyright. All rights reserved.

vancomycin were the most dominant ARG types in these forest soils. Our findings contribute to the growing body of evidence that ARGs are widespread in relatively pristine habitats with less contact with commercial sources of antibiotics (Chen *et al.*, 2013; Segawa *et al.*, 2013; Pawlowski *et al.*, 2016), and highlight that the natural forest ecosystem is a significant reservoir of ARGs and needs to be considered in the risk assessment frameworks of antibiotic resistance. Comparison of the background ARG profiles between the forest ecosystems and some human-impacted soils by using the same HT-qPCR array suggested that the diversity and relative abundance of ARGs in copper-contaminated soils (Hu *et al.*, 2016b), nickel-contaminated soils (Hu *et al.*, 2017), animal manure-amended soils (Zhang *et al.*, 2017), and antibiotics-treated soils (Zhang *et al.*, 2017) are significantly higher than those in forest biomes (Fig. S10). Given that forests are significantly threatened by new anthropogenic pressures such as urbanization, intensive management practices, and land-use conversion to farmland, it is assumed that these pre-existing ARGs might emerge with clinical importance in future if they are transferred to a human pathogen (Martinez *et al.*, 2015).

The widespread occurrence of ARGs in natural forests might be attributed to the inherent presence of antibiotics in soil (Wright, 2010; Nesme *et al.*, 2014), though it is very difficult to measure the natural concentrations of antibiotics (Allen *et al.*, 2010). Most of currently known antibiotics are small molecules produced by certain microorganisms in natural settings (D'Costa *et al.*, 2006). Bacteria can use these antibiotics as weapons (at high concentrations) to compete for limited resources (D'Costa *et al.*, 2006) or as signalling molecules (at low concentrations) to communicate with other microbes (Martinez, 2008). Furthermore, forest soil contains some toxic aromatic compounds derived from degradation processes or specialized antibiotics from plant exudates (Alonso *et al.*, 2001; Yergeau *et al.*, 2014), which may also contribute to the development of antibiotic resistance. Apart from the long-term residence of indigenous ARGs, other unknown selection pressures (i.e. anthropogenic disturbance) can also contribute to the level of ARGs in forests. Although the 300 soil samples in this study were collected in regions that are distant from human industrialized activities and supposedly devoid of anthropogenic disturbance, we cannot completely exclude the possible mechanism of atmospheric deposition of ARGs-bearing microbes that originate from ARGs-rich environments, and the long-distance transport of ARGs by birds and animals (Wang *et al.*, 2016a).

Some ARG subtypes were ubiquitously detected across the five forest ecosystems, such as *aacC*, *mexF*, *blaSFO*, *cphA*, *fox5*, *mphA-01*, *oleC*, *tetR-02*, *vanC-03*, *acrA-05*, *oprD*, and

Wiley-Blackwell and Society for Applied Microbiology
This article is protected by copyright. All rights reserved.

Author Manuscript

Author Manuscript

oprl genes, indicating that natural forest soils are a common reservoir for these genes. The *mexF* gene was the most abundant resistance gene, perhaps driven by multiple antibiotic and toxic compounds in forests. The *mexF* gene has been reported in various environments, such as soils, sediments, rivers and drinking water (De Souza *et al.*, 2006). The *cphA* gene (β-lactam resistance) has been found to be intrinsic in environmental isolates of *A. jandaei* and *Aeromonas hydrophila* (Balsalobre *et al.*, 2009). The *mexF* and *oprD* multidrug resistance genes were proposed to naturally occur in water (Stedtfeld *et al.*, 2016). The *vanC* gene can encode resistance to vancomycin, which is considered as one of the last-resort defence against organisms like *Streptococcus pneumonia* and *Enterococcus* (Gilmore and Hoch, 1999), and this gene is ancient in natural environments (D'Costa *et al.*, 2011). These findings suggest that natural forests are sharing many ARGs within and across biomes, with a possibility of gene flow and exchange among different ecosystems. Therefore, we argue that increasing attention should be paid to the importance of natural resistome and its potential transmission routes to human commensals and pathogens.

Although the HT-qPCR array used in this study can simultaneously profile 285 ARGs and 10 MGEs of interest in a high-throughput manner (Looft *et al.*, 2012; Su *et al.*, 2015), the need for pre-selected primer sets limits its use to known genes or to genes with high nucleotide similarity to known ones (Karkman *et al.*, 2018). For example, the relative dominance of the ARG classes or resistance mechanisms detected using the HT-qPCR technique mostly relies on the representation of primer sets on the array. Metagenomics, especially deep sequencing, is not restricted to prior knowledge of few selected genes, and is able to capture the whole resistome in the environment (Li *et al.*, 2015). However, the annotation of ARGs is still dependent on sets of known verified genes in ARG databases (Gibson *et al.*, 2015). Functional metagenomics can overcome the limitations of HT-qPCR and metagenomic sequencing in detecting mostly known ARGs, but the difficulty in the cloning and expression of environmental DNA in a laboratory host is the main disadvantage of this technique (Karkman *et al.*, 2018). Despite the intrinsic limitations of these methodologies, a combination of HT-qPCR, metagenomics and functional metagenomics is still powerful in the determination of antibiotic resistance in all its dimensions.

Potential for horizontal gene transfer of ARGs in forest soils

The abundant and diverse ARGs detected in forest soils, together with abundant MGEs, offer a likelihood for ARGs being horizontally transferred to other bacteria and pathogens. Firstly, the class 1 integron-integrase gene *intI1* and the transposon-transposase gene *tnpA*

Wiley-Blackwell and Society for Applied Microbiology
This article is protected by copyright. All rights reserved.

Author Manuscript

were detected in every forest soil sample. Class 1 integrons are widely distributed in over 70% of Gram-negative bacteria of clinical importance (Stalder *et al.*, 2012), and have been repeatedly reported to be an important proxy for acquisition and dissemination of ARGs in the environment (Gillings *et al.*, 2015). Integrons have a site recombination system which could capture and express gene cassettes, and they are often located on plasmids and transposons which further facilitate their mobility potential (Heuer *et al.*, 2012). Secondly, the HT-qPCR detected two integrase genes and seven transposase genes, with their total diversity and relative abundances positively correlated with those of ARGs. In particular, the relative abundances of *intI*, *intI1* and *tnpA-02* genes had significantly positive relationships with the relative abundances of all individual classes of ARGs (Table S3), and the compositions of ARGs and MGEs were also closely correlated as revealed by the Mantel test. These results suggested that there is a potential risk of HGT mediated by MGEs for ARGs in forest soils. However, it was found that the diversity of both ARGs and MGEs was significantly correlated with the diversity of bacteria, which means that the significantly positive relationships observed between ARGs and MGEs might be indirect and would not necessarily lead to high HGT potential.

Of particular interest, we found that the *intI1* and *intI* genes had the most intensive connections with ARG subtypes (Fig. 4a), and co-occurred with multiple ARGs conferring resistance to all the major classes of antibiotics (Fig. S6). Some of these ARG subtypes have been reported to be located within integron gene cassettes, such as the *aac* and *aad* families of ARGs (Zhu *et al.*, 2017). Indeed, the *aac(6')-lb*, *aadA5*, and *aadA9* genes in the same cluster with the *intI1* and *intI* genes are known components of integron gene cassettes (Partridge *et al.*, 2009). The co-occurring ARGs and MGEs suggest that these ARGs might be genetically linked together in integrons, transposons and microbial genomes conferring resistance to multiple antibiotics, and likely subject to a synchronous dispersal. However, it should be noted that although there are strong correlations between the diversity, abundance and compositions of ARGs and MGEs, it remains unknown whether HGT is actually happening in forest ecosystems and would disperse environmental ARGs.

Despite of the genetic potential of HGT, the actual frequencies of HGT for soil ARGs are also influenced by a range of abiotic factors (e.g. soil pH, moisture content, temperature, and oxygen levels) and biotic factors (e.g. commensal, antagonistic and mutualistic relationships between soil organisms) in the environment (Aminov, 2011). It remains largely unknown regarding the frequency of HGT in natural settings with relatively lower selection

Wiley-Blackwell and Society for Applied Microbiology
This article is protected by copyright. All rights reserved.

Author Manuscript

pressure compared with contaminated environments or agricultural soils (Johnson *et al.*, 2016; Hu *et al.*, 2017). In fact, it is assumed that the rate of HGT is very low in natural soils, where ARGs content and phylogenetic compositions are strongly associated (Forsberg *et al.*, 2014) suggesting that HGT is not frequent to effectively decouple resistomes from phylogeny. The strong correlation between the ARG content and the bacterial community composition was also observed (Fig. 5 and Table S5), suggestive of a low rate of HGT in forests. However, ARGs are not linked to the phylogenetic composition in animal agriculture (Johnson *et al.*, 2016) and heavy metal-contaminated soils (Hu *et al.*, 2017), indicating that genetic potential for HGT might be enhanced when natural soils are subjected to anthropogenic activities.

Bacterial and herb communities as important determinants of ARGs in forests

The diverse and abundance ARGs detected in natural forests prompted us to identify the major factors shaping their distribution patterns, enabling prediction of the emergence and evolution of natural soil resistome in future. By compiling all the data together, our SEM demonstrated that the positive direct effects of bacterial communities and the positive indirect effects of herbaceous plants on ARG patterns were maintained even when simultaneously accounting for multiple drivers (climate, spatial predictors and edaphic factors). All these findings led to the conclusion that bacterial and herb communities are important predictors of ARG contents in natural forest biomes. Previous studies have reported that the microbial phylogenetic and taxonomic structure is an important determinant of ARG compositions (Forsberg *et al.*, 2014; Hu *et al.*, 2016b; Wang *et al.*, 2016a), because soil microbes are not only the major sources of antibiotic compounds or their derivatives (Allen *et al.*, 2010; Wright and Poinar, 2012) but are also hosts to many ARGs. For example, Actinobacteria are well-known groups of bacteria producing antibiotics; Proteobacteria and Actinobacteria are among the most prevalent predicted hosts of multi-resistant ARGs in metagenomics studies of soils (D'Costa *et al.*, 2006; Forsberg *et al.*, 2014), which is supported by the network analysis in our study (Fig. S7). The strong impact of microbial community structure on ARG contents might be also attributed to low rate of HGT which have not decoupled ARG contents and bacterial phylogeny.

Interestingly, we found that the herb community can indirectly affect ARG contents via its positive impacts on the microbial diversity (Fig. 5a), which might be due to the long-evolved intimate relationships between plants and microbes (Bai *et al.*, 2015). Although plants can exude specialized antibiotic compounds, which can lead to shifts of ARG contents (Yergeau *et al.*, 2014), we did not observe direct impacts of plant communities (including

Wiley-Blackwell and Society for Applied Microbiology

This article is protected by copyright. All rights reserved.

Author Manuscript

trees and herbs) on ARG contents in this study. Because our sampling strategy had tried to collect bulk soils between trees, the selective pressure imposed by the plant exudation of antimicrobials in the plant rhizosphere might have been overlooked. Although trees are the dominant primary producers responsible for >90% of the forest primary production and provide the majority of soil organic materials through root exudates and litters (Baldrian, 2017), understory herbs, rather than trees, were found to be a more important determinant of ARG patterns. Previous studies have also shown that herbs, rather than trees, are better predictors of soil prokaryotic diversity in forest ecosystems (Wang *et al.*, 2016b) and in grasslands (Prober *et al.*, 2015). The possible explanation might be that soil microbes in the topsoil are in more contact with herbs, because of the spatial separation between the roots of herbs and trees (Scholes and Archer, 1997); and herb residues are more easily decomposed by free-living soil microorganisms, compared to tree residues and litter (Hattenschwiler *et al.*, 2005). The impacts of herbaceous plants on the diversity and structure of soil microbial communities will ultimately have consequences on the soil ARG profiles.

Despite of the significant impacts of bacterial and herb communities on ARG contents, a range of other drivers including climate (AMT, annual mean temperature; MTCM, mean temperature of coldest month; and PDM, precipitation of driest month), edaphic factors (mainly soil pH) and geographic locations are also found to influence antibiotic resistome in natural forests (Fig. 5). Soil pH and climate have been recognized as the main driver of soil microbial diversity patterns across different ecosystems (Fierer and Jackson, 2006), implying that the factors driving the distribution of bacteria could underlie the diversity and distribution of ARGs. In addition, soil fungi are also important microbial communities in various forest ecosystems (Baldrian, 2017), and can produce antibiotics and compete with other soil microorganisms for space and nutrients (Chen *et al.*, 2013). Therefore, we argue that global environmental changes such as climate warming, extreme weather events, nitrogen deposition that have strong impacts on the plant, bacterial and fungal communities will also have consequences on soil resistome. Characterizing relationships between ARGs, bacteria, fungi, and plants in various ecosystems across broad environmental gradients will be critical to predicting the patterns of environmental resistome in a changing world.

Conclusions

To our knowledge, this is the first large-scale profiling of a broad spectrum of ARGs in natural forest biomes, providing an overview of how the interactions of aboveground plants and belowground microbes as well as soil properties and climate shape the distribution

Wiley-Blackwell and Society for Applied Microbiology
This article is protected by copyright. All rights reserved.

Author Manuscript

patterns of antibiotic resistance. Our results suggest that natural forest soils are a significant source of ARGs, and their large-scale distribution is regulated by the diversity of bacteria and herbaceous plants even when simultaneously accounting for multiple environmental drivers. These findings contribute to a growing understanding of the magnitude of environmental ARGs, and have important implications for future prediction of the evolution, dissemination and persistence of antibiotic resistance over broad spatial scales. Given the transmission potential of environmental ARGs to human commensals and pathogens, an improved knowledge of the factors maintaining environmental resistome will help the development of effective mitigation options to minimise the spread of antibiotic resistance.

Experimental procedures

Plant inventory and soil sampling campaign

We carried out a large-scale investigation in 2012 across a latitudinal gradient ranging from N18°15' to N53°18' over 4000 km in China (Fig. 1a). The selected sampling sites span over five major forest biomes from north to south: boreal forest, temperature mixed coniferous forest, temperate deciduous forest, subtropical (evergreen broadleaf) forest, and tropical (rainforest) forest. The sampling sites were selected from National Nature Reserves that are distant from anthropogenic disturbance. Our sampling strategy provides a wide variation in soil properties with soil pH from 4.03 to 8.57, total nitrogen (TN) from 0.05 to 1.16%, and soil organic carbon (SOC) from 0.39 to 19.21%. The mean annual temperature (MAT) ranges from -5.6°C to 25.9°C and the mean annual precipitation (MAP) varies from 452 to 2085 mm. The detailed information about the sampling sites and climate is shown in Table S1.

We established 60 independent large quadrats (20 × 20 m^2) in each forest biome, resulting in a total of 300 quadrats for the five forest biomes. Within each quadrat, we recorded all trees with diameter at breast height ≥2.5 cm and classified them into the species level. Three small plots (1 × 1 m^2) were randomly established within each quadrat for the census of herbaceous plants (Fig. 1a). Taxonomic classification of trees and herbs was performed based on the Angiosperm Phylogeny Group III System (Chase and Reveal, 2009). Soil samples were collected from each quadrat by thoroughly mixing 15 random soil cores (5 cm in diameter and 10 cm in depth), sieved through 1.0 mm mesh and pooled into a composite sample for each quadrat. A total of 300 soil samples were sealed in plastic bags

This article is protected by copyright. All rights reserved.

and shipped on ice to the laboratory. A portion of soils was used for physicochemical analysis, and another portion was freeze-dried and stored at -80°C until DNA extraction.

