



OPEN

DATA DESCRIPTOR

A database on the abundance of environmental antibiotic resistance genes

Wenjuan Xu^{1,2,6}, Zhizhen Pan^{1,6}, Yangyu Wu¹, Xin-Li An¹, Weiyi Wang³, Boris Adamovich⁴,
Yong-Guan Zhu¹, Jian-Qiang Su¹✉ & Qiansheng Huang^{1,5}✉

Antimicrobial resistance (AMR) poses a severe threat to global health. The wide distribution of environmental antibiotic resistance genes (ARGs), which can be transferred between microbiota, especially clinical pathogens and human commensals, contributed significantly to AMR. However, few databases on the spatiotemporal distribution, abundance, and health risk of ARGs from multiple environments have been developed, especially on the absolute level. In this study, we compiled the ARG occurrence data generated by a high-throughput quantitative PCR platform from 1,403 samples in 653 sampling sites across 18 provinces in China. The database possessed 291,870 records from five types of habitats on the abundance of 290 ARGs, as well as 8,057 records on the abundance of 30 mobile genetic elements (MGEs) from 2013 to 2020. These ARGs conferred resistance to major common types of antibiotics (a total of 15 types) and represented five major resistance mechanisms, as well as four risk ranks. The database can provide information for studies on the dynamics of ARGs and is useful for the health risk assessment of AMR.

Background & Summary

Antimicrobial resistance (AMR) has been recognized as a serious threat to global public health^{1,2}. A global study published in *The Lancet* showed that AMR contributed directly to about 1.27 million deaths in 2019³. A report entrusted by the UK Government estimated that AMR may carry off 10 million people each year by 2050⁴. The increasing AMR risk is largely attributed to the spread of antibiotic resistance genes (ARGs), which bacteria can acquire to develop resistance to antibiotics^{5,6}. ARGs present in the environment can be horizontally and vertically transferred within the microbial community, thus facilitating the spread of AMR^{6–8}. ARGs confer resistance to a variety of antibiotics, including aminoglycoside, tetracycline, macrolide, chloramphenicol, glycopeptide, sulfonamide, trimethoprim, beta-lactams, fluoroquinolone, quinolones and so on^{9,10}. Six major resistance mechanisms were found in the antibiotic mechanism of ARGs, including reducing intracellular antibiotic concentrations by efflux pumps, inactivating the antibiotics by hydrolysis or modification, altering, replacing, or protecting the antibiotic target, as well as cellular protection with poor permeability of antibiotics^{11–13}.

Actually, ARGs have already existed in environmental bacteria for millions of years, predating the discovery and use of antibiotics. The presence of several ARGs were found from the ancient layers of the Canadian high Arctic permafrost¹⁴. Anthropogenic activities, particularly the overuse and misuse of antibiotics in agriculture, human medicine, and animal husbandry, have accelerated the enrichment and dissemination of ARGs^{15–18}. As emerging contaminants, ARGs have been globally detected in a diversity of environmental compartments, including water, soil, and air^{19–27}. The abundance of ARGs varied across different regions, whereas hotspots are found in antibiotic-enriched areas, including hospitals, antibiotic wastewater and farms^{16,28,29}. It has been confirmed that the ARG abundance increased with the anthropogenic activity³⁰. With the aid of mobile genetic elements (MGEs), ARGs can be horizontally transferred from environmental antibiotic-resistant bacteria to clinical pathogens, further exacerbating the risks to human health^{30,31}. Consequently, the risk assessment of

¹Xiamen Key Laboratory of Indoor Air and Health, Key Lab of Urban Environment and Health, Institute of Urban Environment, Chinese Academy of Sciences, Xiamen, 361021, China. ²College of Resources and Environment, University of Chinese Academy of Sciences, Beijing, 100049, China. ³College of Life Sciences, Institute of Life Sciences and Green Development, Hebei University, Baoding, 071002, China. ⁴Research Laboratory of Aquatic Ecology, Belarusian State University, Minsk, 220030, Belarus. ⁵National Basic Science Data Center, Beijing, 100190, China. ⁶These authors contributed equally: Wenjuan Xu, Zhizhen Pan. ✉e-mail: jqsu@iue.ac.cn; qshuang@iue.ac.cn

ARGs has become a global urgency, where the primary step is to reveal the occurrence, distribution, and abundance of ARGs in the environment.

