Metagenomic analysis reveals the shared and distinct features of the soil resistome across tundra, temperate prairie and tropical ecosystems

CURRENT STATUS: UNDER REVIEW

Xun Qian
Michigan State University

Santosh Gunturu
Michigan State University

Jiarong Guo
Michigan State University

Benli Chai
Michigan State University

James R. Cole
Michigan State University

Jie Gu
Northwest Agriculture and Forestry University

James Tiedje
Michigan State University

tiedje@msu.edu Corresponding Author
ORCiD: https://orcid.org/0000-0002-8992-6218

DOI:
10.21203/rs.3.rs-22597/v1

SUBJECT AREAS
General Microbiology

KEYWORDS
Soil resistome, geographical distribution, background ARG, clinical ARG, anthropogenic impact, metagenomic analysis
Abstract

Background
Soil is likely the largest reservoir of antibiotic resistance genes (ARGs), but their distribution across soil resistomes of different ecosystems is unknown. We used a metagenomic approach to investigate ARG types and amounts in soil DNA of three native ecosystems: Alaskan tundra, US Midwestern prairie, and Amazon rainforest, as well as the effect of conversion of the latter two to agriculture and pasture, respectively.

Results
High diversity (242 ARG subtypes) and abundance (0.184–0.242 ARG copies per 16S rRNA gene copy) were observed irrespective of ecosystem, with multidrug resistance genes and efflux pump the dominant class and mechanism. We identified 55 “background ARGs” which were shared by all 26 soil metagenomes of all three ecosystems, which accounted for more than 81% of resistome abundance. No significant differences in both ARG diversity and abundance were observed between native prairie soil and adjacent long-term cultivated agriculture soil. Conversion of Amazon rainforest to pasture significantly increased soil ARG diversity, with eight ARGs significantly enriched. We chose 12 clinically important ARGs to evaluate at the sequence level and found them to be distinct from those in human pathogens, and when assembled they were even more dissimilar. Significant correlation was found between bacterial community structure and resistome profile, suggesting that variance in resistome profile was mainly driven by the bacterial community composition. There was also wide co-occurrence among ARGs in all samples.

Conclusions
Our results define common background ARGs, quantify resistance classes, provide sequence information suggestive of very low risk but also revealing resistance gene variants that might emerge in the future.

Background
Antibiotic resistance is a global threat to public health, which led to an estimated 700,000 yearly deaths in the world and is predicted to cause 10 million deaths by 2050 if unchecked [1]. Soils are likely the most significant ARG reservoirs, and large amounts and diversities of ARGs have been found in soils throughout the world [2–4], including some in Antarctic surface soils [5]. Many clinically-
associated antibiotic resistance genes (ARGs) originated from the soil resistome via horizontal gene transfer [6-8]. Further studies demonstrate that some anthropogenic activities significantly enrich the abundance of indigenous ARGs in soils [9-11]. Understanding soil resistomes at a broad geographic scale and across various ecosystems, especially in native soils which haven't been exposed to anthropogenic activities, can help better define the background levels and types of ARGs, which is essential for assessing the impact and potential risk of new human activities.

Different land use practices can significantly alter the soil physicochemical as well as biological properties. Conversion of native soil to crop cultivation has been the most common land use change. For example, 100 years of continuous cultivation significantly changed microbial diversity and structure in a consistent though not major way in North American Midwest prairie soil [12]. Further, the conversion of Amazon rainforest to cattle pasture is another land use change that has expanded in the tropics [13], and has led to homogenization of microbial communities [14]. As such, considering that microbes are carriers of ARGs, altering the microbial community could also affect the soil resistome irrespective of the introduction of selection pressure for antibiotic resistance. Since land use change has been and continues to be the most extensive alteration of the terrestrial environment, its impact on the soil resistome is important to understand.

Most of our knowledge of ARGs in soils has come from targeting genes by real-time quantitative PCR (RT-qPCR) [15, 16]. In recent years, the number of genes and samples possible to assay has been expanded by highly parallel qPCR platforms [17, 18]. For example, Wang et al. [9] detected 147 ARGs in park soil samples, and found the application of reclaimed irrigation water resulted in a 99 to 8655 fold enrichment in total resistome abundance compared to respective controls. Furthermore, manure application to croplands temporarily increases ARG type and concentrations [19, 20]. While qPCR is more sensitive, it is limited to the known genes and the specificities imposed by primers.

Metagenomics (shotgun sequencing), however is now more affordable and it provides a more comprehensive overview of environmental resistomes [21, 22]. Furthermore, new ARG bioinformatic analysis tools have recently become available to efficiently analyze this large amount of data [23-25]. Since different ARGs have different levels of risk it's necessary to categorize ARGs by their functional
roles, the necessary components for their resistance function and assess their potential risk separately rather than merely the abundance of total ARGs.

We used the ARGs-OAP v2.0 pipeline, consisting of a hierarchical structured (ARGs type-subtype-reference sequence) database [26, 27] to detect ARGs in 26 soil metagenome samples from Alaska tundra, U.S. Midwest prairie and Amazon rainforest ecosystems. While a few other studies have surveyed metagenomic data in public databases for ARGs [4, 28], we focused on those representing a polar to tropic latitudinal range and soils that were native and matched for edaphic traits with ones that had undergone land use change, all with well-characterized metadata. We addressed the following objectives: i) what are the types and quantities of ARGs in soils of three climate regions (tundra, temperate and tropical), ii) what are the impacts of major land use changes on the ARG profiles, and iii) define which ARGs are common, perhaps universal background, and within the ARGs, which are most frequent, do they commonly co-occur and their relevance to risk. The results of this study should improve our understanding of the background level and classes of soil ARGs and allow for better evaluation of the public health risk of ARGs in the environment.

Materials And Methods
Experimental design and site description
The metagenomic sequence data used herein were from our previous studies, and used to assess the impact of land use change [14, 29] and global warming [30, 31] on soil microbial communities. The sample site locations of the 26 soil metagenomes used in this study are depicted in Additional file 1: Fig. S1. The Alaskan soils were sampled at 15–25 cm depth (active layer; above permafrost boundary) at a moist acidic tundra area in Interior Alaska near Denali National Park (63°52’59” N, 149°13’32” W) in May 2010. The Oklahoma soils were collected by soil core 0-15 cm deep from a tallgrass prairie located at the Great Plain Apiaries in McClain County, Oklahoma, United States (34°59’N, 97°31’W) in 2011, 2012 and 2013. The site was abandoned from field cropping 40 years ago with light grazing until 2008. Three additional tallgrass prairie ecosystem sites were sampled in the summer of 2009 in the US Midwest from a 750 km transect from Kansas through Iowa to Wisconsin. At each site soil was sampled from a native (never tilled) prairie and an adjacent cultivated (> 100 years) soil matched for
soil edaphic traits and landform. The native prairie soils had been grazed by cattle. All cultivated soils had received manure application. The Amazon soils were sampled at the Amazon Rainforest Microbial Observatory site (10°10′05″ S and 62°49′27″ W) in April 2009. Five soil cores of 0-10 cm deep were collected from a primary rainforest and an adjacent 38-year-old converted pasture. The Amazon rainforest soils were never grazed, while the pasture had been continuously used for beef cattle grazing since conversion. Detailed information about the climate, vegetation, soil type and chemistry at each sampling site were described previously [12, 29, 31, 32].