Soil physicochemical characterization

Soil moisture content was measured by oven-drying fresh soil samples at 105°C. Soil pH was determined in soil suspension with a soil to water ratio of 2.5 using a Delta 320 pH-meter (Mettler-Toledo Instruments, Columbus, OH, USA). TN was measured on a Vario EL III Element Analyser (Elementar, Hanau, Germany). SOC was determined on the Element Analyser (Elementar) after treatment with H_2SO_4 to remove carbonate. Soil inorganic nitrogen (including NH_4^+-N and NO_3^--N) was extracted from soils with 1 M KCl and measured using a SAN++ Continuous Flow Analyser (Skalar, Breda, Netherlands). Soil available phosphorous (AP) was extracted from soils by 0.5 M $NaHCO_3$ solution and measured on the Continuous Flow Analyser (Skalar). Soil available potassium (AK) was determined using Inductively Coupled Plasma-Optical Emission Spectroscopy (ICP-OES, Vista axial, Varian). Soil particle size was determined using a Laser Particle Size Analyser (Malvern Instruments, Malvern, UK).

DNA extraction and HT-qPCR analysis

The MoBio PowerSoil DNA Isolation Kit (MoBio Laboratories, Carlsbad, CA, USA) was used to extract soil DNA from 0.25 g of soil as per the manufacturer's recommendations. DNA quality was evaluated using IMPI EN P-330 NanoPhotometer UV/VIS spectrophotometer (IMPLEN, Munich, Germany). A HT-qPCR array was performed to determine the diversity and abundance of ARGs on the Wafergen SmartChip Real-Time PCR System (Fremont, CA, USA) as described previously (Zhu *et al.*, 2013; Su *et al.*, 2015). The Wafergen system is a high-throughput platform with the capability to run 5184 reactions each with a volume of 100 nl using the SensiMix SYBR No-ROX reagent (Bioline, London, UK). The HT-qPCR array contained a total of 296 primer sets (Looft *et al.*, 2012; Su *et al.*, 2015), including 285 primer sets targeting eight major classes of ARGs, 10 primer sets targeting MGEs, and one 16S rRNA gene as the internal control (Table S2). The specificity of all primers was verified through amplicon sequencing, BLAST search and amplification efficiency check (Looft *et al.*, 2012; Johnson *et al.*, 2016).

Thermal-cycling conditions were as follows: 95°C for 10 min, followed by 40 cycles of 95°C for 30 s and 60°C for 30 s. All HT-qPCR reactions were performed in three technical replicates. ARGs or MGEs were regarded as positive detection according to the following

Wiley-Blackwell and Society for Applied Microbiology

This article is protected by copyright. All rights reserved.

Author Manuscript

criteria: (i) a threshold cycle value (C_T) of 31 was used as the detection limit to differentiate between positive amplification and primer-dimers (Su *et al.*, 2015); (ii) samples with three technical replicates all above the detection limit were regarded as positive quantification; (iii) amplicons with multiple melting curves were removed from the analysis; and (iv) reactions with efficiency beyond the range 1.7-2.3 and R^2 values below 0.99 were discarded. The $2^{-\Delta CT}$ method where $\Delta C_T = (C_{T\ \text{detected ARGs}} - C_{T\ \text{16S rRNA gene}})$ was used to calculate the relative abundances of ARGs and MGEs compared to the 16S rRNA gene abundance based on a comparative C_T method (Schmittgen and Livak, 2008).

Illumina Miseq sequencing and data processing

The V4 region of the bacterial 16S rRNA gene was amplified using the primer set 515F/806R (Bates *et al.*, 2011) linked with the Illumina adaptor and a 12-bp barcode sequence. PCR amplification was performed in 50 μl mixtures containing 25 μl Premix Taq (TaKaRa, Dalian, China), 1 μl each primer (10 μM), and 3 μl template DNA (~10 ng). The barcoded PCR products were purified using Wizard SV Gel and PCR Clean-up System (Promega, Madison, WI, USA), quantified, pooled into one mixture, and sequenced using 2× 150-bp paired-end strategy on an Illumina MiSeq sequencer. Raw and paired-end sequences were quality filtered, assembled using the Fast Length Adjustment of Short Reads (FLASH) (Magoè and Salzberg, 2011), de-multiplexed, and assigned to individual samples using Quantitative Insights Into Microbial Ecology (QIIME) analysis (Caporaso *et al.*, 2010). Assembled sequences were binned into operational taxonomic units (OTUs) at the 97% sequence identity using a chimera filtering approach UPARSE (Edgar, 2013). An even number of sequences per sample (45,070 reads) was randomly selected from each sample to compensate for variability in sequencing depth before the downstream analysis. Taxonomic identity was assigned to OTUs using the naïve Bayesian classifier Ribosome Database Project (Wang *et al.*, 2007). The assembled bacterial 16S rRNA gene sequences have been deposited in the European Nucleotide Archive under the accession number PRJEB11871.

Local similarity analysis and network construction

Network analysis has been widely used to explore the underlying interaction/association among microbial taxa, to predict the host of ARGs in complex microbial communities (Hu *et al.*, 2016b; Zhu *et al.*, 2017). The Local Similarity (LS) correlations (ranked Pearson's correlations) among bacteria, herbs and ARGs were determined by the LS analysis (eLSA) (Chow *et al.*, 2014). A correlation matrix was constructed by calculating the pairwise

Wiley-Blackwell and Society for Applied Microbiology
This article is protected by copyright. All rights reserved.

Author Manuscript

correlation coefficients among ARG subtypes and bacterial and herb taxa at the family level which occur in ≥ 50% of all the samples. This filtering step eliminated those poorly represented ARGs and bacterial and herb taxa to minimize the artificial association bias. We calculated *P*-values of the LS correlations using statistical approximation followed by permutation testing (Xia *et al*., 2013) following the protocol as previously described (Chow *et al.*, 2014). We imported the resultant LS correlations into Cytoscape (version 2.8.3) (Shannon *et al.*, 2003) for initial visualization, and explored the network topology in the Gephi platform (Bastian *et al.*, 2009).

Statistical analysis

We calculated the tree diversity by counting the number of tree species in 400 m^2 quadrats, and the herb diversity by counting the number of herb species in 1 m^2 plots. Soil bacterial diversity was calculated for each sample using the number of OTUs at the 97% sequence similarity with a depth of 45,070 sequences per sample. One-way analysis of variance (ANOVA) was performed to test the difference in the diversity of ARGs, MGEs, trees, herbs, and bacteria, as well as the relative abundances of ARGs and MGEs across forest biomes in SPSS 20. Spearman's correlation analysis was conducted to test the relationships between the relative abundances of total MGEs with total ARGs and individual ARG types. The data were normalized using log-transformation when necessary prior to statistical analysis.

We conducted regression analysis to test the relationships between the diversity of ARGs with that of MGEs, bacteria, trees, and herbs as well as soil properties and climatic factors using the "ggplot2" package in the R platform (R Development Core Team, 2008). The curve fitting was estimated with linear and nonlinear models (e.g., quadratic and cubic models), and the best models were selected with the highest proportion of variance in the diversity of ARGs that can be explained by the independent variables. A heat map visualization of the relative abundances of individual ARG subtypes was generated using the "gplots" package in the R platform (version 3.3.1). The changes in ARG compositions across forest biomes were visualized by non-metric multidimensional scaling (NMDS) ordinations based on the Bray-Curtis dissimilarity distances, and statistically evaluated by permutational multivariate analysis of variance (PerMANOVA) using the "vegan" package with 999 permutations in R. Procrustes test for correlation analysis between ARGs and bacterial communities based on Bray-Curtis dissimilarity matrices was performed using "vegan" in R. Mantel test was performed to test the correlations between soil properties, MGEs, the

Wiley-Blackwell and Society for Applied Microbiology

This article is protected by copyright. All rights reserved.

Author Manuscript

community composition of trees and herbs with ARGs content with the Bray-Curtis method in "vegan" with 999 permutations. Canonical correspondence analysis (CCA) between ARG contents and environmental parameters was also performed in R using "cca" function from the "vegan" package. The threshold for significance was $P < 0.05$.

Random Forest analysis and Structural equation modelling

A classification Random Forest analysis was conducted to identify the most important predictors influencing the ARGs content among the following variables: climatic factors, spatial factors, soil properties, herb diversity, and bacterial diversity. The importance of each predictor was determined by assessing the increase in the mean square error (MSE) between observations and out-of-bag predictions as described previously (Delgado-Baquerizo *et al.*, 2016). The significance of the importance of each predictor on ARG patterns was assessed using the "rfPermute" package in R.

We constructed structural equation model (SEM) to determine the direct and indirect effects of soil properties (including soil pH), climatic conditions (including AMT), the diversity of bacteria, trees and herbs, and MGEs, on ARG patterns. SEM can partition direct and indirect effects that one variable might have on another, estimate the strengths of multiple effects, and ultimately provide mechanistic information on the drivers of ARGs pattern (Grace, 2006). We only included those variables that were identified as statistically significant predictors of the ARG patterns from the Random Forest analysis. The data used for the model construction were classified into several groups, including climate, spatial factors, soil properties, plant diversity, MGEs diversity, and ARGs diversity.

We established an *a priori* model based on the known relationships among the drivers of ARG patterns (Fig. S9). Mantel test was performed to calculate the pairwise correlations among these variables using the "Ecodist" package in R, and the matrix of resultant *R* values was parameterized into the AMOS 22 software for the SEM construction using the maximum-likelihood estimation method. Foe the mediation testing, we used a 1000-time bootstrap to control the false discovery rate (FDR) and set the confidence intervals at 95% for the significance assessment. The overall goodness of model fit was assessed using Akaike information criteria (AIC), goodness-of-fit index (> 0.90), χ^2 test ($P > 0.05$), and root mean square errors of approximation (RMSEA < 0.05). We also calculated the standardized total effects of each factor on ARG patterns by summing all direct and indirect pathways between the factor and ARGs.

Wiley-Blackwell and Society for Applied Microbiology

This article is protected by copyright. All rights reserved.

Author Manuscript

Author Manuscript

Acknowledgements

This work was financially supported by Chinese Academy of Sciences (XDB15020200), Australian Research Council (DP170103628, DE150100870), and Natural Science Foundation of China (41601256).

References

Allen, H.K., Moe, L.A., Rodbumrer, J., Gaarder, A., and Handelsman, J. (2009) Functional metagenomics reveals diverse beta-lactamases in a remote Alaskan soil. *ISME J* **3**: 243-251.

Allen, H.K., Donato, J., Wang, H.H., Cloud-Hansen, K.A., Davies, J., and Handelsman, J. (2010) Call of the wild: antibiotic resistance genes in natural environments. *Nat Rev Microbiol* **8**: 251-259.

Alonso, A., Sánchez, P., and Martínez, J.L. (2001) Environmental selection of antibiotic resistance genes. *Environ Microbiol* **3**: 1-9.

Aminov, R.I. (2011) Horizontal gene exchange in environmental microbiota. *Front Microbiol* **2**: 158.

Ashbolt, N.J., Amezquita, A., Backhaus, T., Borriello, P., Brandt, K.K., Collignon, P. et al. (2013) Human health risk assessment (HHRA) for environmental development and transfer of antibiotic resistance. *Environ Health Perspect* **121**: 993-1001.

Bai, Y., Muller, D.B., Srinivas, G., Garrido-Oter, R., Potthoff, E., Rott, M. et al. (2015) Functional overlap of the Arabidopsis leaf and root microbiota. *Nature* **528**: 364-369.

Baldrian, P. (2017) Forest microbiome: diversity, complexity and dynamics. *FEMS Microbiol Rev* **41**: 109-130.

Balsalobre, L.C., Dropa, M., Lincopan, N., Mamizuka, E.M., Matte, G.R., and Matte, M.H. (2009) Detection of metallo-β-lactamases encoding genes in environmental isolates of Aeromonas hydrophila and Aeromonas jandaei. *Lett Appl Microbiol* **49**: 142-145.

Bastian, M., Heymann, S., and Jacomy, M. Gephi: an open source software for exploring and manipulating networks. In: International AAAI Conference on Weblogs and Social Media: San Jose, California 2009.

This article is protected by copyright. All rights reserved.

Author Manuscript

Bates, S.T., Berg-Lyons, D., Caporaso, J.G., Walters, W.A., Knight, R., and Fierer, N. (2011) Examining the global distribution of dominant archaeal populations in soil. *ISME J* **5**: 908-917.

Bonan, G.B. (2008) Forests and climate change: forcings, feedbacks, and the climate benefits of forests. *Science* **320**: 1444-1449.

Berendonk, T.U., Manaia, C.M., Merlin, C., Fatta-Kassinos, D., Cytryn, E., Walsh, F. et al. (2015) Tackling antibiotic resistance: the environmental framework. *Nat Rev Microbiol* **13**: 310-317.

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K. et al. (2010) QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* **7**: 335-336.

Chase, M.W., and Reveal, J.L. (2009). A phylogenetic classification of the land plants to accompany ARG III. *Bot J Linean Soc* **161**: 122-127.

Chen, B., Yang, Y., Liang, X., Yu, K., Zhang, T., and Li, X. (2013) Metagenomic profiles of antibiotic resistance genes (ARGs) between human impacted estuary and deep ocean sediments. *Environ Sci Technol* **47**: 12753-12760.

Chen, Q.L., An, X.L., Li, H., Zhu, Y.G., Su, J.Q., and Cui, L. (2017) Do manure-borne or indigenous soil microorganisms influence the spread of antibiotic resistance genes in manured soil? *Soil Biol Biochem* **114**: 229-237.

Chow, C.E.T., Kim, D.Y., Sachdeva, R., Caron, D.A., and Fuhrman, J.A. (2014) Top-down controls on bacterial community structure: microbial network analysis of bacteria, T4-like viruses and protists. *ISME J* **8**: 816-829.

D'Costa, V.M., McGrann, K.M., Hughes, D.W., and Wright, G.D. (2006) Sampling the antibiotic resistome. *Science* **311**: 374-377.

D'Costa, V.M., King, C.E., Kalan, L., Morar, M., Sung, W.W.L., Schwarz, C. et al. (2011) Antibiotic resistance is ancient. *Nature* **477**: 457-461.

De Souza, M.J., Nair, S., Bharathi, P.L., and Chandramohan, D. (2006) Metal and antibiotic-resistance in psychrotrophic bacteria from Antarctic marine waters. *Ecotoxicology* **15**: 379-384.

Delgado-Baquerizo, M., Maestre, F.T., Reich, P.B., Jeffries, T.C., Gaitan, J.J., Encinar, D. et al. (2016) Microbial diversity drives multifunctionality in terrestrial ecosystems. *Nat Commun* **7**: 10541.

Edgar, R.C. (2013) UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat Methods* **10**: 996-998.

Wiley-Blackwell and Society for Applied Microbiology

This article is protected by copyright. All rights reserved.

Fierer, N., and Jackson, R.B. (2006) The diversity and biogeography of soil bacterial communities. *Proc Natl Acad Sci USA* **103**: 626-631.

Forsberg, K.J., Patel, S., Gibson, M.K., Lauber, C.L., Knight, R., Fierer, N., et al. (2014) Bacterial phylogeny structures soil resistomes across habitats. *Nature* **509**: 612-616.

Forsberg, K.J., Reyes, A., Wang, B., Selleck, E.M., Sommer, M.O.A., and Dantas, G. (2012) The shared antibiotic resistome of soil bacteria and human pathogen. *Science* **337**: 1107-1111.

Gibson, M.K., Forsberg, K.J., and Dantas, G. (2015) Improved annotation of antibiotic resistance determinants reveals microbial resistomes cluster by ecology. *ISME J* **9**: 207-216.

Gillings, M.R., Gaze, W.H., Pruden, A., Smalla, K., Tiedje, J.M., and Zhu, Y.G. (2015) Using the class 1 integron-integrase gene as a proxy for anthropogenic pollution. *ISME J* **9**: 1269-1279.