Currently, metagenomic analysis and quantitative PCR are two popular approaches for detecting the composition and abundance of ARGs^{28,32}. The profiles of ARGs can be predicted using metagenomic sequencing data annotated by blasting against the reference database, such as Antibiotic Resistance Genes Database³³, the Comprehensive Antibiotic Research Database³⁴ and the Structured Antibiotic Resistance Gene³⁵. For instance, Zheng, *et al.*³⁶ generated a global abundance map of 558 ARGs in the soil using metagenomic sequencing. Hendriksen, *et al.*³⁷ use metagenomic analysis of sewage in 60 countries for the prediction of AMR. This method has the capacity to provide a comprehensive overview of ARG profiles in a specific sample. In contrast, quantitative PCR, especially high-throughput quantitative PCR (HT-qPCR), showed better detection limits, lower cost, reduced sample quantity requirement, and the ability for absolute quantification^{38,39}. Willms, *et al.*⁴⁰ utilized real-time quantitative PCR to quantify medically relevant ARGs and MGEs in the soil in Germany, and Delgado-Baquerizo, *et al.*⁴¹ generated a global atlas of soil ARG abundance detected by HT-qPCR. The HT-qPCR method was established in our lab and widely used^{22,42,43}. With this technology, large amounts of data about the levels of ARGs in the environment were produced.

To the best of our knowledge, there are few databases gathering comprehensive information on the distribution and abundance of ARGs from various environments in a time- and space-dependent way, especially regarding the absolute levels of ARGs. In this study, a database was provided on the occurrence of environmental ARGs through HT-qPCR in China from 2013 to 2020. In detail, the data on the relative and absolute abundance of ARGs and MGEs from five types of habitats (aquatic, edaphic, sedimentary, dusty, and atmospheric environments) were presented in a spatiotemporal way. This database can serve as a valuable resource to provide fundamental data for facilitating the understanding of the diversity and abundance of environmental ARGs, as well as for evaluating the health risk of ARGs.

Methods

Sample collection. The soil samples were collected from the top 20 cm of soil by taking, gathering, and mixing several soil cores at each sampling site after plant residue and stones had been removed⁴³. Water samples were aseptically taken from 10 to 20 cm below the water surface using sterilized containers^{44,45}. Sediment samples were collected from the top 15 cm of sediment and stored in well-sealed and sterile plastic bags²². Particulate matter (PM_{2.5} and PM₁₀) samples were pooled using the portable atmospheric particulate matter samplers with PM_{2.5} and PM₁₀ fractionating inlets, respectively, and the samples were enriched with quartz microfiber filters⁴⁶. These PM samples were described as being collected from the air habitat in this study. Dust was collected from outdoor and indoor settings with a clean sterile brush⁴⁶. All the samples were promptly frozen, transported to the laboratory, and stored at 4 °C or –20 °C until next process.

Determination of gene abundance. The DNA was extracted from samples using commercial DNA extraction Kit following the manufacturer's instructions. Then the abundance of ARGs and MGEs were both analysed by HT-qPCR with the SmartChip Real-time PCR system (Warfegen Inc. USA)^{22,42}. A total of 414 primer pairs were used to target 290 subtypes of ARGs, 16 transposases, 6 plasmids, 5 insertion sequences, 3 integrases and the 16S rRNA gene (Table S1). To verify the specificity of each primer set, a non-template negative control was performed. The thermal cycle consisted of an initial denaturation at 95 °C for 10 min, followed by 40 cycles of denaturation at 95 °C for 30 s and annealing at 60 °C for 30 s, finally with a melting curve analysis. The PCR reaction was conducted in triplicate for each primer set for each sample. The detection limit was set up at the threshold cycle (*Ct*) lower than 31. If the value of *Ct* exceeded 31, the corresponding data about the gene abundance would be regarded as 0. For each sample, only data with more than two repeated technical replicates above the detection limit were regarded as positive⁴⁷. The gene copy number was calculated by Eq. (1), and the relative abundance of each gene represented the ratio of its copy number to 16S rRNA gene copy number (Eq. (2))⁴⁸. The level of 16S rRNA was widely used to show the abundance of bacteria. Its absolute copy number was determined by plotting a standard curve using the Roche 480 system. The standard curve was obtained by a series of 10-fold dilutions 10-fold dilution of a standard plasmid containing a cloned and sequenced 16S rRNA gene fragment (1.39 × 10¹⁰ copies/L). The PCR reaction was run in triplicate with negative and positive controls. The absolute abundance of ARGs was calculated by multiplying with 16S rRNA gene absolute copy number as shown in Eq. (3)⁴⁷.