Shotgun Sequencing
Sequencing of Alaska, Oklahoma, and Amazon soils was performed on the Illumina HiSeq platform with 150-bp paired-end strategy by the Joint Genome Institute (JGI). Illumina paired-end augmented with some 454 Titanium sequencing were used for the Iowa, Kansas, and Wisconsin soils. The Alaskan soils were also sequenced by Illumina using paired ends but by Los Alamos National Laboratory. Identification Of Args In Shotgun Data
Adapter reads in the sequence data were removed, the remaining reads were filtered to discard bases with a quality score < 20 and length < 50 base pair (bp) by SolexaQA v.3.1.7.1 [33]. To eliminate the differences caused by variations in the sequencing depth among samples, the least number (200 million) of sequences obtained were randomly picked from each sample. The retrieved sequences were then used to search for ARGs following ARGs-OAP v2.0 pipeline as described by Yin et al. [27]. The parameters used for ARG identification were: alignment length cut-off of 75 nucleotides, alignment e-value cut-off of \(10^{-7}\), and alignment identity of 80%. The abundances of ARGs were normalized by 16S rRNA gene expressed as: Abundance= \[26\]. Analysis of metagenome data of the Earth Microbiome Project shows that the 16S rRNA copy number of all soils is very narrow, with a mean of 2.2 16S rRNA copies per cell [34].

Identification Of Clinical Args
Protein sequences of human disease associated bacterial genomes were collected from the Pathosystems Resource Integration Center (PATRIC) [35]. The collected protein sequences were searched against SARG database and hits with identity \(\geq 80\%\) and length coverage (alignment length/reference ARG length) \(\geq 80\%\) for clinical ARGs. Reads annotated as ARGs by SARG database
were extracted from metagenomes of Amazon rainforest and pasture soils, and they were evaluated by BLAST against the protein sequences of clinical ARGs. The hits with identity ≥ 80%, alignment length ≥ 75 bp nucleotides, and e-value ≤ $10^{-7}$ were regarded as clinical ARGs. The abundance of a clinical ARG was calculated with the same formula to calculate ARG abundance in ARGs-OAP v2.0 pipeline.

To comprehensively understand the homology between ARGs in soil and ARGs found in clinical settings, 12 ARGs in Amazon data sets were assembled with Xander, a target gene assembler [36, 37]. The seed sequences were obtained from SARG database [27]. Hidden Markov Models (HMMs) of target genes were built and used to search against nr database. Hits with e-value < $10^{-7}$ and HMM coverage > 80% were kept and the protein sequences were retrieved from nr database. Nucleotide sequences were fetched by accession number of the protein sequences. The protein and nucleotide sequences, seed sequences and HMM were then used with Xander assembler pipeline. Assembled protein contigs of ≥ 100 amino acids were evaluated by BLAST against clinical ARGs.

**Soil Bacterial Community**

Bacterial taxonomic classification and abundance quantification were analyzed following the SSUsearch pipeline [38]. Briefly, a 16S rRNA gene HMM was used to search against metagenomic data and the hits were annotated with SILVA database. The sequences aligned to a part of 16S V4 variable region (577–657) and with lengths greater than 70 bp were extracted. The extracted 16S rRNA gene sequences were then clustered to estimate OTU number at 95% identity.

**Statistical Analyses**

Only ARGs detected with more than two reads across all samples were retained for further analyses. Bray-Curtis distance-based principal coordinates analysis (PCoA) was performed to estimate the variance of resistome profiles. Procrustes analysis was used to assess the relationship between resistome profile and bacterial community structure, 9999 permutations were used to test the significance. ARGs detected in at least in five samples with a maximum read number > 5 in at least one sample, were kept for ANOVA and network analyses. The differences in ARG abundances across three soils were tested by ANOVA analysis (Least Significant Difference, $p < 0.05$). Spearman’s
correlation coefficients were calculated based on the read number among soils. Network analysis was performed in Cytoscape to identify ARG clusters; only ARGs with significant (Least Significant Difference, $p < 0.01$) and strong Spearman coefficients ($> 0.9$) were used. The ANOVA and Spearman correlation analyses were conducted using SPSS 23.0. PCoA and Procrustes analysis were done with R3.5.1.

Results
Detected ARGs and regulatory genes
The SARG database identifies ARG types (antibiotic class) and within that class subtypes (e.g. a subtype having $>80\%$ identical aligned bases based on HMM model). Diversity data are derived from the number of subtypes. A total of 268 ARG subtypes potentially conferring resistance to 21 classes of antibiotics were detected in the soils, with most of them belonging to antibiotic deactivation (106 ARGs) and efflux pump (93 ARGs) mechanisms (Fig. 1). More than 58% of the resistome abundance was contributed by efflux pump genes, while only 16% was from the deactivation and cellular protection classes. Ten regulatory genes ($mtrR$, $gadX$, $tetR$, $mexT$, $cAMP$-regulatory protein, $arlR$, $ompR$, $vanS$, $cpxR$, $vanR$) were recognized as “ARGs” by the SARG database. These genes accounted for 13 to 35% of resistome abundances in the studied soils and among them $arlR$, $cpxR$, $ompR$, $vanR$, $vanS$ were dominant and observed in all soils (Additional file 1: Fig. S2). Since regulatory genes do not directly confer resistance and they inflate the quantitation, they are not included in further ARGs analyses in this report.

Multidrug resistance genes were most abundant (57.1% of resistome abundance) with 67 subtypes observed, followed by macrolides-lincosamides-streptogramines (MLS) resistance genes, with 28 subtypes that comprised 11.2% of resistome abundance (Fig. 1B). There were 71 beta-lactam, 24 tetracycline and 20 aminoglycoside resistance gene subtypes detected in these soils, but they on average only accounted for 2.6%, 3.0% and 1.6% of resistome abundances, respectively.