Gilmore, M.S., and Hoch, J.A. (1999) Antibiotic resistance: a vancomycin surprise. *Nature* **399**: 524-527.

Grace, J.B. (2006) Structural Equation Modelling and Natural Systems. Cambridge University Press.

Han, X.M., Hu, H.W., Shi, X.Z., Wang, J.T., Han, L.L., Chen, D. et al. (2016) Impacts of reclaimed water irrigation on soil antibiotic resistome in urban parks of Victoria, Australia. *Environ Pollut* **211**: 48-57.

Hattenschwiler, S., Tiunov, A.V., and Scheu, S. (2005) Biodiversity and litter decomposition in terrestrial ecosystems. *Annu Rev Ecol Evol* **36**: 191-218.

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

Hu, H.W., Han, X.M., Shi, X.Z., Wang, J.T., Han, L.L., Chen, D. et al. (2016a) Temporal changes of antibiotic resistance genes and bacterial communities in two contrasting soils treated with cattle manure. *FEMS Microbiol Biol* **92**: fiv169.

Hu, H.W., Wang, J.T., Li, J., Li, J.J., Ma, Y.B., Chen, D. et al. (2016b) Field-based evidence for copper contamination induced changes of antibiotic resistance in agricultural soils. *Environ Microbiol* **18**: 3896-3909.

Hu, H.W., Wang, J.T., Li, J., Shi, X.Z., Ma, Y.B., Chen, D. et al. (2017) Long-term nickel contamination increases the occurrence of antibiotic resistance genes in agricultural soils. *Environ Sci Technol* **51**: 790-800.

Wiley-Blackwell and Society for Applied Microbiology

This article is protected by copyright. All rights reserved.

Author Manuscript

Karkman, A., Do T.T., Walsh, F., and Virta, M.P.J. (2018) Antibiotic-resistance genes in waste water. *Trends Microbiol* **26**: 220-228.

Johnson, T.A., Stedtfeld, R.D, Wang, Q., Cole, J.R., Hashsham, S.A., Looft, T. et al. (2016) Clusters of antibiotic resistance genes enriched together stay together in swine agriculture. *mBio* **7**: e02214-15.

Li, B., Yang, Y., Ma, L., Ju, F., Guo, F., Tiedje, J.M., et al. (2015) Metagenomic and network analysis reveal wide distribution and co-occurrence of environmental antibiotic resistance genes. *ISME J* **9**: 2490-2502.

Looft, T., Johnson, T.A., Allen, H.K., Bayles, D.O., Alt, D.P., Stedtfeld, R.D., et al. (2012) In-feed antibiotic effects on the swine intestinal microbiome. *Proc Nat Acad Sci USA* **109**: 1691-1696.

Magoè, T., and Salzberg, S.L. (2011) FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* **27**: 2957-5963.

Martinez, J.L. (2008) Antibiotics and antibiotic resistance genes in natural environments. *Nature* **321**: 365-367.

Martinez, J.L., Coque, T.M., and Baquero, F. (2015) What is a resistance gene? Ranking risk in resistomes. *Nat Rev Microbiol* **13**: 116-123.

McEachran, A.D., Blackwell, B.R., Hanson, J.D., Wooten, K.J., Mayer, G.D., Cox, S.B. et al. (2015) Antibiotics, bacteria, and antibiotic resistance genes: aerial transport from cattle feed yards via particulate matter. *Environ Health Perspect* **123**: 337-343.

Nesme, J., Cécillon, S., Delmont, T.O., Monier, J.M., Vogel, T.M., Simonet, P. (2014) Large-scale metagenomic-based study of antibiotic resistance in the environment. *Curr Biol* **24**: 1096-1100.

Nesme, J., and Simonet, P. (2014) The soil resistome: a critical review on antibiotic resistance origins, ecology and dissemination potential in telluric bacteria. *Environ Microbiol* **17**: 913-930.

Pawlowski, A.C., Wang, W., Koteva, K., Barton, H.A., McArthur, A.G., and Wright, G.D. (2016) A diverse intrinsic antibiotic resistome from a cave bacterium. *Nat Commun* **7**: 13803.

Partridge, S.R., Tsafnat, G., Coiera, E., and Iredell, J.R. (2009) Gene cassettes and cassette arrays in mobile resistance integrons. *FEMS Microbiol Rev* **33**: 757-784.

Prober, S.M., Leff, J.W., Bates, S.T., Borer, E.T., Firn, J., Harpole, W.S. et al. (2015) Plant diversity predicts beta but not alpha diversity of soil microbes across grasslands worldwide. *Ecol Lett* **18**: 85-95.

Wiley-Blackwell and Society for Applied Microbiology

This article is protected by copyright. All rights reserved.

Author Manuscript

Pruden, A., Pei, R., Storteboom, H., and Carlson, K.H. (2006) Antibiotic resistance genes as emerging contaminants: studies in Northern Colorado. *Environ Sci Technol* **40**: 7445-7450.

R Development Core Team. R: A language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria. URL: http://www.R-project.org. 2008.

Schmittgen, T.D., and Livak, J.K. (2008) Analyzing real-time PCR data by the comparative CT method. *Nat Protoc* **3**: 1101-1108.

Scholes, R., and Archer, S. (1997) Tree-grass interactions in savannas. *Annu Rev Ecol Syst* **28**: 517-544.

Segawa, T., Takeuchi, N., Rivera, A., Yamada, A., Yoshimura, Y., Barcaza, G. et al. (2013) Distribution of antibiotic resistance genes in glacier environments. *Environ Microbiol Rep* **5**: 127-134.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D. et al. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **13**: 2498-2504.

Stalder, T., Barraud, O., Casellas, M., Dagot, C., and Ploy, M.C. (2012) Integron involvement in environmental spread of antibiotic resistance. *Front Microbiol* **3**: 119.

Stedtfeld, R.D., Williams, M., Fakher, U., Johnson, T., Stedtfeld, T.M., Wang, F. et al. (2016) Antimicrobial resistance (AR) dashboard application for mapping environmental occurrence and resistant pathogens. *FEMS Microbiol Ecol* **92**: fiw020.

Su, J.Q., Wei, B., Ou-Yang, W.Y., Huang, F.Y., Zhao, Y., Xu, H.J., et al. (2015) Antibiotic resistome and its association with bacterial communities during sewage sludge composting. *Environ Sci Technol* **49**: 7356-7363.

Wang, F., Stedtfeld, R.D., Kin, O.S., Chai, B., Yang, L., Stedtfeld, T.M. et al. (2016a) Influence of soil characteristics and proximity to Antarctic research stations on abundance of antibiotic resistance genes in soils. *Environ Sci Technol* **50**: 12621-12629.

Wang, F.H., Qiao, M., Su, J.Q., Chen, Z., Zhou, X., and Zhu, Y.G. (2014) High throughput profiling of antibiotic resistance genes in urban park soils with reclaimed water irrigation. *Environ Sci Technol* **48**: 9079-9085.

Wang, J.T., Zheng, Y.M., Hu, H.W., Li, J., Zhang, L.M., Chen, B.D. et al. (2016b) Coupling of soil prokaryotic diversity and plant diversity across latitudinal forest ecosystems. *Sci Rep* **6**: 19561.

Author Manuscript

Wiley-Blackwell and Society for Applied Microbiology

This article is protected by copyright. All rights reserved.

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

World Health Organization. (2014) Antibiotic resistance: global report on surveillance. World Health Organization, Geneva, Switzerland.

Wright, G.D. (2010) Antibiotic resistance in the environment: a link to the clinic? *Curr Opin Microbiol* **13**: 589-594.

Wright GD, and Poinar H. (2012) Antibiotic resistance is ancient: implications for drug discovery. *Trends Microbiol* **20**: 157-159.

Xia, L.C., Ai, D., Gram, J.A., Fuhrman, J.A., and Sun, F. (2013) Efficient statistical significance approximation for local association analysis of high-throughput time series data. *Bioinformatics* **29**: 230-237.

Xiang, Q., Chen, Q.L., Zhu, D., An, X.L., Yang, X.R., S, J.Q. et al. (2018) Spatial and temporal distribution of antibiotic resistomes in a peri-urban area is associated with anthropogenic activities. *Environ Pollut* **235**: 525-533.

Yergeau, E., Sanschagrin, S., Mynard, C., St-Arnaud, M., and Greer, C.W. (2014) Microbial expression profiles in the rhizosphere of willows depend on soil contamination. *ISME J* **8**: 344-358.

Zhang, Y.J., Hu, H.W., Gou, M., Wang, J.T., Chen, D.L., and He, J.Z. (2017) Temporal succession of soil antibiotic resistance genes following application of swine, cattle and poultry manures spiked with or without antibiotics. *Environ Pollut* **231**: 1621-1632.

Zhu, Y.G., Zhao, Y., Li, B., Huang, C.L., Zhang, S.Y., Yu, S. et al. (2017) Continental-scale pollution of estuaries with antibiotic resistance genes. *Nat Microbiol* **2**: 16270.

Zhu, Y.G., Johnson, T.A., Su, J.Q., Qiao, M., Guo, G.X., Stedtfeld, R.D. et al. (2013) Diverse and abundant antibiotic resistance genes in Chinese swine farms. *Proc Natl Acad Sci USA* **110**: 3435-3440.

Author Manuscript

Wiley-Blackwell and Society for Applied Microbiology

This article is protected by copyright. All rights reserved.

Figure legends

Figure 1 A schematic diagram showing the census of trees and herbs and soil sampling across five forest biomes over a span of 4000 km (a). The dots with different colours represent soil sampling sites for different forest biomes. We established 60 independent large quadrats (20 × 20 m^2) in each forest biome, and in each quadrat, all trees with diameter at breast height ≥2.5 cm were recorded. Three small plots (1 × 1 m^2) were randomly established within each 400-m^2 quadrat for the census of herbaceous plants. Soil samples were collected from each quadrat by thoroughly mixing 15 random soil cores, resulting in a total of 300 soil samples. Comparison of the diversity and relative of ARGs (b, d) and MGEs (c, e) across forest biomes. Line, average diversity or relative abundance; box, 25th and 75th percentiles; and error bars, 10th and 90th percentiles. Different letters above the boxes indicate a significant difference.

Figure 2 The heat map showing the broad-spectrum quantitative profile of ARGs and MGEs subtypes detected in five forest biomes (a). Soil samples were shown as columns and ARGs/MGEs shown as rows. The dendrograms for columns and rows were calculated based on the Bray-Curtis dissimilarity matrices. Non-metric multidimensional scaling ordination plot depicts the Bray-Curtis dissimilarity distances between soils based on the relative abundance of ARGs across forest biomes (b). The 2D stress value is 0.14, which indicated that the two-dimensional ordinations could well represent the data. Procrustes analyses depict significant correlations between ARG contents (Bray-Curtis) and bacterial composition (Bray-Curtis) (c). Lines connect ARG data and 16S rRNA gene sequence data. The M^2 fit reported is from a Procrustes transformation over the first two principal coordinates while the *P*-value is calculated from a distribution of empirically determined M^2 values over 10,000 Monte Carlo label permutations. (Abbreviations: FCA: fluoroquinolone, quinolone, florfenicol, chloramphenicol and amphenicol resistance genes; MLSB: Macrolide-Lincosamide-Streptogramin B resistance genes.)

Figure 3 Relationships between the diversity of ARGs with the bacterial diversity (a), herb diversity (b), MGEs diversity (c), and soil pH (d). Line in each plot represents the best-fit curve and the shaded area represents its 95% confidence limits.

Figure 4 The networks showing the co-occurrence patterns among the detected ARG subtypes, MGEs, bacteria, and herbs across forest biomes. (a) The nodes coded with different colors represent ARGs, MGEs, bacteria and herbs; (b) The nodes coded with different colors represent different categories of ARGs, and the nodes for MGEs, bacteria and herbs are

Wiley-Blackwell and Society for Applied Microbiology

This article is protected by copyright. All rights reserved.

Author Manuscript

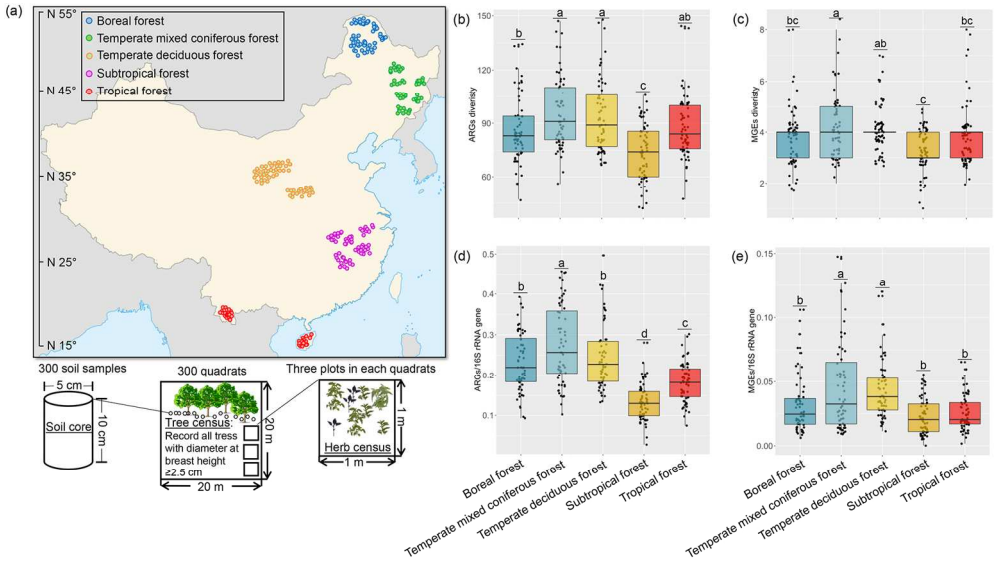
shown in grey colour. The edges connecting nodes correspond to statistically significant correlations between nodes. Node size is proportional to the number of connections between nodes (degree). The thickness of the edges is proportional to the correlation coefficient.

Figure 5 Structural equation modelling of the direct and indirect effects of molecular attributes (bacterial diversity), plant attributes (tree and herb diversity), soil properties, climatic conditions, biome types, and MGEs on the diversity of ARGs (a). The hypothetical models fit our data well, as suggested by the goodness-of-fit statistics: $\chi^2 = 33.0$, $P = 0.11$, d.o.f = 24, GFI = 0.98, AIC = 167.0, and RMSEA = 0.035. Arrows indicate the hypothesized direction of causation. The width of arrows is proportional to the strength of the path coefficients. The numbers adjacent to arrows are standardized path coefficients and indicative of the effect size of the relationship. The R^2 value alongside the response variable indicates the proportion of variance explained by the variable. Only significant path coefficients are displayed. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$. (b) Standardized total effects (the sum of direct plus indirect effects from SEM) of the major factors on the diversity of ARGs. (AMT, annual mean temperature; PDM, Precipitation of driest month; MTCM, Mean temperature of coldest month)

Author Manuscript

Wiley-Blackwell and Society for Applied Microbiology

This article is protected by copyright. All rights reserved.