$$\text{Gene copy number} = 10^{(31 - Ct)/(10/3)} \quad (1)$$

$$\text{Gene relative abundance} = \frac{\text{Gene copy number}}{\text{16S rRNA gene copy number}} \quad (2)$$

$$\text{Gene absolute abundance} = \text{Gene relative abundance} \times \text{16S rRNA gene absolute copies} \quad (3)$$

Data processing. The data were obtained from our and the collaborator's laboratories. We reclassified the ARGs into 15 types according to the antibiotics type to which they conferred resistance, including aminoglycoside, beta-lactams, fluoroquinolone, glycopeptide, macrolide-lincosamide-streptogramin B (MLS_B), sulfonamide, tetracycline, chloramphenicol, trimethoprim, multidrug, rifamycin, fosfomycin, bacitracin, polymyxin and others. Furthermore, all the genes were sorted into a total of 290 different subtypes. As shown in Fig. 1a, a total of 290 ARG subtypes were detected in environmental samples, and the number of ARGs detected in each

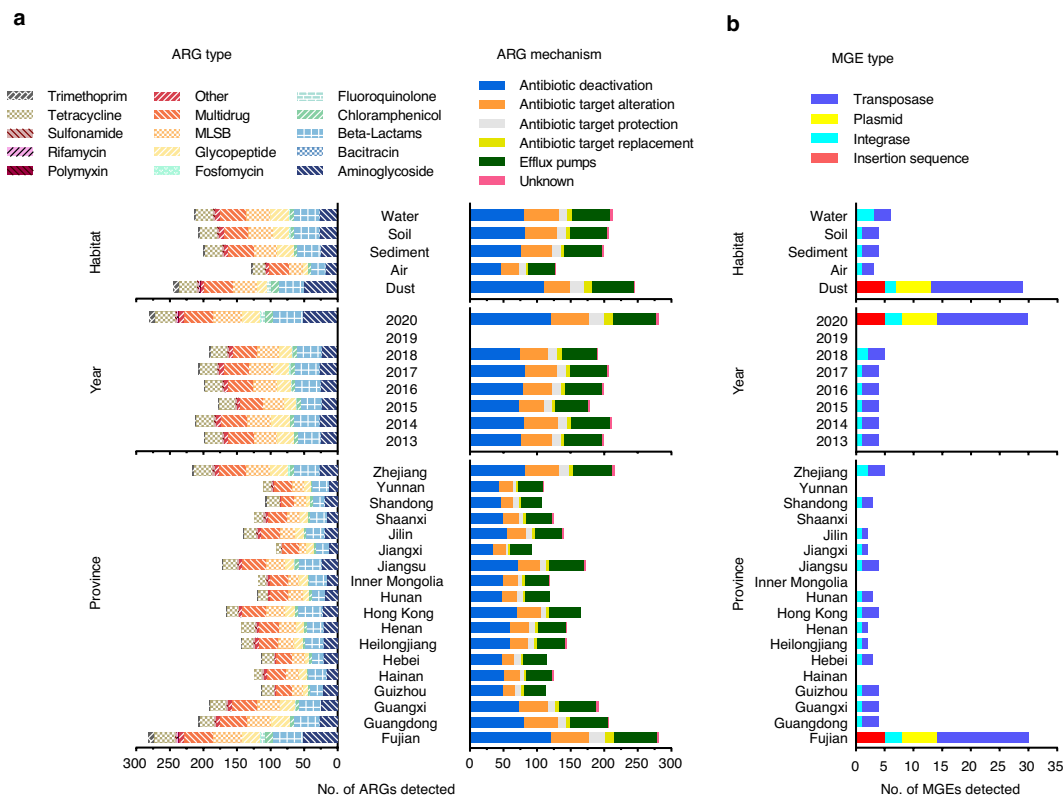


Fig. 1 The number of ARGs (a) and MGEs (b) detected in each habitat, each year from 2013 to 2020 and each province of China. ARGs were classified based on antibiotics to which they conferred resistance and resistance mechanism, respectively.