Shared Args By All Soils
Fifty-five (non-regulatory) ARGs were shared by all 26 soils (Fig. 2), which we define as “background ARGs” hereafter. These background ARGs accounted for 81.5 to 98.6% of resistome abundance across all soils, regardless of native or anthropogenic. The background ARGs consist of nine classes of
antibiotic resistances, but most of them are multidrug resistance genes (31 subtypes with background levels of $1.7 \times 10^{-2} - 6.7 \times 10^{-1}$ copies per 16S rRNA gene copy). There were six background tetracycline resistance genes and five vancomycin resistance genes, and they comprised $1.0 \times 10^{-3} - 1.2 \times 10^{-2}$ of resistome abundance. Only two of the background ARGs belong to the beta-lactam resistance class, and their background concentrations were $4.6 \times 10^{-4} - 3.8 \times 10^{-3}$ copies per 16S rRNA gene copy. The background levels of aminoglycoside and trimethoprim resistance genes differed considerably across the three ecosystems, ranging from $9.4 \times 10^{-5}$ to $2.0 \times 10^{-3}$ copies per 16S rRNA gene copy. Efflux pump was the dominant mechanism in background ARGs, contributing 80.6% of the total background ARG abundance. Four efflux pump genes accounted for at least 10% of all background abundance, including three multidrug resistance genes ($mdtB$, $mdtC$, and $mexF$) and a MLS resistance gene ($macB$).

**Args In Native Soils**

A total of 242 ARG subtypes were observed in the native soils, covering all detected classes of antibiotic resistance (Fig. 3). There were 144, 191, and 215 ARG subtypes with resistome abundances of 0.195, 0.201, and 0.243 copies per 16S rRNA gene detected in Alaskan tundra soil, Midwestern US native prairie soils and Amazon rainforest soils, respectively. There were no significant differences ($p > 0.05$) in ARG diversity nor resistome abundance between native and anthropogenic soils. Eleven ARGs detected in anthropogenic soils were not found in native soils, and they were all beta-lactam resistance genes. The background ARGs accounted for 93.6% of resistome abundance in native soils, but 86.8% in anthropogenic soils. Much of the resistome abundance was contributed by gene components of efflux pump complexes, such as $mdtABC-tolC$, $acrAB$, $mexEF-oprN$, and $rosAB$ (Fig. 4A). Thirteen subtypes of vancomycin resistance genes were detected in native soils, and no significant difference was observed ($p = 0.07$) between native and anthropogenic soils (Fig. 4B). We checked 16 ARGs which were commonly found in plasmids, including eight clinically important ARGs that potentially confer aminoglycoside and beta-lactam resistances. Significant differences were found among all of the three ecosystems, with Alaskan tundra soils harboring the lowest abundance
of these potential mobile ARGs, and Midwest prairie the highest abundance (Fig. 4C).

Soil Resistome At Different Geographical Locations

PCoA analysis clearly demonstrates that soil resistome profiles grouped by geographic location (Fig. 5). ARG diversity in the tundra was significantly lower ($p < 0.01$) than that of temperate and tropical areas (Fig. 3), while no significant difference ($p = 0.91$) was found between temperate and tropical areas. All ARGs detected in Alaskan soils were found in Midwest America and Amazon soils (Additional file 1: Fig. S3). One-hundred and forty ARGs were observed in both temperate and tropical soils, but not in tundra. The temperate and tropical soils shared 87.6% of ARGs, while 27 ARGs were only found in one or the other. The tropical soils had the highest resistome abundance, but no statistical difference ($p > 0.05$) was found among the three areas. Procrustes analysis showed that the ARG profile was significantly correlated with the bacterial community structure (sum of squares $M^2 = 0.183$, $r = 0.904$, $p < 0.01$) (Fig. 5B).

Effect Of Cultivated Agriculture On Soil Resistome

There were no observed significant differences ($p > 0.05$) in either diversity or total ARG abundance between the US Midwest native and the long-term cultivated soils (Fig. 3). The resistome profiles were very similar in the cultivated soils from the three sampling sites (Fig. 5). The vancomycin resistance genes and potentially mobile ARGs in native soil was higher than in cultivated soils, but the difference was not significant ($p = 0.56$) (Fig. 4B and Fig. 4C). norB and mexD were the only two genes significantly ($p < 0.05$) altered by cultivation - they were enriched 2.9 and 1.8 times in cultivated soils, respectively.

Changes of soil resistome during conversion of the Amazon rainforest to pasture

Conversion of Amazon rainforest to pasture significantly altered the soil resistome profile (Fig. 5). The conversion to pasture led to an increase of soil ARG diversity, 23 ARG subtypes could only be detected in pasture soils (Additional file 1: Fig. S4). Most of these genes were beta-lactam (11 subtypes) and aminoglycoside (7 subtypes) resistance genes. All exclusive ARGs (in rainforest or pasture soils), except dfrB2, had a relatively low abundance, ranging from $1.8 \times 10^{-5}$-3.1 $\times 10^{-4}$ per 16S rRNA gene. The resistome abundance in Amazon pasture soil was 11.8% lower than that in the
native rainforest soil, although it was not statistically different ($p = 0.06$). Similarly, no significant difference was observed in the abundance of potentially mobile ARGs between Amazon pasture and rainforest soils (Fig. 4C). Eight ARGs were significantly attenuated, while another eight ARGs were significantly enriched after conversion to pasture (Table 1). $dfrB2$, an integron-encoded dihydrofolate reductase gene, was only found in pasture soil with abundance of $0.0004$ copies per 16S rRNA gene copy. $norB$ was the most enriched ARG (16.4 times) in pasture soil.

**Table 1**

| Type       | Subtype | Abundance in Forest soil | Abundance in Pasture soil | Fold change |
|------------|---------|--------------------------|---------------------------|-------------|
| Beta-lactam| LRA-3   | 0.0029                   | 0.0007                    | 0.24        |
| Beta-lactam| LRA-9   | 0.0014                   | 0.0004                    | 0.26        |
| Tetracycline| tetV    | 0.0004                   | 0.00024                   | 0.40        |
| Beta-lactam| FEZ-1   | 0.0030                   | 0.0013                    | 0.43        |
| Aminoglycoside| $aac(2')$-I | 0.0014               | 0.0006                    | 0.46        |
| Aminoglycoside| $aac(3)$-I | 0.0016               | 0.0009                    | 0.38        |
| Multidrug   | mdtB    | 0.0409                   | 0.0246                    | 0.60        |
| Multidrug   | mdtC    | 0.0406                   | 0.0212                    | 0.61        |
| Multidrug   | oprC    | 0.0003                   | 0.0005                    | 1.56        |
| Multidrug   | mexE    | 0.0007                   | 0.0011                    | 1.61        |
| Vancomycin  | vanA    | 0.0007                   | 0.0012                    | 1.68        |
| Quinolone   | mfpA    | 0.0003                   | 0.0006                    | 2.00        |
| Fosfomycin  | fosB    | 0.0006                   | 0.0024                    | 3.74        |
| Multidrug   | opcM    | 0.0000                   | 0.0001                    | 4.77        |
| Quinolone   | norB    | 0.0001                   | 0.0012                    | 16.35       |
| Trimethoprim| $dfrB2$ | -                        | 0.0004                    | -           |

*Fold Change is the ratio of gene abundance in Amazon pasture soil to that in Amazon forest soil. Genes in bold are in the list of background ARGs.