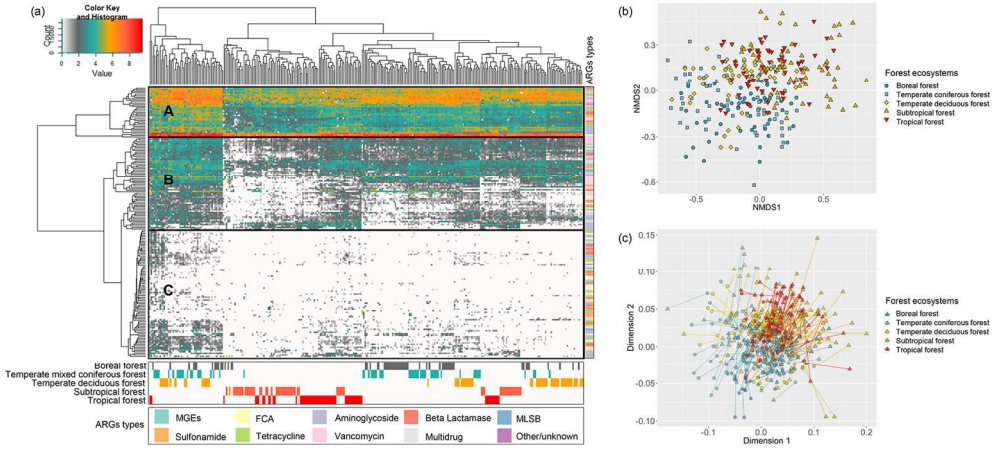


A schematic diagram showing the census of trees and herbs and soil sampling across five forest biomes over a span of 4000 km (a). The dots with different colours represent soil sampling sites for different forest biomes. We established 60 independent large quadrats (20 × 20 m2) in each forest biome, and in each quadrat, all trees with diameter at breast height ≥2.5 cm were recorded. Three small plots (1 × 1 m2) were randomly established within each 400-m2 quadrat for the census of herbaceous plants. Soil samples were collected from each quadrat by thoroughly mixing 15 random soil cores, resulting in a total of 300 soil samples. Comparison of the diversity and relative of ARGs (b, d) and MGEs (c, e) across forest biomes. Line, average diversity or relative abundance; box, 25th and 75th percentiles; and error bars, 10th and 90th percentiles. Different letters above the boxes indicate a significant difference.

169x95mm (300 x 300 DPI)

Author

This article is protected by copyright. All rights reserved.

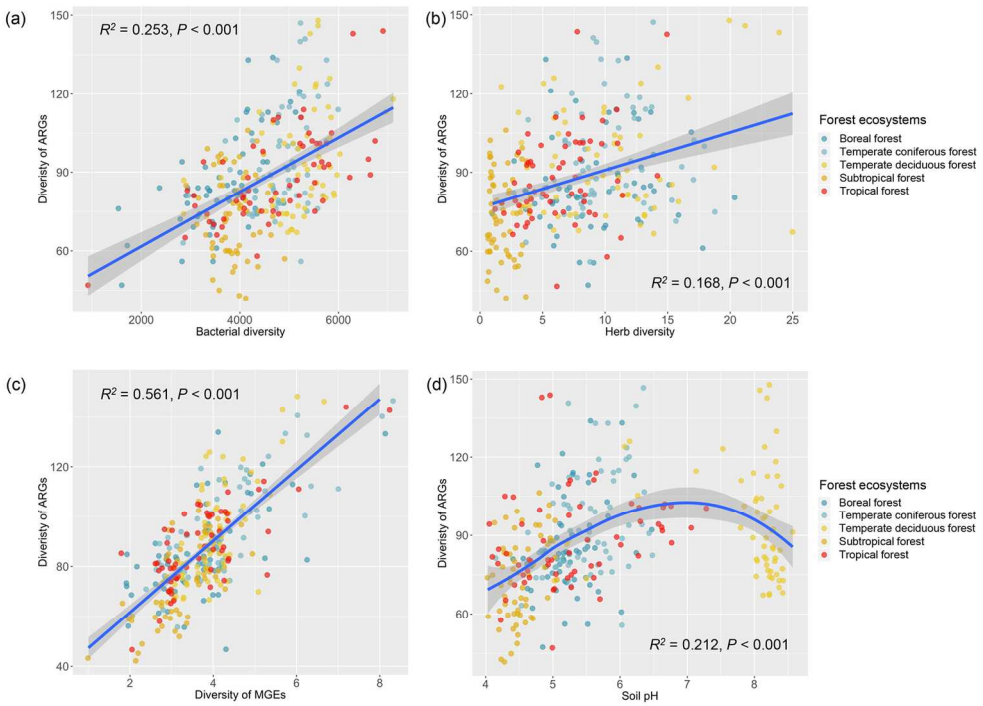


The heat map showing the broad-spectrum quantitative profile of ARGs and MGEs subtypes detected in five forest biomes (a). Soil samples were shown as columns and ARGs/MGEs shown as rows. The dendrograms for columns and rows were calculated based on the Bray-Curtis dissimilarity matrices. Non-metric multidimensional scaling ordination plot depicts the Bray-Curtis dissimilarity distances between soils based on the relative abundance of ARGs across forest biomes (b). The 2D stress value is 0.14, which indicated that the two-dimensional ordinations could well represent the data. Procrustes analyses depict significant correlations between ARG contents (Bray-Curtis) and bacterial composition (Bray-Curtis) (c). Lines connect ARG data and 16S rRNA gene sequence data. The M2 fit reported is from a Procrustes transformation over the first two principal coordinates while the P-value is calculated from a distribution of empirically determined M2 values over 10,000 Monte Carlo label permutations. (Abbreviations: FCA: fluoroquinolone, quinolone, florfenicol, chloramphenicol and amphenicol resistance genes; MLSB: Macrolide-Lincosamide-Streptogramin B resistance genes.)

160x75mm (300 x 300 DPI)

Author

This article is protected by copyright. All rights reserved.

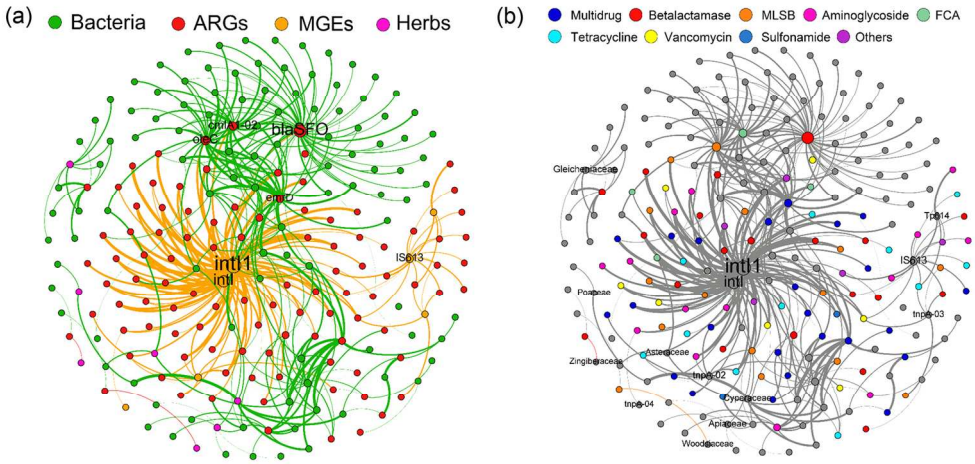


Relationships between the diversity of ARGs with the bacterial diversity (a), herb diversity (b), MGEs diversity (c), and soil pH (d). Line in each plot represents the best-fit curve and the shaded area represents its 95% confidence limits.

160x118mm (300 x 300 DPI)

Author

This article is protected by copyright. All rights reserved.

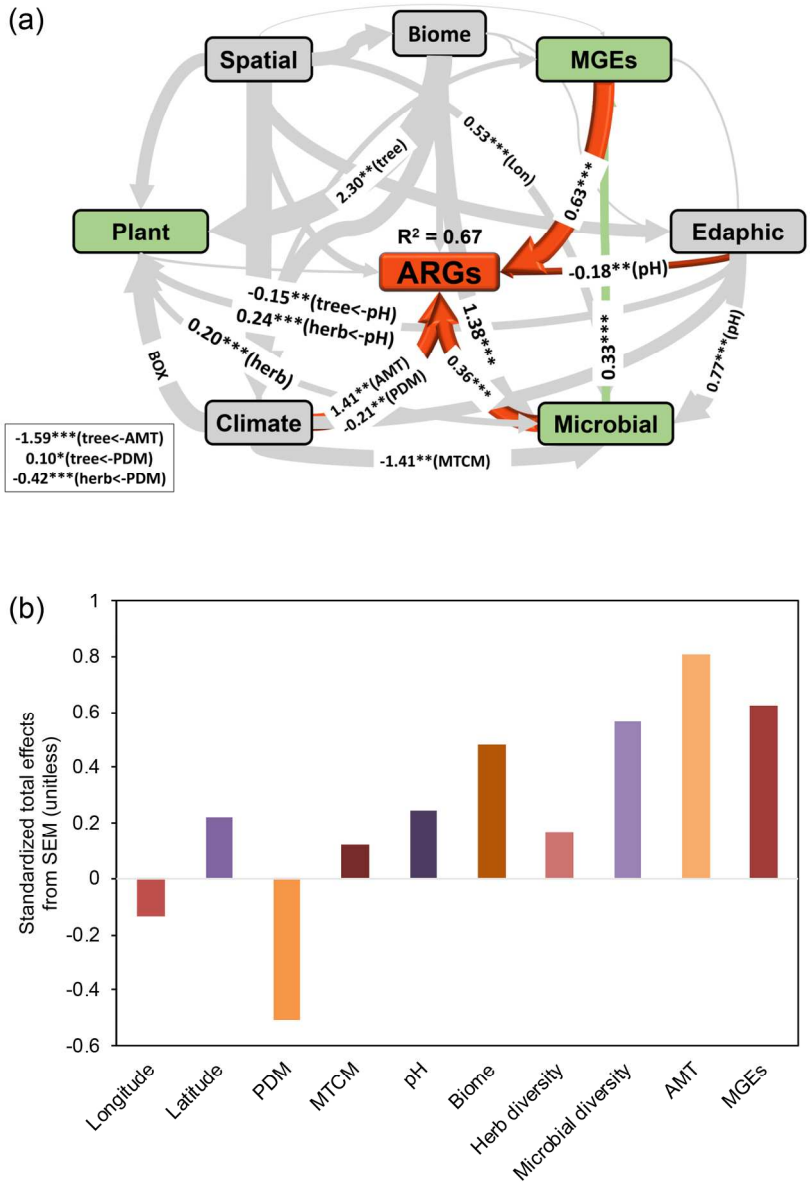


The networks showing the co-occurrence patterns among the detected ARG subtypes, MGEs, bacteria, and herbs across forest biomes. (a) The nodes coded with different colors represent ARGs, MGEs, bacteria and herbs; (b) The nodes coded with different colors represent different categories of ARGs, and the nodes for MGEs, bacteria and herbs are shown in grey colour. The edges connecting nodes correspond to statistically significant correlations between nodes. Node size is proportional to the number of connections between nodes (degree). The thickness of the edges is proportional to the correlation coefficient.

160x79mm (300 x 300 DPI)

Author M

This article is protected by copyright. All rights reserved.



Structural equation modelling of the direct and indirect effects of molecular attributes (bacterial diversity), plant attributes (tree and herb diversity), soil properties, climatic conditions, biome types, and MGEs on the diversity of ARGs (a). The hypothetical models fit our data well, as suggested by the goodness-of-fit statistics: $\chi 2 = 33.0$, $P = 0.11$, d.o.f = 24, GFI = 0.98, AIC = 167.0, and RMSEA = 0.035. Arrows indicate the hypothesized direction of causation. The width of arrows is proportional to the strength of the path coefficients. The numbers adjacent to arrows are standardized path coefficients and indicative of the effect size of the relationship. The R2 value alongside the response variable indicates the proportion of variance explained by the variable. Only significant path coefficients are displayed. *P < 0.05, **P < 0.01, ***P < 0.001. (b) Standardized total effects (the sum of direct plus indirect effects from SEM) of the major factors on the diversity of ARGs. (AMT, annual mean temperature; PDM, Precipitation of driest month; MTCM, Mean temperature of coldest month)

160x231mm (300 x 300 DPI)

This article is protected by copyright. All rights reserved.

Author Manuscript

Wiley-Blackwell and Society for Applied Microbiology

This article is protected by copyright. All rights reserved.

Author Manuscript

Title page

Diversity of herbaceous plants and bacterial communities regulates soil resistome across forest biomes

Hang-Wei Hu[1,2,6], Jun-Tao Wang[1,6], Brajesh K. Singh[3,4], Yu-Rong Liu[1], Yong-Liang Chen[5], Yu-Jing Zhang[2], Ji-Zheng He[1,2,*]

1. State Key Laboratory of Urban and Regional Ecology, Research Centre for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing 100085, China

2. Faculty of Veterinary and Agricultural Sciences, The University of Melbourne, Parkville, Victoria 3010, Australia

3. Hawkersbury Institute for the Environment, Western Sydney University, Penrith South DC, NSW 2751, Australia

4. Global Centre for Land-Based Innovation, Western Sydney University, Penrith South DC, NSW 2751, Australia

5. State Key Laboratory of Vegetation and Environmental Change, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China

6. These authors contributed equally to this work.

***Author for correspondence**

Ji-Zheng He, Research Centre for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing 100085, China. Email: jzhe@rcees.ac.cn

Running title: Plant communities regulate forest soil resistomes

Wiley-Blackwell and Society for Applied Microbiology

This article is protected by copyright. All rights reserved.

Author Manuscript

Originality-Significance statement

This is the first large-scale profiling of a broad spectrum of antibiotic resistance genes (ARGs) in natural forest biomes, providing important information on how the interactions of aboveground plants and belowground microbes as well as soil properties and climate shape the distribution patterns of antibiotic resistance. Our results suggest that natural forest soils are a significant source of ARGs, and their large-scale distribution is regulated by the diversity of bacteria and herbaceous plants even when simultaneously accounting for multiple environmental drivers. This knowledge is critical in understanding the origins and evolution of natural soil resistome, and may help to design effective framework to predict and manage the spread of environmental ARGs.

Summary

Antibiotic resistance is ancient and prevalent in natural ecosystems and evolved long before the utilization of synthetic antibiotics started, but factors influencing the large-scale distribution patterns of natural antibiotic resistance genes (ARGs) remain largely unknown. Here, a large-scale investigation over 4000 km was performed to profile soil ARGs, plant communities, and bacterial communities from 300 quadrats across five forest biomes with minimal human impact. We detected diverse and abundant ARGs in forests, including over 160 genes conferring resistance to eight major categories of antibiotics. The diversity of ARGs was strongly and positively correlated with the diversity of bacteria, herbaceous plants, and mobile genetic elements (MGEs). The ARG composition was strongly correlated with the taxonomic structure of bacteria and herbs. Consistent with this strong correlation, structural equation modelling demonstrated that the positive effects of bacterial and herb communities on ARG patterns were maintained even when simultaneously accounting for multiple drivers (climate, spatial predictors and edaphic factors). These findings suggest a paradigm that the interactions between aboveground and belowground communities shape the large-scale distribution of soil resistomes, providing new knowledge for tackling the emerging environmental antibiotic resistance.

This article is protected by copyright. All rights reserved.

Author Manuscript

Introduction

Emerging prevalence and dissemination of antibiotic resistance genes (ARGs, collectively known as "resistome") represents a major threat to public health in the 21st century (WHO, 2014; Berendonk *et al.*, 2015). A growing recognition of an environmental component of resistome pools has prompted calls for investigation into the significance of antibiotic resistance in natural settings (Ashbolt *et al.*, 2013). It has been reported that some environmental ARGs have perfect sequence identity to those in human pathogens (Forsberg *et al.*, 2012) and human feces (Nesme *et al.*, 2014), indicative of the possible gene flow between environmental and clinical resistomes (Wright, 2010). Environmental ARGs can be disseminated via physical forces such as leaching to water, atmospheric transport/deposition of dust, dissipation by animals or birds, and delivery of ARGs-carrying products (Allen *et al.*, 2010; McEachran *et al.*, 2015~~),~~). ~~and~~ They are also subject to horizontal gene transfer mediated by mobile genetic elements (MGEs) such as integrons, transposons, and plasmids (Gillings *et al.*, 2015). Therefore, environmental ARGs have been recognized as contaminants of global concern (Pruden *et al.*, 2006), and may constitute an important contributor to clinical antibiotic resistance if acquired by human commensals and pathogens.