environment ranged from 128 to 245, with an average of 198. Multidrug, MLSB, and beta-lactams resistance genes were the major types, followed by aminoglycoside, tetracycline, and glycopeptide resistance genes in each habitat. In terms of the time series, the number of ARGs detected from 2013 to 2018 and in 2020 ranged from 178 to 281, with an average of 183. The top six most detected types of ARGs (multidrug, MLSB, beta-lactams, tetracycline, aminoglycoside, and glycopeptide resistance genes) every year were the same in every habitat. Regarding the spatial scale, the number of ARGs detected in each province ranged from 92 to 282, with an average of 149. The dominant types of ARGs in each province were the same as mentioned above. In addition, a total of 30 MGEs (5 insertion sequences, 6 plasmids, 3 integrase genes, and 16 transposase genes) were detected and the highest number of MGEs was detected in the dust samples with all four types detectable. In the air, sediment, and water samples, integrase genes and transposase genes were detected. In a temporal scale, the highest number of detected MGEs occurred in 2020, encompassing all four types, whereas only integrase genes and transposase genes were detected in the other years. In terms of spatial distribution, the highest number of detected MGEs was in Fujian, and all four types were involved, while only two types, integrase genes and transposase genes, were detected in other provinces (Fig. 1b). If the genes detected in the same sample belonged to the same subtype, we would regard the sum of their abundance data as the abundance of that subtype. For example, the sum of the abundance of *acrA-01*, *acrA-02*, *acrA-03*, *acrA-04*, and *acrA-05* is regarded as the abundance of *acrA*.

The ARGs were also divided into six groups based on the mechanism of resistance: antibiotic deactivation, efflux pumps, antibiotic target alteration, antibiotic target protection, antibiotic target replacement, and unknown. Antibiotic deactivation, antibiotic target alteration, and efflux pumps were the three most dominant resistance mechanisms (Fig. 1a). In addition, the health risk of ARGs was categorized into four ranks according to a previous study⁴⁹. Rank I, as the highest risk, refers to mobile ARGs that are present in ESKAPE (*Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, and *Enterobacter* species) pathogens. Rank II indicates mobile ARGs that are not yet present in pathogens. Rank III reflects ARGs that are human-associated but not associated with MGEs, while Rank IV, as the lowest risk, denotes ARGs that are not human-associated. We identified the health risk rank of ARGs based on the subtype. When a subtype possessed more than one rank, the highest risk rank would be assigned to this subtype. In this database, among all the ARGs, 47 subtypes were identified as Rank I ARGs, 9 as Rank II ARGs, 53 as Rank III ARGs, and 79 as Rank IV ARGs. Rank I and Rank II ARGs, as the current and future threats, respectively, deserve public attention. The number of the Rank I ARGs detected in each habitat ranged from 25 to 44. They consisted of 11 types of ARGs, with MLSB and beta-lactams resistance genes being the top two types. For Rank II ARGs, 5 to 7 genes belonging to five types were detected in each environment. Among 18 provinces,

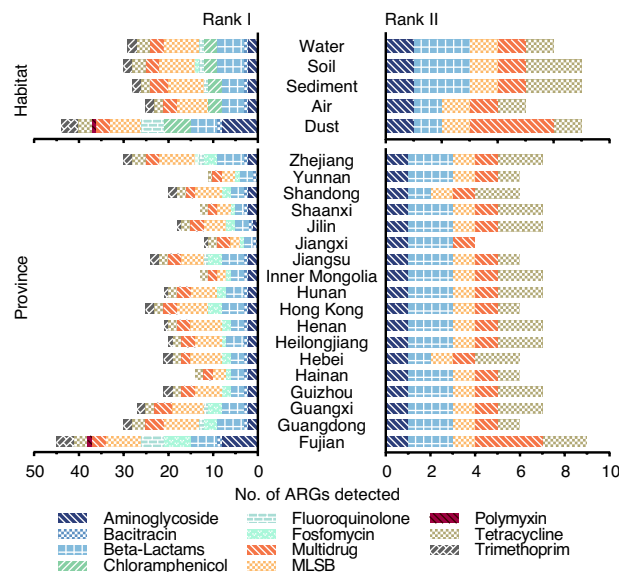


Fig. 2 The number of detected risk health Rank I and II ARGs. ARGs were classified based on antibiotics to which they conferred resistance.

the number of detected Rank I ARGs ranged from 11 to 45, and the top two dominant ARGs were MLSB and beta-lactams resistance genes. The number of detected Rank II ARGs ranged from 4 to 9 (Fig. 2).

Data Records

The data can be downloaded in the form of Excel spreadsheets (.xlsx) from <https://doi.org/10.57760/sciencedb.09803>⁵⁰ and the online platform is available at <http://sdc.iue.ac.cn/abundance/arg/list>.