**Detection Of Clinical Args In Soil Resistome**

To understand the relationship between the soil resistome and clinical ARGs, we analyzed the abundance of 12 clinically important ARGs, including two beta-lactam ($ampC$ and $FEZ-1$), one quinolone ($mfpA$), three aminoglycoside ($aac(2')$-$I$, $aac(6')$-$I$, and $aph(6')$-$I$), and six tetracycline resistance genes in Amazon rainforest and pasture soils (Fig. 6). Clinical-similar reads were detected in all 12 ARGs except $mfpA$ at amino acid identity cut-off of 80%. Among them, more than 74% of $FEZ-1$, $aac(2')$-$I$, $tetC$, $tetO$, and $tetV$ reads had this level of amino acid identity to their clinical types. However, only six ARGs were recovered at 90% sequence similarity. Highly-similar (97% identity) clinical ARG reads were detected for $FEZ-1$, $tetC$, and $tetX$, but they only account for 0.9–4.4% of their environmental (clinical and non-clinical) abundance and none had sequences identical to those ARGs in clinical pathogens. $tetV$ was the most abundant clinical ARG, with abundance of $4.1 \times 10^{-4}$ and $2.4 \times 10^{-4}$ copies per 16S rRNA gene in Amazon rainforest and pasture soils respectively. No statistically
difference ($p > 0.05$) was found in abundance of clinical ARGs at all chosen identity levels between Amazon rainforest and pasture soils. Clinical-similar $ampC$ was only detected in Amazon pasture soil at identity level ≥ 90% but its abundance was as low as $5.3 \times 10^{-7}$ copies per 16S rRNA gene. Only three of the 12 selected ARGs were successfully assembled, namely $ampC$, $tetM$ and $tetO$. Assembled $ampC$ was only found in Amazon forest soils, and it shared 46-56% of amino acid similarity with clinical $ampC$. Assembled $tetM$ and $tetO$ shared 62-71% and 69-74% of similarity to their clinical types found in human pathogens. Consistent to the results by SARG pipeline, no significant differences were found in assembled $tetM$ and $tetO$ abundances between Amazon forest and pasture soils. The abundances of assembled $tetM$ and $tetO$ were 0.015-0.024 (number of reads covering at least one kmer in gene contigs divided by the number of reads covering at least one kmer in the rplB contigs) and 0.064-0.097 in Amazon forest and pasture soils.

Identification Of Arg Clusters In Soils
Network analysis was performed to identify the universal ARG clusters present in all soils (Fig. S5). We found wide co-occurrence of multiple genes among all soil samples, despite their distant geographic locations, various ecosystems, and with or without anthropogenic activities. A total of eight types of ARG clusters were observed, comprised of 67 ARG subtypes potentially conferring resistance to 12 classes of antibiotics. Most of the co-occurred ARGs are multidrug resistance genes (27 subtypes), followed by tetracycline resistance genes (9 subtypes) and beta-lactamase resistance genes (8 subtypes). Four aminoglycoside resistance genes were found in ARG clusters and they had the highest average linkage with 4.8 to other ARGs. The largest ARG cluster was comprised of 25 ARGs, where $ThinB$ and $aac(3)$-IV were the hub genes connecting 10 and 9 other ARGs, respectively.

Discussion
The high diversity and abundance of the soil resistome in Alaskan tundra, temperate prairie and tropical ecosystems support the view that ARGs are naturally ubiquitous, and in widely different terrestrial ecosystems. The ARGs detected in native soils can potentially confer resistance to all major antibiotics used to treat humans and animals, such as beta-lactams ($FEZ$, $Thin-B$, and $LRA$, $PER$, $TEM$, and $OXA$ genes), macrolides and lincosamides ($erm$ genes), quinolones ($qepA$), aminoglycosides ($aac$
and aph genes), and tetracyclines (tet genes). It is not surprising that ARGs naturally exist in native soils [3] because many antibiotics are produced by soil microorganisms, and indeed were the original source of pharmaceutical products [39]. In accord with our observation, previous studies also identified divergent beta-lactamase resistance genes and a novel chloramphenicol resistance gene from undisturbed Alaskan soil [40, 41]. Vancomycin is regarded as the last line of defense against MRSA strains, but thirteen subtypes of vancomycin resistance genes were detected in these native soils, including vanH, vanA and vanX which are found in clinical pathogens S. aureus and vancomycin-resistant Enterococci [42]. Similarly, D’Costa et al. detected the three vancomycin genes in 30,000-year-old permafrost sediments, and further analyses confirmed the similarity in structure and function between the ancient vanA and their modern variants [43]. These results support the view that the primary function of ARGs in native environments is not antibiotic resistance, but other functions beneficial for microbial life [44–46]. Antibiotic concentrations are extremely low in native soils, so harboring antibiotic resistance genes is not necessary for their defense and would even be a burden for microbes, unless it has other beneficial functions. We observed significantly lower abundance of mobile element associated ARGs in Alaskan tundra soils compared to Midwest prairie and Amazon soils, perhaps due to the lower microbial diversity in Alaskan tundra soils. However, antibiotic resistance could become a primary function and be further transferred to pathogens when selection occurs. For example, some clinically relevant resistance genes have been acquired from organisms where their native function is not antibiotic resistance, and which may not even confer a resistance phenotype in their native context [44].

We defined background ARGs as those which were shared by all soils regardless of ecosystem type and geography. These ARGs were also found across various ecosystems in previous studies. Thirty of our background ARGs were detected in paddy soils [4]; ten were found in dryland (peanut) soils [10]; at least sixteen were observed in greenhouse soils [47] and ten were found in Antarctic soils [5]. Most of the background ARGs are multidrug resistance genes with efflux pump as the dominant mechanism, a trait that has other natural physiological functions and found in the cytoplasmic membrane of all bacterial cells [48]. For example, 11 of the background ARGs are involved in Mex-Opr
efflux pump systems, and they are known to play a prominent role in the multidrug resistance of gram-negative bacteria such as *Pseudomonas* [49, 50]. The AcrAB efflux pump plays a physiologic role of pumping out bile acids, fatty acids, and various toxic compounds [51, 52]. *bcrA* confers resistance to bacitracin and is predicted to be responsible for energy coupling to the transport system [53]. Thus, we argue that the background ARGs should be considered as a separate category, generally of low risk, when evaluating ARG risk in soil environments. However, this doesn’t mean that background ARGs are risk free, since some of them have been found in plasmids and can be enriched with anthropogenic activity. For example, *macB* can be easily acquired by mobile elements, and thus spread macrolide resistance [54]. Background ARGs *acrA*, *vanC*, and *mexF* were found significantly enriched by the application of sewage sludge and chicken manure to soil [55]. These results also imply that compared to the abundance of ARGs, assessment of ARG mobility may be more important for ARG risk evaluation since ARG transfer into pathogens is a primary risk factor.