Soil is one of the largest environmental reservoir comprising approximately 30% of known ARGs in public repositories (Forsberg *et al.*, 2012; Nesme *et al.*, 2014), and is one of the most complex ecosystems in terms of ecological niches and biodiversity (Nesme and Simonet, 2014). To date, a large body of studies have documented the prevalence of ARGs in environments under anthropogenic perturbations, such as soils with manure amendments (Zhu *et al.*, 2013; Zhang *et al.*, 2017), contaminated with pharmaceutical residues (Zhu *et al.*, 2017) or heavy metals (Hu *et al.*, 2016b, 2017), and irrigated with reclaimed water (Wang *et al.*, 2014; Han *et al.*, 2016). We have, however, limited knowledge of the background levels of antibiotic resistance and the distribution of ARGs in natural settings with little or no human disturbance. In fact, most antibiotics are natural compounds produced by different taxa of bacteria, fungi~~,~~ or plants (D'Costa *et al.*, 2006; Chen *et al.*, 2013), and the presence of natural antibiotics may have exerted consistent selection pressure on microorganisms since long before the introduction of human-made antibiotics (Allen *et al.*, 2010; Wright and Poinar, 2012). Therefore, it is not surprising that ARGs were identified from ancient permafrost sediment (D'Costa *et al.*, 2011), Alaskan permafrost soil (Allen *et al.*, 2009), glaciers (Segawa *et al.*, 2013), Antarctic soils (Wang *et al.*, 2016a), underground cave (Pawlowski *et al.*, 2016), and deep ocean sediment (Chen *et al.*, 2013). However, we lack a general

Wiley-Blackwell and Society for Applied Microbiology

This article is protected by copyright. All rights reserved.

understanding of the key determinants shaping the diversity and composition of ARGs in natural ecosystems. These knowledge gaps have precluded the inclusion of natural resistomes in management strategies and risk assessment frameworks (Ashbolt *et al.*, 2013).

Forest ecosystems occupy approximately 30% of Earth's land surface (Bonan, 2008) representing important habitats for trees, herbs, and microbes (Baldrian, 2017). Plants can influence the structure of soil microbial communities through plant litters and root exudation (Prober *et al.*, 2015; Wang *et al.*, 2016b), which might indirectly shape soil resistome composition via affecting the taxonomic and phylogenetic structure of ARGs-bearing microbes (Forsberg *et al.*, 2014). On the other hand, forest soils contain antibiotics produced by soil microbes, and potentially toxic aromatic compounds derived from degradation processes, as well as specialized antibiotics (e.g., salicylic acid, phytoalexins~~,~~ and flavonoids) and signalling molecules from plant exudates (Alonso *et al.*, 2001; Yergeau *et al.*, 2014), which are reported to increase expression of antibiotic resistance in the plant rhizosphere (Yergeau *et al.*, 2014). In addition, forest soils are characterized by nutrient scarcity, and represent a highly competitive environment for microbes to survive, interact, and compete for space and nutrients, hence it is expected to observe a wide occurrence of ARGs-bearing microbes (producing low concentrations of antibiotics as signalling molecules to modulate the interactions within microbial communities; or producing high concentrations of antibiotics as a weapon to counter other microbial taxa probably in some micro-environments ~~in the rhizosphere~~) in forest soils (Martinez, 2008). Despite forest soils being hypothesized as an important reservoir of ARGs, the large-scale distribution patterns of ARGs and how their compositions are regulated by environmental factors and plant and bacterial communities remain less understood.

In this study, we used high-throughput quantitative PCR (HT-qPCR) techniques to profile 285 ARG subtypes encoding resistance to eight major classes of antibiotics from 300 soil samples collected across five forest biomes in China. This large-scale investigation through establishing 300 quadrats over 4000 km from boreal forest to tropic rainforest allows simultaneous documentation of plant communities (including trees and ~~understory~~ herbs), soil microorganisms, soil resistome, and a wide range of climatic variables and soil properties. We hypothesized that (i) the natural forest ecosystem ~~can be~~is a significant reservoir of ARGs, due to the antimicrobial molecules produced by the soil microorganism and plant rhizosphere; (ii) the changes in ARG patterns are closely associated with the changes in the bacterial communities, which is assumed as an important determinant of ARGs in natural settings; and

Wiley-Blackwell and Society for Applied Microbiology
This article is protected by copyright. All rights reserved.

Author Manuscript

(iii) the impacts of plant communities on the diversity and structure of soil microbial communities will ultimately have consequences on the soil ARG profiles. Our study provides important information on the background level of ARGs and the key drivers shaping the large-scale~~diversity and~~ biogeographic distribution of antibiotic resistance in natural forest biomes ~~over large scales~~. This knowledge is critical in understanding the origins and evolution of natural soil resistome, and may help to design effective framework to predict and manage the spread of environmental ARGs.

Results

Diversity and relative abundance of ARGs and MGEs

The HT-qPCR array detected a total of 160 ARGs and 9 MGEs (including 2 integrase genes and 7 transposase genes) in the 300 soil samples collected from five forest biomes (Fig. 1a). The average number of detected ARGs considerably changed from 95.5 (± 20.1) in temperate mixed coniferous forest to 73.8 (± 16.1) in subtropical forest (Fig. 1b). The diversity of ARGs in temperate mixed coniferous forest and temperate deciduous forest (94.2 ± 20.7) was significantly higher than that in boreal forest (85.7 ± 18.2) and subtropical forest ($P < 0.05$). The numbers of detected MGEs showed a similar pattern to that of ARGs (Fig. 1c), with the lowest found in subtropical forest (3.3 ± 0.8) and the highest in temperate mixed coniferous forest (4.2 ± 1.3). The diversity of MGEs in temperate mixed coniferous forest and temperate deciduous forest (4.1 ± 0.8) was significantly higher than that in subtropical forest ($P < 0.05$). The detected ARGs could potentially confer resistance to eight major classes of antibiotics, with the most frequently detected ARGs being multidrug and β-lactam resistance (Fig. S1a), accounting for 22.4% and 21.8%, respectively, of the total ARGs. Other frequently encountered ARG types included aminoglycoside (13.5%), MLSB (14.1%), tetracycline (11.2%), and vancomycin (8.2%) resistance, depending on the representation of primer sets on the array. The detected ARGs included all major resistance mechanisms: antibiotic inactivation (42.9%), multidrug efflux pumps (33.5%), and cellular protection (20.0%) (Fig. S1b). Venn diagram showed that a total of 59 ARGs were shared among the five forest ecosystems (Fig. S2a), and these commonly shared ARGs were dominated by the multidrug and β-lactam resistance genes (Fig. S2b).

We further calculated the relative abundance of ARGs and MGEs by normalizing against the bacterial 16S rRNA gene. Similar to the pattern of the ARG ~~s~~ diversity, the relative abundance of ARGs highly varied across forest biomes, with the most abundant

Wiley-Blackwell and Society for Applied Microbiology

This article is protected by copyright. All rights reserved.

Author Manuscript

ARGs observed in temperate mixed coniferous forest, and the lowest in subtropical forest ($P < 0.05$) (Fig. 1d). The relative abundance of ARGs was significantly higher in boreal and temperate (mixed coniferous and deciduous) forests than subtropical and tropical forests ($P < 0.05$). The relative abundance of MGEs significantly shifted across forest biomes (Fig. 1e), and was significantly lower in subtropical and tropical forest than in other three forest types ($P < 0.05$). Spearman's correlation analysis revealed significantly positive relationships between the relative abundances of total MGEs with total ARGs and individual ARG types (Table S3). The relative abundances of specific MGEs, such as *intI*, *intI1*, and *tnpA-02* genes were also significantly and positively correlated with those of total ARGs and individual ARG types.

Gene-specific analyses of ARGs and MGEs subtypes

The patterns of the relative abundance of ARG and MGE subtypes across forest biomes were explored in heat maps based on the HT-qPCR array results (Fig. 2a). Soil samples (columns in the heat map) could be separated into several clusters: samples from subtropical and tropical forest tended to cluster together; and samples from the other three forest biomes were distributed into three separate major clusters with each containing samples from all three forest biomes. This finding was further corroborated by non-metric multidimensional scaling (NMDS) ordinations based on the Bray-Curtis distances of ARGs (Fig. 2b), which demonstrated that subtropical and tropical forests harboured significantly different ARG compositions from temperate and boreal forests (PerMANOVA, $P < 0.05$). The profiles of individual ARG subtypes (rows in the heat map) could be classified into three distinct patterns: (A) These ARG subtypes are prevalent with high relative abundance in the majority of forest soils and tended to be less abundant in subtropical and tropical forest; (B) Compared to pattern A, these ARGs are less abundant in forest soils, and tend to be decreased in subtropical and tropical forests; (C) These ARG subtypes are in low abundance or below the detection limit in most forest soils.

Among the 160 ARGs and 9 MGEs detected, 32 ARGs and 2 MGEs were found in more than 98% of the forest soil samples (Table S4). These ubiquitous genes contained 1 aminoglycoside resistance gene, 9 β-lactam resistance genes, 4 macrolide-lincosamide-streptogramin B (MLSB) resistance genes, 5 tetracycline resistance genes, 5 vancomycin resistance genes, and 8 multidrug resistance genes. Some widely-distributed ARG and MGE subtypes included *intl1*, *aacC*, *mexF*, *blaSFO*, *cphA-01*, *fox5*, *mphA-01*, *oleC*, *tetR-02*, *vanC-03*, *acrA-05*, *oprD*, and *oprl* genes. The *mexF* gene (multidrug resistance) was the most

Wiley-Blackwell and Society for Applied Microbiology

This article is protected by copyright. All rights reserved.

Author Manuscript

abundant resistance gene, up to 20.2% as abundant as the 16S rRNA gene. The *intl1* gene was the most abundant MGE, ranging from 0.02-14.4%.

Soil bacterial, herb and tree communities

Bacterial communities were determined by targeting the 16S rRNA gene on the Miseq platform, which yielded a total of 15,421,630 high-quality sequences from the 300 soils, with an average of 51,405 sequences per sample. These sequences could be clustered to 40,399 OTUs at 97% sequence identity, with the number of OTUs per sample ranging from 918 to 7110. Bacterial communities were dominated by Proteobacteria, Acidobacteria, Actionobacteria, Verrucomicrobia, Planctomycetes, and Chloroflexi at the phylum level (Fig. S3a).

Herb communities were comprised of three major phyla Magnoliophyta, Pteridophyta, and Lycopodiophyta, and their relative abundance highly varied across forest biomes, with Pteridophyta predominant in subtropical forest and Lycopodiophyta found only in tropical forest (Fig. S3b). Tree communities showed substantial differences across forest biomes (Fig. S3c). For example, the Ericales and Fabales families belonging to the Angiospermae phylum were only found in subtropical and tropical forests, while the Rosales family was more dominant in boreal and temperate mixed coniferous forests.

Soil bacterial diversity was calculated for each sample using the numbers of OTUs at 97% sequence similarity. The changes in the diversity of bacteria and herbs across forest biomes were very similar, with the lowest level of diversity recorded in subtropical forest (Figs. S3d and S3e), which mimicked the diversity patterns of ARGs (Fig. 1b). Tree diversity significantly increased from boreal forest (2.08 ± 1.07) to tropical forest (34.0 ± 12.2) (Fig. S3f).

Relationships between ARGs with MGEs, plants, bacteria, and environmental factors

Regression analysis was conducted to explore the relationships between the diversity of ARGs with the diversity of MGEs, bacteria and plants, soil properties and climatic factors. The ARG~~s~~ diversity had strongly positive linear relationships with the diversity of bacteria (Fig. 3a), herbs (Fig. 3b), and MGEs (Fig. 3c) (all $P < 0.001$). Soil pH was the most important edaphic factor that was significantly correlated with the ARG~~s~~ diversity, with the highest diversity predicted around pH 7 (Fig. 3d). The relationships between the ARG~~s~~ diversity with tree diversity, spatial factors, soil properties, and climatic factors were not significant, and

Wiley-Blackwell and Society for Applied Microbiology

This article is protected by copyright. All rights reserved.

Author Manuscript

only the correlations with annual mean precipitation (AMP) and available potassium were statistically significant (Fig. S4).

Mantel tests were further performed to examine the correlations between the ARG~~s~~ composition with the compositions of MGEs, bacteria and plants as well as environmental factors (Table S5). Similarly, we found that MGEs, bacteria, herbs, and soil pH were the most important factors that had significantly positive correlations with the ARG~~s~~ composition. Soil properties including soil organic carbon (SOC), total nitrogen (TN), and moisture content, climatic factors including annual mean temperature (AMT) and AMP, and spatial factor including latitude were also drivers that may significantly influence the patterns of ARGs in forests. This result was further supported by the CCA test which revealed that soil pH, AMT, AMP, latitude, and moisture content were important drivers shaping the ARG~~s~~ patterns (Fig. S5).

Co-occurrence patterns of ARGs, MGEs, bacteria, and herbs

The links between ARGs, MGEs, and bacterial and herb communities were further explored using the local similarity analyses, which were visualized as association networks based on strong ($\rho > 0.8$) and significant ($P < 0.05$) correlations. Significant co-occurring correlations were observed in the ARGs-MGEs-bacteria-herb networks, suggesting their strong interactions in forests (Fig. 4a). Interestingly, the *intI1* and *intI* genes (belonging to MGEs) had the most intensive connections with ARGs (Fig. 4a), which can potentially confer resistance to multiple classes of antibiotics (Fig. 4b). The co-occurrence pattern between ARGs and MGEs was further assessed in a network including only ARG/MGE subtypes (Fig. S6). The entire ARGs-MGEs network was clearly separated into four clusters, with the Module III composed of the *intI1* and *intI* genes that had positive relationships with all the major classes of ARGs (Fig. S6). The Gleicheniaceae, Poaceae, Zingberaceae, Cyperaceae, Apiaceae, and Woodsiaceae families were the key herb taxa that had significant correlations with bacteria and ARGs (Fig. 4).

The *blaSFO* (β-lactam resistance), *cmlA1-02* (FCA resistance), *oleC* (MLSB resistance), and *emrD* (sulfonamide resistance) genes were highly connected with a variety of bacterial taxa (Fig. 4a), which were mainly affiliated with Actinobacteria, Proteobacteria, Chloroflexi, Acidobacteria (Fig. S7), implying that these bacterial phyla are possible ARG hosts. Single bacterial taxa significantly correlated with multiple ARG subtypes, suggestive of a bacterial host to harbour multiple ARGs (Fig. S7). Procrustes analysis further demonstrated significant

Wiley-Blackwell and Society for Applied Microbiology
This article is protected by copyright. All rights reserved.

Author Manuscript

correlations between the bacterial community compositions and the ARG~~s~~ contents with highly significant goodness-of-fit measures (Fig. 2c).