The database consists of two separate tables, providing the abundance of ARGs and MGEs with detailed gene information and sample information, respectively. Each gene was assigned a unique Gene ID. For gene information, the resistance type, resistance mechanism, and health risk rank were introduced to classify different ARG subtypes. The types of MGEs are also provided. Each sample has a unique sample ID, which in turn can be related to a source ID. The source ID can be linked to the information about the publications using the data in the database. The location ID was also introduced, allowing distinguishing samples from different sites. The information about the geographic location of samples comprises the country, province, longitude, and latitude. All the information that is necessary to identify and retrace a sample, including the IDs assigned by us, the original ID used in the original studies, sampling date, sampling sites, and habitat, is provided. An explanatory note can be downloaded together with the database. An online platform with a user-friendly interface has been developed and can provide search service to obtain profiles and abundance of ARGs in a specific period and location.

The database includes 291,870 records on ARGs, and 8,057 records on MGEs. For ARGs, there were 203,138 records from soil habitat, 48,354 records from water habitat, 19,476 records from dust habitat, 14,910 records from sediment habitat, and 5,992 records from air habitat. For MGEs, there were 3,844 records from soil habitat, 694 records from water samples, 3,071 records from dust habitat, 280 records from sediment habitat, and 168 records from air habitat (Fig. 3a). As shown in Fig. 3b, the ARG abundance data for 2017 were the largest ($n = 185,376$), followed by 2018 ($n = 26,750$) and 2020 ($n = 20,974$), while the MGEs abundance data for 2017 were the largest ($n = 3,512$), followed by 2020 ($n = 3,113$) and 2018 ($n = 584$). As shown in Fig. 3c and Fig. 4, there were 653 sampling sites across 18 provinces in China, mainly in the eastern region, which is regarded as the hotspot region of antibiotic pollution^{51–53}. The top four provinces with the highest number of ARG abundance data were Zhejiang ($n = 130,472$), Fujian ($n = 34,312$), Guizhou ($n = 29,610$) and Heilongjiang ($n = 17,430$). For MGEs, abundance data in Fujian ($n = 3,365$) and Zhejiang ($n = 2,276$) were the largest, more than four times the number in other provinces (Fig. 3c). The ARG abundance records cover 15 types of ARGs. Among them, the multidrug ($n = 54,542$), beta-lactams ($n = 50,588$), MLSB resistance genes ($n = 49,375$), aminoglycoside ($n = 37,508$), tetracycline ($n = 36,011$), glycopeptide ($n = 33,235$) had the highest number of the records, and they were most studied in each habitat, each year or each province (Fig. 3d,g). Four types of MGEs were included. The number of abundance data about transposase genes, integrase genes, plasmids, and insertion sequences were 5,356, 1,568, 618, and 515, respectively, and transposase genes were most studied in each habitat, each year or each province (Fig. 3d,g). The ARGs contain six antibiotic mechanisms. Antibiotic deactivation ($n = 112,698$), efflux pumps ($n = 79,248$), and antibiotic target alteration ($n = 66,614$) had the major number of the abundance data (Fig. 3e). In addition, among four health risk ranks of ARGs, Rank IV ARGs ($n = 78,303$) had the largest number of the abundance data, followed by Rank III ARGs ($n = 54,091$) and Rank I ARGs ($n = 42,287$). Rank II ARGs ($n = 8,911$) had the lowest number of the abundance data (Fig. 3f).