The soil resistome profile had a significant geographic pattern, which was greater than land use change. The resistome profile of Amazon pasture soil was more similar to Midwest pasture soil, than Amazon rainforest soil was to Midwest pasture soil, indicating the above-ground vegetation’s influence on the soil resistome. The significant correlation ($p < 0.01$, $R^2 = 0.795$) between ARG diversity and bacterial diversity (Additional file 1: Fig. S5) suggests that the lower bacterial diversity may explain the lower ARG diversity in tundra. We found significant correlation between the resistome profile and microbial community structure, indicating that the bacterial composition shapes the ARG profile in soil. This is consistent to previous studies which also found strong correlation between ARG profile and bacterial community structure in various environmental contexts [56-58]. Thus, the variation of soil resistomes at different geographical locations was probably related to the differing plant diversity, climate, and edaphic factors such as pH and soil organic matter [59-61] which will select different populations (and hence the ARGs they carry) or some ARGs by their alternative function.

It's well known that the introduction of selective or co-selective pressure by human activities is a primarily responsible for the enrichment of ARGs in soils. For example, irrigation with reclaimed water
led to enrichment of 60 ARGs [9]. Long-term application of pig manure significantly enhanced the abundance of tetL, tetB(P), tetO, tetW, sul1, ermB, and ermF as compared with inorganic fertilizers [62]. However, it’s not clear whether the normal agricultural activities such as crop cultivation affect the soil resistome. In this study, no significant change was observed in either ARG diversity, resistome abundance, or abundances of mobile element associated ARGs after long-term continuous cultivation. However, the cultivated soils from the three Midwest sites tend to have similar resistome profiles which may be due to selection for similar adaptations of the bacterial community to grain cultivation. Cropping system type, fertilization and other soil management practices are generally thought to be other factors that can influence the soil resistome [10, 62]. In these study sites antibiotics and heavy metals were not used so external factors would not have provided for selection. norB was one of the two genes significantly changed after continuous cultivation. It confers resistance to multiple drugs but is also involved in soil denitrification. Both N fertilizer and/or manure applications were used at the temperate and tropical sites, which may have favored soil denitrification and thus enriched norB gene [63]. Overall, our results suggest that previous standard cultivation and fertilization practices of US Midwest (primarily moldboard plow, inorganic N.P.K and low levels of manure, e.g. cattle grazing) did not increase the public health risk of ARGs in soil.

Twenty-three new ARGs emerged in soil after converting rainforest to pasture. The grass vegetation (Urochloa brizantha, Urochloa decumbens, Panicum maximum) and/or cattle grazing, which includes their manures, may be responsible for the newly emerged ARGs in pasture soil by selecting different microbial populations and/or increasing their diversity (Additional file 1: Fig. S6) [14]. Vancomycin resistance gene vanA and fosfomycin resistance gene fosB were enriched in pasture soil, corroborating similar results found by Yang et al. [64]. Enterococcus, Bacillus, and Staphylococcus are important hosts of vanA and fosB, and are commonly found in animal manure [65–67]. We observed increased Enterococcus, Bacillus, and Staphylococcus in pasture soils (Additional file 1: Fig. S7), suggesting that enrichment of the two ARGs could be at least in part due to the increase of these taxa. Both enrichment and attenuation of ARGs were observed after land conversion, and we speculate that ARGs are enriched or attenuated following land use change because this selects for
microbes that happen to be ARG carriers (or not) or for an alternative function of the ARG. For example, as noted above, the enriched \textit{norB} in pasture soil may be because the grazing activity as well as the observed higher soil moisture in the rainy season (when the site was sampled) stimulated denitrification [63].

Growing evidence has shown that some ARGs in pathogens are acquired from environmental bacteria through horizontal gene transfer. For example, the CTX-M extended spectrum beta-lactamase originated from chromosomal genes of an environmental genus, \textit{Kluyvera} [68], and the clinical \textit{vanA} has been found in environmental \textit{Bacilli} [69]. Thus, we selected 12 ARGs which are clinically important and could transfer between bacteria [70] and assessed their sequence similarity to clinical ARGs in human pathogens. The ARGs we detected in Amazon soils are distinct from those found in human pathogens, implying that most ARGs in the natural soil resistome are not demonstrated as problematic. Only a few clinical-similar reads (> 90% amino acid identity) of our tested ARGs were observed, but further check with a target gene assembly method confirmed that most of them were aligned to conserved regions of these genes. For example, \textit{ampC} codes clinically important cephalosporinases which cause resistance to cephalothin, cefazolin, cefoxitin, and most penicillins, but the assembled \textit{ampC} in Amazon soils share less than 54% of similarity with those found in human pathogens. It’s noteworthy that most researchers used short-read based BLAST for ARG search, which provides a sensitive detection but will also recover non-functional pseudo genes or conserved domains. By contrast, ARG evaluation with assembled genes will miss some low abundance ARGs but should better reflect the presence, abundance and sequence similarity of potentially functional ARGs. The wide co-occurrence of ARGs observed in this study is consistent with previous findings that soil ARGs commonly exist in clusters [17, 21, 28]. \textit{mdtA}, \textit{mdtB}, and \textit{mdtC} co-occurred, consistent with the fact that they are parts of the MdtABC-TolC efflux pump system. Many background ARGs were found in ARG clusters, indicating that background ARGs can also be co-enriched when selective pressure for that population or trait emerges. For example, \textit{vanA}, \textit{vanD}, and \textit{vanX} were found in the same cluster, and they were found co-enriched in soils during conversion of Amazon rainforest to pasture. The ARG cluster analysis could improve our understanding of the co-occurrence and spread of ARGs in the soil
environment. However, the network analysis is based on a statistical method and may not reflect gene linkage or population level co-selection.