Structural equation modelling accounting for multiple drivers

We used Random Forest modelling to identify the major factors predicting the ARG~~s~~ diversity, and structural equation model (SEM) to test whether the relationships between the diversity of plants, bacteria and ARGs is maintained when accounting for multiple drivers simultaneously. Random Forest analysis indicated that microbial diversity was more important than other environmental factors in predicting the ARG~~s~~ patterns (Fig. S8). Plant diversity (including tree and herb diversity) was as important as most environmental drivers. Based on the known relationships among the drivers of ARG patterns, ~~We~~ we constructed an *a priori* model (Fig. S9) to determine the direct and indirect effects of soil properties (including soil pH), climatic conditions (including AMT), biome types, the diversity of bacteria, trees and herbs, and MGEs, on ARG patterns ~~(Fig. S9)~~. Our final SEM explained 67% of the variance in the ARG~~s~~ diversity across forest biomes (Fig. 5a). Consistent with the correlation analyses (Fig. 3), microbial diversity and herb diversity showed positive effects on the ARG~~s~~ diversity even after accounting for the simultaneous direct and indirect effects of multiple variables (Fig. 5b). No significant and direct effects of plant diversity on ARG patterns were found, but ~~H~~herb diversity had indirect effects on ARG~~s~~ patterns through its positive impacts on microbial diversity (Fig. 5a). Microbial diversity exerted direct positive effects on ARG~~s~~ patterns, and also indirectly affected ARGs through its positive effects on MGEs (Fig. 5a). Not surprisingly, biome types, climate, spatial factors and edaphic factors showed strong direct and indirect effects on ARGs as indicated by the standardized total effects from SEM (Fig. 5b), assuming that these factors have been identified to be critical in shaping the distribution pattern of soil microbes. Collectively, these results demonstrate that microbial and herb diversity play critical roles in shaping the ARG~~s~~ patterns, which are also directly and indirectly affected by other drivers in forests.

Discussion

Ubiquitous ARGs in forest biomes

A key finding of this study is that more than 160 ARGs ~~subtypes~~ (encoding resistance to eight major classes of antibiotics) were detected from the 300 forest soils with minimal anthropogenic disturbance. In line with previous screening of ARGs in polluted estuaries (Zhu *et al.*, 2017), urban environments (Xiang *et al.*, 2018), agricultural soils (Hu *et al.*,

Wiley-Blackwell and Society for Applied Microbiology

This article is protected by copyright. All rights reserved.

Author Manuscript

Author Manuscript

2016a, 2016b; Chen *et al.*, 2017) and Chinese swine farms (Zhu *et al.*, 2013), genes conferring resistance to multidrug, β-lactam, aminoglycoside, MLSB, tetracycline, and vancomycin were the most dominant ARG types in these forest soils. Our findings contribute to the growing body of evidence that ARGs are widespread in relatively pristine habitats with less contact with commercial sources of antibiotics (Chen *et al.*, 2013; Segawa *et al.*, 2013; Pawlowski *et al.*, 2016), and highlight that the natural forest ecosystem is a significant reservoir of ARGs and needs to be considered in the risk assessment frameworks of antibiotic resistance. Comparison of the background ARG profiles between the forest ecosystems and some human-impacted soils by using the same HT-qPCR array suggested that the diversity and relative abundance of ARGs in copper-contaminated soils (Hu *et al.*, 2016b), nickel-contaminated soils (Hu *et al.*, 2017), animal manure-amended soils (Zhang *et al.*, 2017), and antibiotics-treated soils (Zhang *et al.*, 2017) are significantly higher than those in forest biomes (Fig. S10). Given that forests are significantly threatened by new anthropogenic pressures such as urbanization, intensive management practices, and land-use conversion to farmland ~~(Gauthier *et al.*, 2015)~~, it is assumed that these pre-existing ARGs might emerge with clinical importance in future if they are transferred to a human pathogen (Martinez *et al.*, 2015).

The widespread occurrence of ARGs in natural forests might be attributed to the inherent presence of antibiotics in soil (Wright, 2010; Nesme *et al.*, 2014), though it is very difficult to measure the natural concentrations of antibiotics (Allen *et al.*, 2010). Most of currently known antibiotics are small molecules produced by certain microorganisms in natural settings (D'Costa *et al.*, 2006). Bacteria can use these antibiotics as weapons (at high concentrations) to compete for limited resources (D'Costa *et al.*, 2006) or as signalling molecules (at low concentrations) to communicate with other microbes (Martinez, 2008). Furthermore, forest soil contains some toxic aromatic compounds derived from degradation processes or specialized antibiotics from plant exudates (Alonso *et al.*, 2001; Yergeau *et al.*, 2014), which may also contribute to the development of antibiotic resistance. Apart from the long-term residence of indigenous ARGs, other unknown selection pressures (i.e. anthropogenic ~~selection pressures~~disturbance) can also contribute to the level of ARGs in forests. Although the 300 soil samples in this study were collected in regions that are distant from human industrialized activities and supposedly devoid of anthropogenic disturbance, we cannot completely exclude the possible mechanism of atmospheric deposition of ARGs-

Wiley-Blackwell and Society for Applied Microbiology
This article is protected by copyright. All rights reserved.

Author Manuscript

bearing microbes that originate from ARGs-rich environments, and the long-distance transport of ARGs by birds and animals (Wang *et al.*, 2016a).

Some ARG subtypes were ubiquitously detected across the five forest ecosystems, such as *aacC*, *mexF*, *blaSFO*, *cphA*, *fox5*, *mphA-01*, *oleC*, *tetR-02*, *vanC-03*, *acrA-05*, *oprD*, and *oprl* genes, indicating that natural forest soils are a common reservoir for these genes. The *mexF* gene ~~(multidrug resistance)~~ was the most abundant resistance gene, perhaps driven by multiple antibiotic and toxic compounds in forests. The *mexF* gene has been reported in various environments, such as soils, sediments, rivers~~,~~ and drinking water (De Souza *et al.*, 2006). The *cphA* gene (β-lactam resistance) has been found to be intrinsic in environmental isolates of *A. jandaei* and *Aeromonas hydrophila* (Balsalobre *et al.*, 2009). The *mexF* and *oprD* multidrug resistance genes were proposed to naturally occur in water (Stedtfeld *et al.*, 2016). The *vanC* gene can encode resistance to vancomycin, which is considered as one of the last-resort defence against organisms like *Streptococcus pneumonia* and *Enterococcus* (Gilmore and Hoch, 1999), and this gene is ancient in natural environments (D'Costa *et al.*, 2011). These findings suggest that natural forests are sharing many ARGs within and across biomes, with a possibility of gene flow and exchange among different ecosystems. Therefore, we argue that increasing attention should be paid to the importance of natural resistome and its potential transmission routes to human commensals and pathogens.

Although the HT-qPCR array used in this study can simultaneously profile 285 ARGs and 10 MGEs of interest in a high-throughput manner (Looft *et al.*, 2012; Su *et al.*, 2015), the need for pre-selected primer sets limits its use to known genes or to genes with high ~~homology~~ nucleotide similarity to known ones (Karkman *et al.*, 2018). For example, the relative dominance of the ARG classes or resistance mechanisms detected using the HT-qPCR technique mostly relies on the representation of primer sets on the array. Metagenomics, especially deep sequencing, is not restricted to prior knowledge of few selected genes, and is able to capture the whole resistome in the environment (Li *et al.*, 2015). However, the annotation of ARGs is still dependent on sets of known verified genes in ARG databases (Gibson *et al.*, 2015). Functional metagenomics can overcome the limitations of HT-qPCR and metagenomic sequencing in detecting mostly known ARGs, but the difficulty in the cloning and expression of environmental DNA in a laboratory host is the main disadvantage of this technique (Karkman *et al.*, 2018). Despite the intrinsic limitations of these methodologies, a combination of HT-qPCR, metagenomics and functional metagenomics is still powerful in the determination of antibiotic resistance in all its dimensions.

Wiley-Blackwell and Society for Applied Microbiology

This article is protected by copyright. All rights reserved.

Potential for horizontal gene transfer of ARGs in forest soils

The abundant and diverse ARGs detected in forest soils, together with abundant MGEs, offer a likelihood for ARGs being horizontally transferred to other bacteria and pathogens. Firstly, the class 1 integron-integrase gene *intI1* and the transposon-transposase gene *tnpA* were detected in every forest soil sample. Class 1 integrons are widely distributed in over 70% of Gram-negative bacteria of clinical importance (Stalder *et al.*, 2012), and have been repeatedly reported to be an important proxy for acquisition and dissemination of ARGs in the environments (Gillings *et al.*, 2015). Integrons have a site recombination system which could capture and express gene cassettes, and they are often located on plasmids and transposons which further facilitate their mobility potential (Heuer *et al.*, 2012). Secondly, the HT-qPCR detected two integrase genes and seven transposase genes, with their total diversity and relative abundances positively correlated with those of ARGs. In particular, the relative abundances of *intI*, *intI1* and *tnpA-02* genes had significantly positive relationships with the relative abundances of all individual classes of ARGs (Table S3), and the compositions of ARGs and MGEs were also closely correlated as revealed by the Mantel test. These results suggested that there is a potential risk of HGT mediated by MGEs for ARGs in forest soils. However, it was found that the diversity of both ARGs and MGEs was significantly correlated with the diversity of bacteria, which means that the significantly positive relationships observed between ARGs and MGEs might be indirect and would not necessarily lead to high HGT potential.

Of particular interest, we found that the *intI1* and *intI* genes had the most intensive connections with ARG subtypes (Fig. 4a), and co-occurred with multiple ARGs conferring resistance to all the major classes of antibiotics (Fig. S6). Some of these ARG subtypes have been reported to be located within integron gene cassettes, such as the *aac* and *aad* families of ARGs (Zhu *et al.*, 2017). Indeed, the *aac(6')-Ib*, *aadA5*, and *aadA9* genes in the same cluster with the *intI1* and *intI* genes are known components of integron gene cassettes (Partridge *et al.*, 2009). The co-occurring ARGs and MGEs suggest that these ARGs might be genetically linked together in integrons, transposons and microbial genomes conferring resistance to multiple antibiotics, and likely subject to a synchronous dispersal during HGT. However, it should be noted that although there are strong correlations between the diversity, abundance and compositions of ARGs and MGEs, it remains unknown whether HGT is actually happening in forest ecosystems and would disperse environmental ARGs.

Wiley-Blackwell and Society for Applied Microbiology

This article is protected by copyright. All rights reserved.

Author Manuscript

Author Manuscript

Despite of the genetic potential of HGT, the actual frequencies of HGT for soil ARGs are also influenced by a range of abiotic factors (e.g. soil pH, moisture content, temperature, and oxygen levels) and biotic factors (e.g. commensal, antagonistic and mutualistic relationships between soil organisms) in the environment (Aminov, 2011). It remains largely unknown regarding the frequency of HGT in natural settings with relatively lower selection pressure compared with contaminated environments or agricultural soils (Johnson *et al.*, 2016; Hu *et al.*, 2017). In fact, it is assumed that the rate of HGT is very low in natural soils, where ARGs content and phylogenetic compositions are strongly associated (Forsberg *et al.*, 2014) suggesting that HGT is not frequent to effectively decouple resistomes from phylogeny. The strong correlation between the ARG content and the bacterial community composition was also observed (Fig. 5 and Table S5), suggestive of a low rate of HGT in forests. However, ARGs are not linked to the phylogenetic composition in animal agriculture (Johnson *et al.*, 2016) and heavy metal-contaminated soils (Hu *et al.*, 2017), indicating that genetic potential for HGT might be enhanced when natural soils are subjected to anthropogenic activities.

Bacterial and herb communities as important determinants of ARGs in forests

The diverse and abundance ARGs detected in natural forests prompted us to identify the major factors shaping their distribution patterns, enabling prediction of the emergence and evolution of natural soil resistome in future. By compiling all the data together, our SEM demonstrated that the positive direct effects of bacterial communities and the positive indirect effects of herbaceous plants on ARG patterns were maintained even when simultaneously accounting for multiple drivers (climate, spatial predictors and edaphic factors). All these findings led to the conclusion that bacterial and herb communities are important predictors of ARG contents in natural forest biomes.

Previous studies have reported that the microbial phylogenetic and taxonomic structure is an important determinant of ARG compositions (Forsberg *et al.*, 2014; Hu *et al.*, 2016b; Wang *et al.*, 2016a), because soil microbes are not only the major sources of antibiotic compounds or their derivatives (Allen *et al.*, 2010; Wright and Poinar, 2012) but are also hosts to many ARGs. For example, Actinobacteria are well-known groups of bacteria producing antibiotics; Proteobacteria and Actinobacteria are among the most prevalent predicted hosts of multi-resistant ARGs in metagenomics studies of soils (D'Costa *et al.*, 2006; Forsberg *et al.*, 2014), which is supported by the network analysis in our study (Fig. S7). The strong impact of microbial community structure on ARG contents might be also

Wiley-Blackwell and Society for Applied Microbiology

This article is protected by copyright. All rights reserved.

attributed to low rate of HGT which have not decoupled ARG contents and bacterial phylogeny.

Interestingly, we found that the herb community can indirectly affect ARG contents via its positive impacts on the microbial diversity (Fig. 5a), which might be due to the long-evolved intimate relationships between plants and microbes (Bai *et al.*, 2015). Although plants can exude specialized antibiotic compounds, which can lead to shifts of ARG contents (Yergeau *et al.*, 2014), we did not observe direct impacts of plant communities (including trees and herbs) on ARG contents in this study. Because our sampling strategy had tried to collect bulk soils between trees, the selective pressure imposed by the plant exudation of antimicrobials in the plant rhizosphere might have been overlooked. Although trees are the dominant primary producers responsible for >90% of the forest primary production and provide the majority of soil organic materials through root exudates and litters (Baldrian, 2017), understory herbs, rather than trees, were found to be a more important determinant of ARG patterns. Previous studies have also shown that herbs, rather than trees, are better predictors of soil prokaryotic diversity in forest ecosystems (Wang *et al.*, 2016b) and in grasslands (Prober *et al.*, 2015). The possible explanation might be that soil microbes in the topsoil are in more contact with herbs, because of the spatial separation between the roots of herbs and trees (Scholes and Archer, 1997); and herb residues are more easily decomposed by free-living soil microorganisms, compared to tree residues and litter (Hattenschwiler *et al.*, 2005). The impacts of herbaceous plants on the diversity and structure of soil microbial communities will ultimately have consequences on the soil ARG profiles.

Despite of the significant impacts of bacterial and herb communities on ARG contents, a range of other drivers including climate (AMT, annual mean temperature; MTCM, mean temperature of coldest month; and PDM, precipitation of driest month), edaphic factors (mainly soil pH) and geographic locations are also found to influence antibiotic resistome in natural forests (Fig. 5). Soil pH and climate have been recognized as the main driver of soil microbial diversity patterns across different ecosystems (Fierer and Jackson, 2006), implying that the factors driving the distribution of bacteria could underlie the diversity and distribution of ARGs. In addition, soil fungi are also important microbial communities in various forest ecosystems (Baldrian, 2017), and can produce antibiotics and compete with other soil microorganisms for space and nutrients (Chen *et al.*, 2013). Therefore, we argue that global environmental changes such as climate warming, extreme weather events, nitrogen deposition that have strong impacts on the plant, bacterial and fungal communities will also have

Wiley-Blackwell and Society for Applied Microbiology
This article is protected by copyright. All rights reserved.

Author Manuscript

consequences on soil resistome. Characterizing relationships between ARGs, bacteria, fungi, and plants in various ecosystems across broad environmental gradients will be critical to predicting the patterns of environmental resistome in a changing world.

Conclusions

To our knowledge, this is the first large-scale profiling of a broad spectrum of ARGs in natural forest biomes, providing an overview of how the interactions of aboveground plants and belowground microbes as well as soil properties and climate shape the distribution patterns of antibiotic resistance. Our results suggest that natural forest soils are a significant source of ARGs, and their large-scale distribution is regulated by the diversity of bacteria and herbaceous plants even when simultaneously accounting for multiple environmental drivers. These findings contribute to a growing understanding of the magnitude of environmental ARGs, and have important implications for future prediction of the evolution, dissemination and persistence of antibiotic resistance over broad spatial scales. Given the transmission potential of environmental ARGs to human commensals and pathogens, an improved knowledge of the factors maintaining environmental resistome will help the development of ~~both the risk assessment and~~ effective mitigation options to minimise the spread of antibiotic resistance.