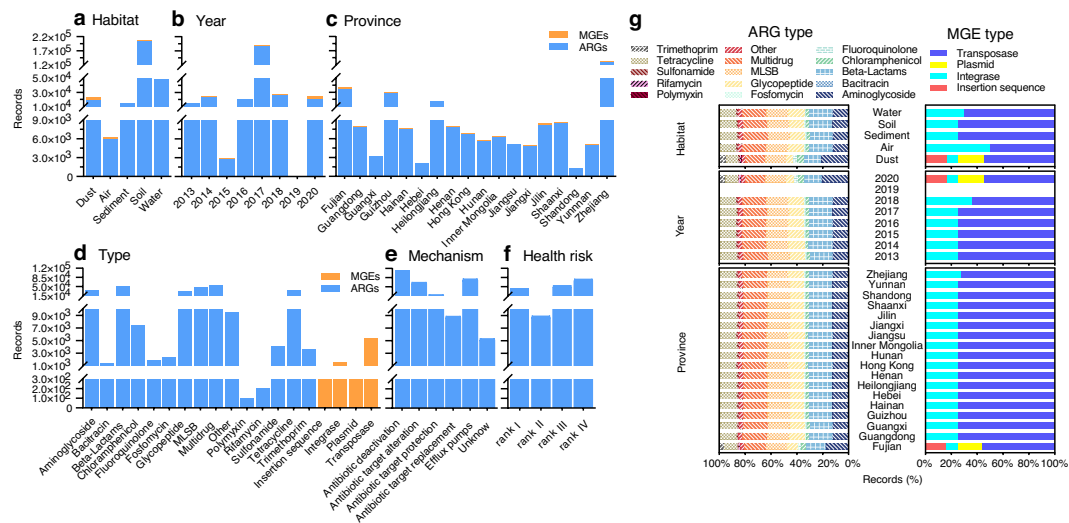


Fig. 3 The abundance records of ARGs and MGEs in the database. The number of abundance records across habitats (a), years (2013–2020) (b), and provinces in China (c). The number of abundance records across ARG and MGE types (d), and ARG mechanisms (e). (f) The number of ARG abundance records across four health risk ranks. (g) The proportion of abundance records for ARG and MGE type in each habitat, each year and each province of China.

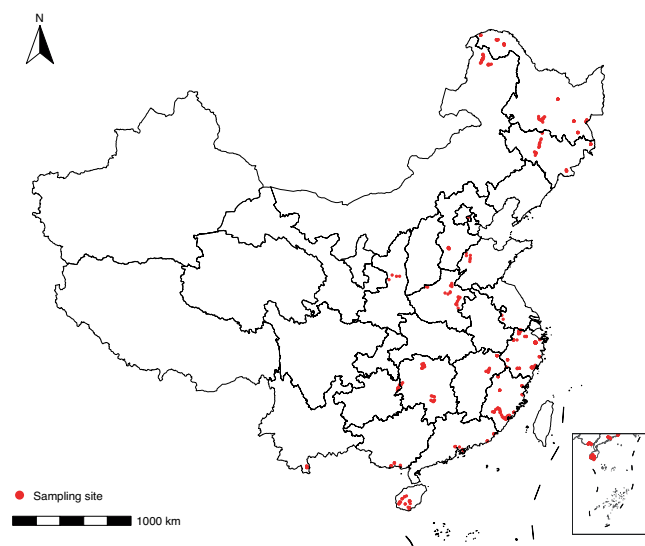


Fig. 4 Map of sampling areas across China. The red dots indicate the sampling sites.

Technical Validation

All the data were uniformly generated using HT-qPCR under the same standard operational protocols, thus showing high comparability among samples. For each gene, at least two replicates producing the threshold cycle lower than 31 were obtained. With this technology, large amounts of data were produced. Approximately 48.63% of the data were applied to the research paper published in *Nature Microbiology*, *Environmental International*, and other journals^{22,42–46,54,55}.

Abundance of ARGs. The relative abundance and absolute abundance of ARGs are available for the five habitats in the database, except for soil which only the relative abundance data can be obtained. The abundance of ARGs in various habitats is shown in Fig. 5. The average absolute abundance of ARGs in dust, air, sediment, and water were 6.22×10^8 copies/g, 4.11×10^3 copies/L, 2.67×10^7 copies/g, 5.61×10^{12} copies/L, respectively. The median absolute abundance of ARGs in dust, air, sediment, and water were 9.86×10^7 copies/g, 3.92×10^2 copies/L, 3.48×10^6 copies/g, 4.54×10^{10} copies/L respectively (Fig. 5a). The proportion of absolute abundance of ARG types was different in each habitat. For example, in the sediment habitat, multidrug and aminoglycoside resistance genes accounted for 43.64% and 32.40%, respectively, while in dust habitat, aminoglycoside (24.27%),