In comparison to traditional the qPCR method, the metagenomic approach is more comprehensive for assessing the entire known resistome. The ARGs-OAP approach used in this study can not only detect known ARGs but also identify ARG variants not yet recognized but might be the natural reservoir for emerging new subtypes since the Hidden Markov Models detect sequence diversity within the subclasses [27]. It is noteworthy that some regulatory genes were contained in the SARG database as well as the widely used CARD database [71]. It’s problematic as to whether regulatory genes should be counted as ARGs since they only control expression of ARGs. For example, vanR and vanS cannot confer resistance to vancomycin, but vanR can promote co-transcription of vanA, vanH, and vanX when activated by vanS [72]. A high abundance of regulatory genes was detected and they differed in soils from the several ecosystems. We removed regulatory genes from our further analyses since the potential risk of ARGs is largely from the horizontal transfer of structural genes which code for functional proteins. In addition to the regulatory genes, some ARGs are components of a functional complex, for which an individual ARG cannot code antibiotic resistance without others. For example, many ARGs detected in our study are components of mdtABC-toIC, acrAB-toIC, and mexEF-oprN efflux complexes. Thus, the addition of ARGs belonging to a complex can inflate the total resistome abundance. Granted, soil as an important reservoir of ARGs; it harbors background ARGs that may or may not become problematic, probably harbors ARGs not yet emerged, and can harbor clinical ARGs, most likely to have entered soil from human or animal waste disposal. We recommend that more attention be paid to ARG genes or gene sets necessary for resistance function, for their status relative to common ARG backgrounds, for linkage to mobile genetic elements and their correspondence or linkage to host populations. Sequence similarity may or may not be indicative of potential ARG function but it is a strong indicator of whether the ARG source was from a known clinical resistance and detectable by methods targeting the clinical gene variant.

Conclusions
ARGs are a worldwide public health concern and especially so because of possible horizontal gene
transfers to pathogens. Soil harbors ARGs that may or may not become problematic, and some that are yet to emerge. Interrogating soil DNA using a comprehensive ARG database is the most comprehensive way to reveal soil’s native and human influenced resistome. We show that the ARG reservoir in soil is global, huge, and exhibits significant geographic patterns. We define "background ARGs" as those in common in all samples of these diverse ecosystems and suggest that they be considered as a separate category for health risk evaluation. We recommend that more attention be paid to ARG genes or gene sets necessary for resistance function, for their status relative to common ARG backgrounds, for linkage to mobile genetic elements and their correspondence or linkage to host populations to evaluate risk.

Declarations

AVAILABILITY OF DATA AND MATERIALS

All sequence data used in the study is available at European Nucleotide Archive (no. PRJEB10725) and JGI (Project Ids 1077701-1077706 and 1080879-1080888).

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Not applicable.

CONSENT FOR PUBLICATION

Not applicable.

COMPETING INTERESTS

The authors declare that they have no competing interests.

FUNDING

This study was supported by Interdisciplinary Research Center for Soil Microbial Ecology and Land Sustainable Productivity in Dry Areas at Northwest A&F University and by Center for Health Impacts in Agriculture at Michigan State University.

CONTRIBUTIONS

JMT designed the experiment. XQ analyzed the data and wrote the manuscript. SG, GJR, CBL, and JRC guided the bioinformatic analyses. JMT and JG provided advice and revised this manuscript. All authors read and approved the final manuscript.

ACKNOWLEDGMENTS

We thank Professor Tong Zhang and his students Xiaole Yin and Xiaotao Jiang at The University of
Hong Kong for providing help on using ARGs-OAP pipeline. We thank Dr. Lisa Boughner for assistance with manuscript editing.

References

1.  O'Neill J. The review on antimicrobial resistance. In: Tackling drug-resistant infections globally: Final report and recommendations. London: HM Government and the Wellcome Trust; 2016.

2.  Braga LPP, Alves RF, Dellias MTF, Navarrete AA, Basso TO, Tsai SM. Vinasse fertirrigation alters soil resistome dynamics: an analysis based on metagenomic profiles. Biodata Min. 2017;10:17.

3.  Cytryn E. The soil resistome: The anthropogenic, the native, and the unknown. Soil Biol Biochem. 2013;63:18–23.

4.  Xiao KQ, Li B, Ma LP, Bao P, Zhou X, Zhang T, et al. Metagenomic profiles of antibiotic resistance genes in paddy soils from South China. FEMS Microbiol Ecol. 2016. doi:

5.  Wang F, Stedtfeld RD, Kim OS, Chai B, Yang L, Stedtfeld TM, et al. Influence of soil characteristics and proximity to Antarctic research stations on abundance of antibiotic resistance genes in soils. Environ Sci Technol. 2016;50:12621–9.

6.  Canton R. Antibiotic resistance genes from the environment: a perspective through newly identified antibiotic resistance mechanisms in the clinical setting. Clin Microbiol Infec. 2009;15:20–5.

7.  Dantas G, Sommer MOA, Oluwasegun RD, Church GM. Bacteria subsisting on antibiotics. Science. 2008;320:100–3.

8.  Wellington EMH, Boxall ABA, Cross P, Feil EJ, Gaze WH, Hawkey PM, et al. The role of the natural environment in the emergence of antibiotic resistance in gram-negative bacteria. Lancet Infect Dis. 2013;13:155–65.
9. Wang FH, Qiao M, Su JQ, Chen Z, Zhou X, Zhu YG. High throughput profiling of antibiotic resistance genes in urban park soils with reclaimed water irrigation. Environ Sci Technol. 2014;48:9079-85.

10. Wang F, Xu M, Stedtfeld RD, Sheng H, Fan J, Liu M, et al. Long-term effect of different fertilization and cropping systems on the soil antibiotic resistome. Environ Sci Technol. 2018;52:13037-46.

11. Wu D, Huang XH, Sun JZ, Graham DW, Xie B. Antibiotic resistance genes and associated microbial community conditions in aging landfill systems. Environ Sci Technol. 2017;51:12859-67.

12. Mackelprang R, Grube AM, Lamendella R, Jesus EDC, Copeland A, Liang C, et al. Response of the soil microbiome to cultivation in native tallgrass prairie soils of the Midwestern United States. Front Microbiol. 2018;9:1775.

13. Soares BS, Nepstad DC, Curran LM, Cerqueira GC, Garcia RA, Ramos CA. et a. Modelling conservation in the Amazon basin. Nature. 2006;440:520-3.

14. Rodrigues JLM, Pellizari VH, Mueller R, Baek K, Jesus ED, Paula FS, et al. Conversion of the Amazon rainforest to agriculture results in biotic homogenization of soil bacterial communities. Proc. Natl. Acad. Sci. U.S.A. 2013;110:988-993.

15. Knapp C, Callan A, Aitken B, Shearn R, Koenders A, Hinwood A. Relationship between antibiotic resistance genes and metals in residential soil samples from Western Australia. Environ Sci Pollut R. 2017;24:2484-94.

16. Sandberg KD, LaPara TM. The fate of antibiotic resistance genes and class 1 integrons following the application of swine and dairy manure to soils. FEMS Microbiol Ecol. 2016. doi:

17. Johnson TA, Stedtfeld RD, Wang Q, Cole JR, Hashsham SA, Looft T, et al. Clusters of antibiotic resistance genes enriched together stay together in swine agriculture.
18. Zhu YG, Johnson TA, Su JQ, Qiao M, Guo GX, Stedtfeld RD, et al. Diverse and abundant antibiotic resistance genes in Chinese swine farms. Proc. Natl. Acad. Sci. U.S.A. 2013;110:3435-3440.