Experimental procedures

Plant inventory and soil sampling campaign

We carried out a large-scale investigation in 2012 across a latitudinal gradient ranging from N18°15' to N53°18' over 4000 km in China (Fig. 1a). The selected sampling sites span over five major forest biomes from north to south: boreal forest, temperature mixed coniferous forest, temperate deciduous forest, subtropical (evergreen broadleaf) forest, and tropical (rainforest) forest. The sampling sites were selected from National Nature Reserves that are distant from anthropogenic disturbance. Our sampling strategy provides a wide variation in soil properties with soil pH from 4.03 to 8.57, total nitrogen (TN) from 0.05 to 1.16%, and soil organic carbon (SOC) from 0.39 to 19.21%. The mean annual temperature (MAT) ranges from -5.6°C to 25.9°C and the mean annual precipitation (MAP) varies from 452 to 2085 mm. The detailed information about the sampling sites and climate is shown in Table S1.

We established 60 independent large quadrats (20 × 20 m^2) in each forest biome, resulting in a total of 300 quadrats for the five forest biomes. Within each quadrat, we

Wiley-Blackwell and Society for Applied Microbiology
This article is protected by copyright. All rights reserved.

Author Manuscript

recorded all trees with diameter at breast height ≥2.5 cm and classified them into the species level. Three small plots (1 × 1 m^2) were randomly established within each quadrat for the census of herbaceous plants (Fig. 1a). Taxonomic classification of trees and herbs was performed based on the Angiosperm Phylogeny Group III System (Chase and Reveal, 2009). Soil samples were collected from each quadrat by thoroughly mixing 15 random soil cores (5 cm in diameter and 10 cm in depth), sieved through 1.0 mm mesh and pooled into a composite sample for each quadrat. A total of 300 soil samples were sealed in plastic bags and shipped on ice to the laboratory. A portion of soils was used for physicochemical analysis, and another portion was freeze-dried and stored at -80°C until DNA extraction.

Soil physicochemical characterization

Soil moisture content was measured by oven-drying fresh soil samples at 105°C. Soil pH was determined in soil suspension with a soil to water ratio of 2.5 using a Delta 320 pH-meter (Mettler-Toledo Instruments, Columbus, OH, USA). TN was measured on a Vario EL III Element Analyser (Elementar, Hanau, Germany). SOC was determined on the Element Analyser (Elementar) after treatment with H_2SO_4 to remove carbonate. Soil inorganic nitrogen (including NH_4^+-N and NO_3^--N) was extracted from soils with 1 M KCl and measured using a SAN++ Continuous Flow Analyser (Skalar, Breda, Netherlands). Soil available phosphorous (AP) was extracted from soils by 0.5 M $NaHCO_3$ solution and measured on the Continuous Flow Analyser (Skalar). Soil available potassium (AK) was determined using Inductively Coupled Plasma-Optical Emission Spectroscopy (ICP-OES, Vista axial, Varian). Soil particle size was determined using a Laser Particle Size Analyser (Malvern Instruments, Malvern, UK).

DNA extraction and HT-qPCR analysis

The MoBio PowerSoil DNA Isolation Kit (MoBio Laboratories, Carlsbad, CA, USA) was used to extract soil DNA from 0.25 g of soil as per the manufacturer's recommendations. DNA quality was evaluated using IMPI EN P-330 NanoPhotometer UV/VIS spectrophotometer (IMPLEN, Munich, Germany). A HT-qPCR array was performed to determine the diversity and abundance of ARGs on the Wafergen SmartChip Real-Time PCR System (Fremont, CA, USA) as described previously (Zhu *et al.*, 2013; Su *et al.*, 2015). The Wafergen system is a high-throughput platform with the capability to run 5184 reactions each with a volume of 100 nl using the SensiMix SYBR No-ROX reagent (Bioline, London, UK). The HT-qPCR array contained a total of 296 primer sets (Looft *et al.*, 2012; Su *et al.*, 2015),

Wiley-Blackwell and Society for Applied Microbiology
This article is protected by copyright. All rights reserved.

Author Manuscript

including 285 primer sets targeting eight major classes of ARGs, 10 primer sets targeting MGEs, and one 16S rRNA gene as the internal control (Table S2). The specificity of all primers was verified through amplicon sequencing, BLAST search and amplification efficiency check (Looft *et al.*, 2012; Johnson *et al.*, 2016).

Thermal-cycling conditions were as follows: 95°C for 10 min, followed by 40 cycles of 95°C for 30 s and 60°C for 30 s. All HT-qPCR reactions were performed in three technical replicates. ARGs or MGEs were regarded as positive detection according to the following criteria: (i) a threshold cycle value (C_T) of 31 was used as the detection limit to differentiate between positive amplification and primer-dimers (Su *et al.*, 2015); (ii) samples with three technical replicates all above the detection limit were regarded as positive quantification; (iii) amplicons with multiple melting curves were removed from the analysis; and (iv) reactions with efficiency beyond the range 1.7-2.3 and R^2 values below 0.99 were discarded. The $2^{-\Delta CT}$ method where $\Delta C_T = (C_{T\ \text{detected ARGs}} - C_{T\ 16S\ \text{rRNA gene}})$ was used to calculate the relative abundances of ARGs and MGEs compared to the 16S rRNA gene abundance based on a comparative C_T method (Schmittgen and Livak, 2008).

Illumina Miseq sequencing and data processing

The V4~~–V5~~ region of the bacterial 16S rRNA gene was amplified using the primer set 515F/806R (Bates *et al.*, 2011) linked with the Illumina adaptor and a 12-bp barcode sequence. PCR amplification was performed in 50 μl mixtures containing 25 μl Premix Taq (TaKaRa, Dalian, China), 1 μl each primer (10 μM), and 3 μl template DNA (~10 ng). The barcoded PCR products were purified using Wizard SV Gel and PCR Clean-up System (Promega, Madison, WI, USA), quantified, pooled into one mixture, and sequenced using 2× 150-bp paired-end strategy on an Illumina MiSeq sequencer. Raw and paired-end sequences were quality filtered, assembled using the Fast Length Adjustment of Short Reads (FLASH) (Magoè and Salzberg, 2011), de-multiplexed, and assigned to individual samples using Quantitative Insights Into Microbial Ecology (QIIME) analysis (Caporaso *et al.*, 2010). Assembled sequences were binned into operational taxonomic units (OTUs) at the 97% sequence identity using a chimera filtering approach UPARSE (Edgar, 2013). An even number of sequences per sample (45,070 reads) was randomly selected from each sample to compensate for variability in sequencing depth before the downstream analysis. Taxonomic identity was assigned to OTUs using the naïve Bayesian classifier Ribosome Database Project (Wang *et al.*, 2007). The assembled bacterial 16S rRNA gene sequences have been deposited in the European Nucleotide Archive under the accession number PRJEB11871.

Wiley-Blackwell and Society for Applied Microbiology

This article is protected by copyright. All rights reserved.

Author Manuscript

Author Manuscript

Local similarity analysis and network construction

Network analysis has been widely used to explore the underlying interaction/association among microbial taxa, to predict the host of ARGs in complex microbial communities (Hu *et al.*, 2016b; Zhu *et al.*, 2017). The Local Similarity (LS) correlations (ranked Pearson's correlations) among bacteria, herbs and ARGs were determined by the LS analysis (eLSA) (Chow *et al.*, 2014). A correlation matrix was constructed by calculating the pairwise correlation coefficients among ARG subtypes and bacterial and herb taxa at the family level which occur in ≥ 50% of all the samples. This filtering step eliminated those poorly represented ARGs and bacterial and herb taxa to minimize the artificial association bias. We calculated *P*-values of the LS correlations using statistical approximation followed by permutation testing (Xia *et al.*, 2013) following the protocol as previously described (Chow *et al.*, 2014). We imported the resultant LS correlations into Cytoscape (version 2.8.3) (Shannon *et al.*, 2003) ~~plug-in CoNet (version 1.1)~~ for initial visualization, and explored the network topology ~~using the Circular algorithms~~ in the Gephi platform (Bastian *et al.*, 2009).

Statistical analysis

We calculated the tree diversity by counting the number of tree species in 400 m^2 quadrats, and the herb diversity by counting the number of herb species in 1 m^2 plots. Soil bacterial diversity was calculated for each sample using the number of OTUs at the 97% sequence similarity with a depth of 45,070 sequences per sample. One-way analysis of variance (ANOVA) was performed to test the difference in the diversity of ARGs, MGEs, trees, herbs, and bacteria, as well as the relative abundances of ARGs and MGEs across forest biomes in SPSS 20. Spearman's correlation analysis was conducted to test the relationships between the relative abundances of total MGEs with total ARGs and individual ARG types. The data were normalized using log-transformation when necessary prior to statistical analysis.

We conducted regression analysis to test the relationships between the diversity of ARGs with that of MGEs, bacteria, trees, and herbs as well as soil properties and climatic factors using the "ggplot2" package in the R platform (R Development Core Team, 2008). The curve fitting was estimated with linear and nonlinear models (e.g., quadratic and cubic models), and the best models were selected with the highest proportion of variance in the diversity of ARGs that can be explained by the independent variables. A heat map visualization of the relative abundances of individual ARG subtypes was generated using the

Wiley-Blackwell and Society for Applied Microbiology
This article is protected by copyright. All rights reserved.

Author Manuscript

"gplots" package in the R platform (version 3.3.1). The changes in ARG compositions across forest biomes were visualized by non-metric multidimensional scaling (NMDS) ordinations based on the Bray-Curtis dissimilarity distances, and statistically evaluated by permutational multivariate analysis of variance (PerMANOVA) using the "vegan" package with 999 permutations in R. Procrustes test for correlation analysis between ARGs and bacterial communities based on Bray-Curtis dissimilarity matrices was performed using "vegan" in R. Mantel test was performed to test the correlations between soil properties, MGEs, the community composition of trees and herbs with ARGs content with the Bray-Curtis method in "vegan" with 999 permutations. Canonical correspondence analysis (CCA) between ARG contents and environmental parameters was also performed in R using "cca" function from the "vegan" package. The threshold for significance was $P < 0.05$.

Random Forest analysis and Structural equation modelling

A classification Random Forest analysis was conducted to identify the most important predictors influencing the ARGs content among the following variables: climatic factors, spatial factors, soil properties, herb diversity, and bacterial diversity. The importance of each predictor was determined by assessing the increase in the mean square error (MSE) between observations and out-of-bag predictions as described previously (Delgado-Baquerizo *et al.*, 2016). The significance of the importance of each predictor on ARG patterns was assessed using the "rfPermute" package in R.

We constructed structural equation model (SEM) to determine the direct and indirect effects of soil properties (including soil pH), climatic conditions (including AMT), the diversity of bacteria, trees and herbs, and MGEs, on ARG patterns. SEM can partition direct and indirect effects that one variable might have on another, estimate the strengths of multiple effects, and ultimately provide mechanistic information on the drivers of ARGs pattern (Grace, 2006). We only included those variables that were identified as statistically significant predictors of the ARG patterns from the Random Forest analysis. The data used for the model construction were classified into several groups, including climate, spatial factors, soil properties, plant diversity, MGEs diversity, and ARGs diversity.

We established an *a priori* model based on the known relationships among the drivers of ARG patterns (Fig. S9). Mantel test was performed to calculate the pairwise correlations among these variables using the "Ecodist" package in R, and the matrix of resultant *R* values was parameterized into the AMOS 22 software for the SEM construction using the maximum-

Wiley-Blackwell and Society for Applied Microbiology
This article is protected by copyright. All rights reserved.

likelihood estimation method. Foe the mediation testing, ~~W~~we used a 1000-time bootstrap to control the false discovery rate (FDR) and set the confidence intervals at 95% for the significance assessment. The overall goodness of model fit was assessed using Akaike information criteria (AIC), goodness-of-fit index (> 0.90), χ^2 test ($P > 0.05$), and root mean square errors of approximation (RMSEA < 0.05). We also calculated the standardized total effects of each factor on ARG patterns by summing all direct and indirect pathways between the factor and ARGs.

Acknowledgements

This work was financially supported by Chinese Academy of Sciences (XDB15020200), Australian Research Council (DP170103628, DE150100870), and Natural Science Foundation of China (41601256).

References

Allen, H.K., Moe, L.A., Rodbumrer, J., Gaarder, A., and Handelsman, J. (2009) Functional metagenomics reveals diverse beta-lactamases in a remote Alaskan soil. *ISME J* **3**: 243-251.

Allen, H.K., Donato, J., Wang, H.H., Cloud-Hansen, K.A., Davies, J., and Handelsman, J. (2010) Call of the wild: antibiotic resistance genes in natural environments. *Nat Rev Microbiol* **8**: 251-259.

Alonso, A., Sánchez, P., and Martínez, J.L. (2001) Environmental selection of antibiotic resistance genes. *Environ Microbiol* **3**: 1-9.

Aminov, R.I. (2011) Horizontal gene exchange in environmental microbiota. *Front Microbiol* **2**: 158.

Ashbolt, N.J., Amezquita, A., Backhaus, T., Borriello, P., Brandt, K.K., Collignon, P. et al. (2013) Human health risk assessment (HHRA) for environmental development and transfer of antibiotic resistance. *Environ Health Perspect* **121**: 993-1001.

Bai, Y., Muller, D.B., Srinivas, G., Garrido-Oter, R., Potthoff, E., Rott, M. et al. (2015) Functional overlap of the Arabidopsis leaf and root microbiota. *Nature* **528**: 364-369.

Baldrian, P. (2017) Forest microbiome: diversity, complexity and dynamics. *FEMS Microbiol Rev* **41**: 109-130.

Wiley-Blackwell and Society for Applied Microbiology

This article is protected by copyright. All rights reserved.

Author Manuscript

Balsalobre, L.C., Dropa, M., Lincopan, N., Mamizuka, E.M., Matte, G.R., and Matte, M.H. (2009) Detection of metallo-β-lactamases encoding genes in environmental isolates of Aeromonas hydrophila and Aeromonas jandaei. *Lett Appl Microbiol* **49**: 142-145.

Bastian, M., Heymann, S., and Jacomy, M. Gephi: an open source software for exploring and manipulating networks. In: International AAAI Conference on Weblogs and Social Media: San Jose, California 2009.

Bates, S.T., Berg-Lyons, D., Caporaso, J.G., Walters, W.A., Knight, R., and Fierer, N. (2011) Examining the global distribution of dominant archaeal populations in soil. *ISME J* **5**: 908-917.

Bonan, G.B. (2008) Forests and climate change: forcings, feedbacks, and the climate benefits of forests. *Science* **320**: 1444-1449.

Berendonk, T.U., Manaia, C.M., Merlin, C., Fatta-Kassinos, D., Cytryn, E., Walsh, F. et al. (2015) Tackling antibiotic resistance: the environmental framework. *Nat Rev Microbiol* **13**: 310-317.

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K. et al. (2010) QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* **7**: 335-336.

Chase, M.W., and Reveal, J.L. (2009). A phylogenetic classification of the land plants to accompany ARG III. *Bot J Linean Soc* **161**: 122-127.

Chen, B., Yang, Y., Liang, X., Yu, K., Zhang, T., and Li, X. (2013) Metagenomic profiles of antibiotic resistance genes (ARGs) between human impacted estuary and deep ocean sediments. *Environ Sci Technol* **47**: 12753-12760.

Chen, Q.L., An, X.L., Li, H., Zhu, Y.G., Su, J.Q., and Cui, L. (2017) Do manure-borne or indigenous soil microorganisms influence the spread of antibiotic resistance genes in manured soil? *Soil Biol Biochem* **114**: 229-237.

Chow, C.E.T., Kim, D.Y., Sachdeva, R., Caron, D.A., and Fuhrman, J.A. (2014) Top-down controls on bacterial community structure: microbial network analysis of bacteria, T4-like viruses and protists. *ISME J* **8**: 816-829.