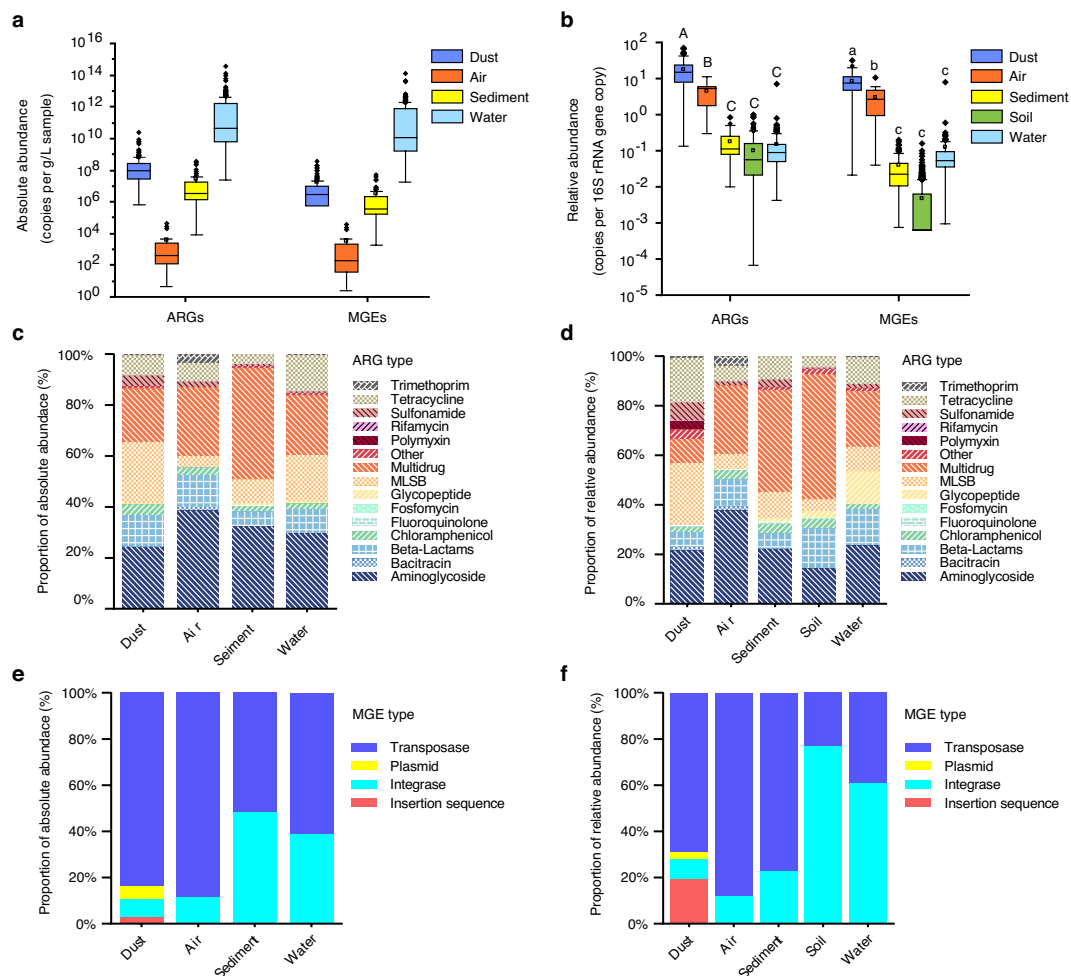


Fig. 5 The abundance of ARGs and MGEs in various habitats. Absolute abundance (a) and relative abundance (b) of ARGs and MGEs in various habitats. The middle horizontal line and square represent the median and mean, the box denotes the 25% and 75% percentiles, the whiskers refer to the range within $1.5 \times \text{IQR}$ (interquartile range), data beyond the whiskers are outlying points. Different letters in (b) indicate statistically significant differences among different habitats ($P < 0.05$, one-way ANOVA with Tukey's test). Capital letters are for ARGs and lowercase letters are for MGEs. The proportion of absolute abundance (c) and relative abundance (d) of ARG types. The proportion of absolute abundance (e) and relative abundance (f) of MGE types.

MLSB (24.18%) and multidrug resistance genes (20.87%) were the most abundant ARGs (Fig. 5c). To compare the relative abundance of ARGs among different habitats, one-way ANOVA with Tukey B post hoc analysis was conducted. The difference was considered statistically significant at a P value less than 0.05. The dust had the highest relative abundance of total ARGs (average = 18.00 copies per 16S rRNA gene copy, median = 15.02 copies per 16S rRNA gene copy), followed by atmospheric environment (average = 4.54 copies per 16S rRNA gene copy, median = 5.21 copies per 16S rRNA gene copy) (Fig. 5b). To be noted, the major dust samples were collected from the hospital, which could be the reason for relative high abundance of ARGs. The relative composition of different types of ARGs varied from one habitat to another. For example, genes conferring multidrug resistance (50.63%) dominated in the soil environment, while ARGs conferring resistance to aminoglycoside (37.62%) had a high proportion in the atmospheric environment (Fig. 5d).