19. McKinney CW, Dungan RS, Moore A, Leytem AB. Occurrence and abundance of antibiotic resistance genes in agricultural soil receiving dairy manure. FEMS Microbiol Ecol. 2018. doi:

20. Xie W, Yuan S, Xu M, Yang X, Shen Q, Zhang W, et al. Long-term effects of manure and chemical fertilizers on soil antibiotic resistome. Soil Biol Biochem. 2018;122:111-9.

21. Li B, Yang Y, Ma LP, Ju F, Guo F, Tiedje JM, et al. Metagenomic and network analysis reveal wide distribution and co-occurrence of environmental antibiotic resistance genes. ISME J. 2015;9:2490-502.

22. Ma L, Xia Y, Li B, Yang Y, Li LG, Tiedje JM, et al. Metagenomic assembly reveals hosts of antibiotic resistance genes and the shared resistome in pig, chicken, and human feces. Environ Sci Technol. 2016;50:420-7.

23. Arango-Argoty G, Garner E, Pruden A, Heath LS, Vikesland P, Zhang L. DeepARG: a deep learning approach for predicting antibiotic resistance genes from metagenomic data. Microbiome. 2018. doi:

24. Garner E, Chen C, Xia K, Bowers J, Engelthaler DM, McLain J, et al. Metagenomic characterization of antibiotic resistance genes in full-scale reclaimed water distribution systems and corresponding potable systems. Environ Sci Technol. 2018;52:6113-25.

25. Li L, Yin X, Zhang T. Tracking antibiotic resistance gene pollution from different sources using machine-learning classification. Microbiome. 2018;6:93.
26. Yang Y, Jiang XT, Chai BL, Ma LP, Li B, Zhang AN, et al. ARGs-OAP: online analysis pipeline for antibiotic resistance genes detection from metagenomic data using an integrated structured ARG-database. Bioinform. 2016;32:2346–51.

27. Yin XL, Jiang XT, Chai BL, Li LG, Yang Y, Cole JR, et al. ARGs-OAP v2.0 with an expanded SARG database and Hidden Markov Models for enhancement characterization and quantification of antibiotic resistance genes in environmental metagenomes. Bioinform. 2018;34:2263–70.

28. Pal C, Bengtsson-Palme J, Kristiansson E, Larsson DGJ. Co-occurrence of resistance genes to antibiotics, biocides and metals reveals novel insights into their co-selection potential. BMC Genom. 2015;16:964.

29. Howe AC, Jansson JK, Malfatti SA, Tringe SG, Tiedje JM, Brown CT. Tackling soil diversity with the assembly of large, complex metagenomes. P Natl Acad Sci USA. 2014;111:4904–9.

30. Cheng L, Zhang NF, Yuan MT, Xiao J, Qin YJ, Deng Y, et al. Warming enhances old organic carbon decomposition through altering functional microbial communities. ISME J. 2017;11:1825–35.

31. Johnston ER, Rodriguez-R LM, Luo C, Yuan MM, Wu L, He Z, et al. Metagenomics reveals pervasive bacterial populations and reduced community diversity across the Alaska tundra ecosystem. Front Microbiol. 2016;7:579.

32. Natali SM, Schuur EAG, Trucco C, Pries CEH, Crummer KG, Lopez AFB. Effects of experimental warming of air, soil and permafrost on carbon balance in Alaskan tundra. Glob Chang Biol. 2011;17:1394-407.

33. Cox MP, Peterson DA, Biggs PJ. SolexaQA. At-a-glance quality assessment of Illumina second-generation sequencing data. BMC Bioinform. 2010;11:485.

34. Thompson LR, Sanders JG, McDonald D, Amir A, Ladau J, Locey KJ, et al. A communal
catalogue reveals Earth’s multiscale microbial diversity. Nature. 2017;551:457–63.

35. Wattam AR, Davis JJ, Assaf R, Boisvert S, Brettin T, Bun C, et al. Improvements to PATRIC, the all-bacterial bioinformatics database and analysis resource center. Nucleic Acids Res. 2017;45:535–42.

36. Wang Q, Fish JA, Gilman M, Sun Y, Brown CT, Tiedje JM, et al. Xander: employing a novel method for efficient gene-targeted metagenomic assembly. Microbiome. 2015;3:32.

37. Guo J, Quensen J, Sun Y, Wang Q, Brown CT, Cole JR, et al. Review, evaluation and directions for gene-targeted assembly for ecologic analyses of metagenomes. Front Genet. 2019;10:957.

38. Guo JR, Cole JR, Zhang QP, Brown CT, Tiedje JM. Microbial community analysis with ribosomal gene fragments from shotgun metagenomes. Appl Environ Microbiol. 2016;82:157–66.

39. Davies J, Davies D. Origins and evolution of antibiotic resistance. Microbiol Mol Biol R. 2010;74:417–33.

40. Allen HK, Moe LA, Rodbumrer J, Gaarder A, Handelsman J. Functional metagenomics reveals diverse beta-lactamases in a remote Alaskan soil. ISME J. 2009;3:243–51.

41. Lang KS, Anderson JM, Schwarz S, Williamson L, Handelsman J, Singer RS. Novel florfenicol and chloramphenicol resistance gene discovered in Alaskan soil by using functional metagenomics. Appl Environ Microbiol. 2010;76:5321–6.

42. Wright GD. The antibiotic resistome: The nexus of chemical and genetic diversity. Nat Rev Microbiol. 2007;5:175–86.

43. D’Costa VM, King CE, Kalan L, Morar M, Sung WWL, Schwarz C, et al. Antibiotic resistance is ancient. Nature. 2011;477:457–61.

44. Dantas G, Sommer MO. Context matters—the complex interplay between resistome
genotypes and resistance phenotypes. Curr Opin Microbiol. 2012;15:577–82.

45. Gillings MR. Evolutionary consequences of antibiotic use for the resistome, mobilome, and microbial pangenome. Front Microbiol. 2013;4:4.

46. Perry JA, Westman EL, Wright GD. The antibiotic resistome: What's new? Curr Opin Microbiol. 2014;21:45–50.

47. Fang H, Wang H, Cai L, Yu Y. Prevalence of antibiotic resistance genes and bacterial pathogens in long-term manured greenhouse soils as revealed by metagenomic survey. Environ Sci Technol. 2015;49:1095–104.

48. Spengler G, Kincses A, Gajdacs M, Amaral L. New roads leading to old destinations: Efflux pumps as targets to reverse multidrug resistance in bacteria. Mol. 2017;22:468.