D'Costa, V.M., McGrann, K.M., Hughes, D.W., and Wright, G.D. (2006) Sampling the antibiotic resistome. *Science* **311**: 374-377.

D'Costa, V.M., King, C.E., Kalan, L., Morar, M., Sung, W.W.L., Schwarz, C. et al. (2011) Antibiotic resistance is ancient. *Nature* **477**: 457-461.

Author Manuscript

Wiley-Blackwell and Society for Applied Microbiology

This article is protected by copyright. All rights reserved.

De Souza, M.J., Nair, S., Bharathi, P.L., and Chandramohan, D. (2006) Metal and antibiotic-resistance in psychrotrophic bacteria from Antarctic marine waters. *Ecotoxicology* **15**: 379-384.

Delgado-Baquerizo, M., Maestre, F.T., Reich, P.B., Jeffries, T.C., Gaitan, J.J., Encinar, D. et al. (2016) Microbial diversity drives multifunctionality in terrestrial ecosystems. *Nat Commun* **7**: 10541.

Edgar, R.C. (2013) UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat Methods* **10**: 996-998.

Fierer, N., and Jackson, R.B. (2006) The diversity and biogeography of soil bacterial communities. *Proc Natl Acad Sci USA* **103**: 626-631.

Forsberg, K.J., Patel, S., Gibson, M.K., Lauber, C.L., Knight, R., Fierer, N., et al. (2014) Bacterial phylogeny structures soil resistomes across habitats. *Nature* **509**: 612-616.

Forsberg, K.J., Reyes, A., Wang, B., Selleck, E.M., Sommer, M.O.A., and Dantas, G. (2012) The shared antibiotic resistome of soil bacteria and human pathogen. *Science* **337**: 1107-1111.

~~Gauthier, S., Bernier, P., Kuuluvainen, T., Shvidenko, A.Z., and Schepaschenko, D.G. (2015) Boreal forest health and global change. *Science* **349**: 819-822.~~

Gibson, M.K., Forsberg, K.J., and Dantas, G. (2015) Improved annotation of antibiotic resistance determinants reveals microbial resistomes cluster by ecology. *ISME J* **9**: 207-216.

Gillings, M.R., Gaze, W.H., Pruden, A., Smalla, K., Tiedje, J.M., and Zhu, Y.G. (2015) Using the class 1 integron-integrase gene as a proxy for anthropogenic pollution. *ISME J* **9**: 1269-1279.

Gilmore, M.S., and Hoch, J.A. (1999) Antibiotic resistance: a vancomycin surprise. *Nature* **399**: 524-527.

Grace, J.B. (2006) Structural Equation Modelling and Natural Systems. Cambridge University Press.

Han, X.M., Hu, H.W., Shi, X.Z., Wang, J.T., Han, L.L., Chen, D. et al. (2016) Impacts of reclaimed water irrigation on soil antibiotic resistome in urban parks of Victoria, Australia. *Environ Pollut* **211**: 48-57.

Hattenschwiler, S., Tiunov, A.V., and Scheu, S. (2005) Biodiversity and litter decomposition in terrestrial ecosystems. *Annu Rev Ecol Evol* **36**: 191-218.

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

Wiley-Blackwell and Society for Applied Microbiology

This article is protected by copyright. All rights reserved.

Author Manuscript

Hu, H.W., Han, X.M., Shi, X.Z., Wang, J.T., Han, L.L., Chen, D. et al. (2016a) Temporal changes of antibiotic resistance genes and bacterial communities in two contrasting soils treated with cattle manure. *FEMS Microbiol Biol* **92**: fiv169.

Hu, H.W., Wang, J.T., Li, J., Li, J.J., Ma, Y.B., Chen, D. et al. (2016b) Field-based evidence for copper contamination induced changes of antibiotic resistance in agricultural soils. *Environ Microbiol* **18**: 3896-3909.

Hu, H.W., Wang, J.T., Li, J., Shi, X.Z., Ma, Y.B., Chen, D. et al. (2017) Long-term nickel contamination increases the occurrence of antibiotic resistance genes in agricultural soils. *Environ Sci Technol* **51**: 790-800.

Karkman, A., Do T.T., Walsh, F., and Virta, M.P.J. (2018) Antibiotic-resistance genes in waste water. *Trends Microbiol* **26**: 220-228.

Johnson, T.A., Stedtfeld, R.D, Wang, Q., Cole, J.R., Hashsham, S.A., Looft, T. et al. (2016) Clusters of antibiotic resistance genes enriched together stay together in swine agriculture. *mBio* **7**: e02214-15.

Li, B., Yang, Y., Ma, L., Ju, F., Guo, F., Tiedje, J.M., et al. (2015) Metagenomic and network analysis reveal wide distribution and co-occurrence of environmental antibiotic resistance genes. *ISME J* **9**: 2490-2502.

Looft, T., Johnson, T.A., Allen, H.K., Bayles, D.O., Alt, D.P., Stedtfeld, R.D., et al. (2012) In-feed antibiotic effects on the swine intestinal microbiome. *Proc Nat Acad Sci USA* **109**: 1691-1696.

Magoè, T., and Salzberg, S.L. (2011) FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* **27**: 2957-5963.

Martinez, J.L. (2008) Antibiotics and antibiotic resistance genes in natural environments. *Nature* **321**: 365-367.

Martinez, J.L., Coque, T.M., and Baquero, F. (2015) What is a resistance gene? Ranking risk in resistomes. *Nat Rev Microbiol* **13**: 116-123.

McEachran, A.D., Blackwell, B.R., Hanson, J.D., Wooten, K.J., Mayer, G.D., Cox, S.B. et al. (2015) Antibiotics, bacteria, and antibiotic resistance genes: aerial transport from cattle feed yards via particulate matter. *Environ Health Perspect* **123**: 337-343.

Nesme, J., Cécillon, S., Delmont, T.O., Monier, J.M., Vogel, T.M., Simonet, P. (2014) Large-scale metagenomic-based study of antibiotic resistance in the environment. *Curr Biol* **24**: 1096-1100.

Author Manuscript

Wiley-Blackwell and Society for Applied Microbiology

This article is protected by copyright. All rights reserved.

Nesme, J., and Simonet, P. (2014) The soil resistome: a critical review on antibiotic resistance origins, ecology and dissemination potential in telluric bacteria. *Environ Microbiol* **17**: 913-930.

Pawlowski, A.C., Wang, W., Koteva, K., Barton, H.A., McArthur, A.G., and Wright, G.D. (2016) A diverse intrinsic antibiotic resistome from a cave bacterium. *Nat Commun* **7**: 13803.

Partridge, S.R., Tsafnat, G., Coiera, E., and Iredell, J.R. (2009) Gene cassettes and cassette arrays in mobile resistance integrons. *FEMS Microbiol Rev* **33**: 757-784.

Prober, S.M., Leff, J.W., Bates, S.T., Borer, E.T., Firn, J., Harpole, W.S. et al. (2015) Plant diversity predicts beta but not alpha diversity of soil microbes across grasslands worldwide. *Ecol Lett* **18**: 85-95.

Pruden, A., Pei, R., Storteboom, H., and Carlson, K.H. (2006) Antibiotic resistance genes as emerging contaminants: studies in Northern Colorado. *Environ Sci Technol* **40**: 7445-7450.

R Development Core Team. R: A language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria. URL: http://www.R-project.org. 2008.

Schmittgen, T.D., and Livak, J.K. (2008) Analyzing real-time PCR data by the comparative CT method. *Nat Protoc* **3**: 1101-1108.

Scholes, R., and Archer, S. (1997) Tree-grass interactions in savannas. *Annu Rev Ecol Syst* **28**: 517-544.

Segawa, T., Takeuchi, N., Rivera, A., Yamada, A., Yoshimura, Y., Barcaza, G. et al. (2013) Distribution of antibiotic resistance genes in glacier environments. *Environ Microbiol Rep* **5**: 127-134.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D. et al. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **13**: 2498-2504.

Stalder, T., Barraud, O., Casellas, M., Dagot, C., and Ploy, M.C. (2012) Integron involvement in environmental spread of antibiotic resistance. *Front Microbiol* **3**: 119.

Stedtfeld, R.D., Williams, M., Fakher, U., Johnson, T., Stedtfeld, T.M., Wang, F. et al. (2016) Antimicrobial resistance (AR) dashboard application for mapping environmental occurrence and resistant pathogens. *FEMS Microbiol Ecol* **92**: fiw020.

Su, J.Q., Wei, B., Ou-Yang, W.Y., Huang, F.Y., Zhao, Y., Xu, H.J., et al. (2015) Antibiotic resistome and its association with bacterial communities during sewage sludge composting. *Environ Sci Technol* **49**: 7356-7363.

Wiley-Blackwell and Society for Applied Microbiology

This article is protected by copyright. All rights reserved.

Author Manuscript

Wang, F., Stedtfeld, R.D., Kin, O.S., Chai, B., Yang, L., Stedtfeld, T.M. et al. (2016a) Influence of soil characteristics and proximity to Antarctic research stations on abundance of antibiotic resistance genes in soils. *Environ Sci Technol* **50**: 12621-12629.

Wang, F.H., Qiao, M., Su, J.Q., Chen, Z., Zhou, X., and Zhu, Y.G. (2014) High throughput profiling of antibiotic resistance genes in urban park soils with reclaimed water irrigation. *Environ Sci Technol* **48**: 9079-9085.

Wang, J.T., Zheng, Y.M., Hu, H.W., Li, J., Zhang, L.M., Chen, B.D. et al. (2016b). Coupling of soil prokaryotic diversity and plant diversity across latitudinal forest ecosystems. *Sci Rep* **6**: 19561.

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

World Health Organization. (2014) Antibiotic resistance: global report on surveillance. World Health Organization, Geneva, Switzerland.

Wright, G.D. (2010) Antibiotic resistance in the environment: a link to the clinic? *Curr Opin Microbiol* **13**: 589-594.

Wright GD, and Poinar H. (2012) Antibiotic resistance is ancient: implications for drug discovery. *Trends Microbiol* **20**: 157-159.

Xia, L.C., Ai, D., Gram, J.A., Fuhrman, J.A., and Sun, F. (2013) Efficient statistical significance approximation for local association analysis of high-throughput time series data. *Bioinformatics* **29**: 230-237.

Xiang, Q., Chen, Q.L., Zhu, D., An, X.L., Yang, X.R., S, J.Q. et al. (2018) Spatial and temporal distribution of antibiotic resistomes in a peri-urban area is associated with anthropogenic activities. *Environ Pollut* **235**: 525-533.

Yergeau, E., Sanschagrin, S., Mynard, C., St-Arnaud, M., and Greer, C.W. (2014) Microbial expression profiles in the rhizosphere of willows depend on soil contamination. *ISME J* **8**: 344-358.

Zhang, Y.J., Hu, H.W., Gou, M., Wang, J.T., Chen, D.L., and He, J.Z. (2017) Temporal succession of soil antibiotic resistance genes following application of swine, cattle and poultry manures spiked with or without antibiotics. *Environ Pollut* **231**: 1621-1632.

Zhu, Y.G., Zhao, Y., Li, B., Huang, C.L., Zhang, S.Y., Yu, S. et al. (2017) Continental-scale pollution of estuaries with antibiotic resistance genes. *Nat Microbiol* **2**: 16270.

Author Manuscript

Wiley-Blackwell and Society for Applied Microbiology

This article is protected by copyright. All rights reserved.

Zhu, Y.G., Johnson, T.A., Su, J.Q., Qiao, M., Guo, G.X., Stedtfeld, R.D. et al. (2013) Diverse and abundant antibiotic resistance genes in Chinese swine farms. *Proc Natl Acad Sci USA* **110**: 3435-3440.

Author Manuscript

Wiley-Blackwell and Society for Applied Microbiology

This article is protected by copyright. All rights reserved.

Figure legends

Figure 1 A schematic diagram showing the census of trees and herbs and soil sampling across five forest biomes over a span of 4000 km (a). The dots with different colours represent soil sampling sites for different forest biomes. We established 60 independent large quadrats (20 × 20 m^2) in each forest biome, and in each quadrat, all trees with diameter at breast height ≥2.5 cm were recorded. Three small plots (1 × 1 m^2) were randomly established within each 400-m^2 quadrat for the census of herbaceous plants. Soil samples were collected from each quadrat by thoroughly mixing 15 random soil cores, resulting in a total of 300 soil samples. Comparison of the diversity and relative of ARGs (b, d) and MGEs (c, e) across forest biomes. Line, average diversity or relative abundance; box, 25th and 75th percentiles; and error bars, 10th and 90th percentiles. Different letters above the boxes indicate a significant difference.

Figure 2 The heat map showing the broad-spectrum quantitative profile of ARGs and MGEs subtypes detected in five forest biomes (a). Soil samples were shown as columns and ARGs/MGEs shown as rows. The dendrograms for columns and rows were calculated based on the Bray-Curtis dissimilarity matrices. Non-metric multidimensional scaling ordination plot depicts the Bray-Curtis dissimilarity distances between soils based on the relative abundance of ARGs across forest biomes (b). The 2D stress value is 0.14, which indicated that the two-dimensional ordinations could well represent the data. Procrustes analyses depict significant correlations between ARG contents (Bray-Curtis) and bacterial composition (Bray-Curtis) (c). Lines connect ARG data and 16S rRNA gene sequence data. The M^2 fit reported is from a Procrustes transformation over the first two principal coordinates while the *P*-value is calculated from a distribution of empirically determined M^2 values over 10,000 Monte Carlo label permutations. (Abbreviations: FCA: fluoroquinolone, quinolone, florfenicol, chloramphenicol and amphenicol resistance genes; MLSB: Macrolide-Lincosamide-Streptogramin B resistance genes.)

Figure 3 Relationships between the diversity of ARGs with the bacterial diversity (a), herb diversity (b), MGEs diversity (c), and soil pH (d). Line in each plot represents the best-fit curve and the shaded area represents its 95% confidence limits.

Figure 4 The networks showing the co-occurrence patterns among the detected ARG subtypes, MGEs, bacteria, and herbs across forest biomes. (a) The nodes coded with different colors represent ARGs, MGEs, bacteria and herbs; (b) The nodes coded with different colors represent different categories of ARGs, and the nodes for MGEs, bacteria and herbs are

Wiley-Blackwell and Society for Applied Microbiology

This article is protected by copyright. All rights reserved.

Author Manuscript

shown in grey colour. The edges connecting nodes correspond to statistically significant correlations between nodes. Node size is proportional to the number of connections between nodes (degree). The thickness of the edges is proportional to the correlation coefficient.

Figure 5 Structural equation modelling of the direct and indirect effects of molecular attributes (bacterial diversity), plant attributes (tree and herb diversity), soil properties, climatic conditions, biome types, and MGEs on the diversity of ARGs (a). The hypothetical models fit our data well, as suggested by the goodness-of-fit statistics: $\chi^2 = 33.0$, $P = 0.11$, d.o.f = 24, GFI = 0.98, AIC = 167.0, and RMSEA = 0.035. Arrows indicate the hypothesized direction of causation. The width of arrows is proportional to the strength of the path coefficients. The numbers adjacent to arrows are standardized path coefficients and indicative of the effect size of the relationship. The R^2 value alongside the response variable indicates the proportion of variance explained by the variable. Only significant path coefficients are displayed. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$. (b) Standardized total effects (the sum of direct plus indirect effects from SEM) of the major factors on the diversity of ARGs. (AMT, annual mean temperature; PDM, Precipitation of driest month; MTCM, Mean temperature of coldest month)

Author Manuscript

Wiley-Blackwell and Society for Applied Microbiology

This article is protected by copyright. All rights reserved.