49. Aendekerk S, Diggle SP, Song Z, Hoiby N, Cornelis P, Williams P, et al. The MexGHI-OpmD multidrug efflux pump controls growth, antibiotic susceptibility and virulence in *Pseudomonas aeruginosa* via 4-quinolone-dependent cell-to-cell communication. Microbiol. 2005;151:1113–25.

50. Li Y, Mima T, Komori Y, Morita Y, Kuroda T, Mizushima T, et al. A new member of the tripartite multidrug efflux pumps, MexVW-OprM, in *Pseudomonas aeruginosa*. J Antimicrob Chemoth. 2003;52:572–5.

51. Okusu H, Ma D, Nikaido H. AcrAB efflux pump plays a major role in the antibiotic resistance phenotype of *Escherichia coli* multiple-antibiotic-resistance (Mar) mutants. J Bacteriol. 1996;178:306–8.

52. Yu EW, McDermott G, Zgurskaya HI, Nikaido H, Koshland DEJ. Structural basis of multiple drugbinding capacity of the *AcrB* multidrug efflux pump. Science. 2003;300:976–80.

53. Podlesek Z, Comino A, Herzog-Velikonja B, Zgr-Bertok D, Komel R, Grabnar M.
Bacillus licheniformis bacitracin-resistance ABC Transporter: relationship to mammalian multidrug resistance. Mol Microbiol. 1995;16:969-76.

54. Roberts MC. Update on macrolide-lincosamide-streptogramin, ketolide, and oxazolidinone resistance genes. FEMS Microbiol Lett. 2008;282:147-59.

55. Chen QL, An XL, Li H, Su JQ, Ma YB, Zhu YG. Long-term field application of sewage sludge increases the abundance of antibiotic resistance genes in soil. Environ Int. 2016;92-93:1-10.

56. Gibson MK, Forsberg KJ, Dantas G. Improved annotation of antibiotic resistance determinants reveals microbial resistomes cluster by ecology. ISME J. 2015;9:207-16.

57. Guo N, Wang YK, Tong TZ, Wang SG. The fate of antibiotic resistance genes and their potential hosts during bio-electrochemical treatment of high-salinity. Water Res. 2018;133:79-86.

58. Qian X, Sun W, Gu J, Wang XJ, Sun JJ, Yin YN, et al. Variable effects of oxytetracycline on antibiotic resistance gene abundance and the bacterial community during aerobic composting of cow manure. J Hazard Mater. 2016;315:61-9.

59. Fierer N, Jackson RB. The diversity and biogeography of soil bacterial communities. P Natl Acad Sci USA. 2006;103:626-31.

60. Nottingham AT, Fierer N, Turner BL, Whitaker J, Ostle NJ, McNamara NP, et al. Microbes follow Humboldt: temperature drives plant and soil microbial diversity patterns from the Amazon to the Andes. Ecol. 2018;99:2455-66.

61. Prober SM, Leff JW, Bates ST, Borer ET, Firn J, Harpole WS, et al. Plant diversity predicts beta but not alpha diversity of soil microbes across grasslands worldwide. Ecol Lett. 2015;18:85-95.

62. Peng S, Feng YZ, Wang YM, Guo XS, Chu HY, Lin XG. Prevalence of antibiotic resistance genes in soils after continually applied with different manure for 30 years.
63. Hofstra N, Bouwman AF. Denitrification in agricultural soils: summarizing published data and estimating global annual rates. Nutr Cycl Agroecosys. 2005;72:267-78.

64. Cetinkaya Y, Falk P, Mayhall CG. Vancomycin-resistant Enterococci. Clin Microbiol Rev. 2000;13:686-707.

65. Thompson MK, Goodman M, Keithly M, Hammer N, Cook P, Jagessar K, et al. Structure and function of the genomically-encoded fosfomycin resistance enzyme, FosB, from Staphylococcus aureus. Biophys J. 2014;106:755-65.

66. Thompson MK, Keithly ME, Harp J, Cook PD, Jagessar KL, Sulikowski GA, et al. Structural and chemical aspects of resistance to the antibiotic fosfomycin conferred by fosB from Bacillus cereus. Biochem. 2013;52:7350-62.

67. Poirel L, Kampfer P, Nordmann P. Chromosome-encoded Ambler class A beta-lactamase of Kluyvera georgiana, a probable progenitor of a subgroup of CTX-M extendedspectrum beta-lactamases. Antimicrob Agents Chemother. 2002;46:4038-40.

68. Patel R, Piper K, Cockerill FR, Steckelberg JM, Yousten AA. The biopesticide Paenibacillus popilliae has a vancomycin resistance gene cluster homologous to the Enterococcal VanA vancomycin resistance gene cluster. Antimicrob Agents Chemother. 2000;44:705-9.

69. van Hoek AH, Mevius D, Guerra B, Mullany P, Roberts AP, Aarts HJ. Acquired antibiotic resistance genes: an overview. Front Microbiol. 2011;2:203.

70. Jia BF, Raphenya AR, Alcock B, Waglechner N, Guo PY, Tsang KK, et al. Card 2017: Expansion and model-centric curation of the comprehensive antibiotic resistance database. Nucleic Acids Res. 2017;45:566-73.

71. Arthur M, Molinas C, Courvalin P. The VanS-VanR two-component regulatory system
controls synthesis of depsipeptide peptidoglycan precursors in \textit{Enterococcus faecium} BM4147. J Bacteriolo. 1992;174:2582–91.

Figures

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure1.png}
\caption{Composition of ARGs and regulator genes in 26 soil metagenomes. a Resistance mechanism, b Antibiotic classes.}
\end{figure}
Figure 2

Heatmap showing abundances of 55 background ARGs in 26 soil metagenomes. MLS: Macrolides-lincosamides-streptogramines.
Figure 3

Diversity and abundance and of ARGs among soils of three ecosystems. a ARG diversity of different mechanisms, b ARG abundance of different mechanisms, c ARG abundance of different antibiotic classes. MLS: Macrolides-lincosamides-streptogramines.
Abundance of selected ARGs among soils of the three ecosystems. 

a Efflux pump complex, 

b vancomycin resistance genes, c ARGs commonly found in mobile elements.
Soil resistome profiles. a PCoA analysis showing profiles of soil resistomes in three ecosystems. Circles represent native soils, and triangles represent anthropogenic soils, b Procrustes analysis of bacterial community and resistome profile.
Figure 6

Abundance of 12 clinical ARGs in Amazon rainforest and pasture soils. The blue column represents the ARG abundance quantified with SARG database. The yellow and red columns are abundances of clinical ARGs at 80%, 90%, 95% and 97% amino acid identities. Bars are standard errors.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.
AdditionalFile1.docx