As shown in Fig. 6a, the relative abundance of total ARGs ranged from 7.40×10^{-3} to 11.30 copies per 16S rRNA gene copy across 18 provinces in China. The map showed that the average relative abundance of ARGs in Fujian was the highest, 2–4 orders of magnitude higher than in other provinces. This may be because most of the dust samples were collected in Fujian. Among 18 provinces, Fujian, Jiangsu, Zhejiang, Heilongjiang, Shaanxi, Inner Mongolia, and Yunnan were the top seven provinces with the highest relative abundance of ARGs, while Shandong, Hebei, Hunan, Guizhou, Hongkong, and Jiangxi had the lowest abundance. For the relative composition of ARGs, as shown in Fig. 6b, in most provinces, the multidrug resistance genes were the dominant ARGs (above 42.41%), while the glycopeptide resistance genes were the major ARGs in Jiangsu (37.40%). Aminoglycoside and multidrug resistance genes dominated in Zhejiang, Hong Kong, and Guangdong (60.84%, 53.95%, and 60.22%, respectively). The genes conferring resistance to MLSB (24.37%), aminoglycoside (21.06%), and tetracycline (17.59%) were the top three ARGs in Fujian. Figure 7 displays the relative abundance

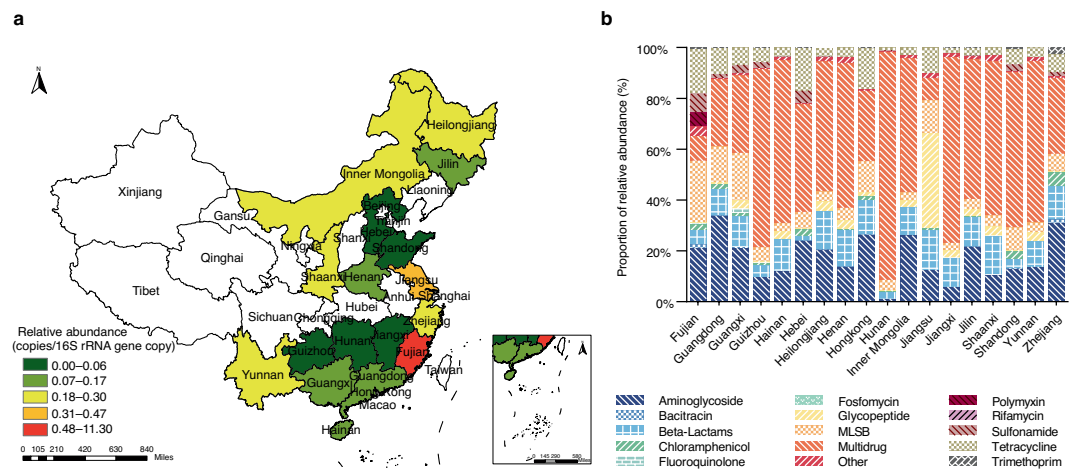


Fig. 6 Distribution of ARGs in China. **(a)** Spatial distribution of total ARGs relative abundance in China. **(b)** Relative abundance of ARG types (%) in China.

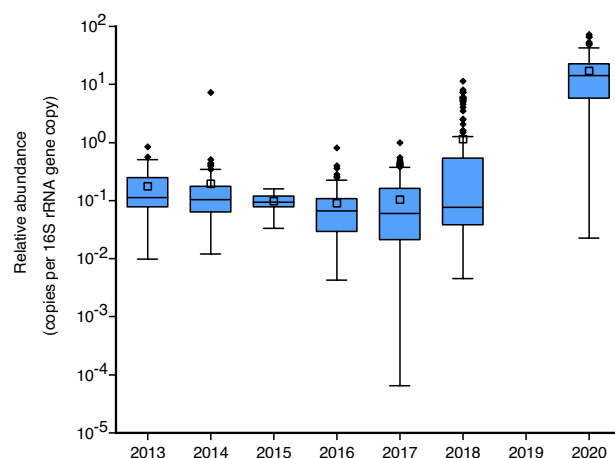


Fig. 7 Relative abundance of total ARGs from 2013 to 2020. The middle horizontal line and square represent the median and mean, the box denotes the 25% and 75% percentiles, the whiskers refer to the range within $1.5 \times \text{IQR}$, data beyond the whiskers are outlying points.

of total ARGs from 2013 to 2020, ranging from 9.09×10^{-2} to 16.94 copies per 16S rRNA gene copy. The average relative abundance of ARGs in 2018 and 2020 was higher than in other years, to a great extent, due to the dust and air samples which were obtained in 2018 and 2020.

Abundance of MGEs. Both the relative abundance and absolute abundance of MGEs are available for the five habitats in the database, except soil which only the relative abundance data can be obtained. The average absolute abundance in dust, air, sediment, and water were 2.01×10^8 copies/g, 3.53×10^3 copies/L, 3.35×10^6 copies/g, 1.96×10^{12} copies/L, respectively. The median absolute abundance of MGEs in dust, air, sediment, and water were 4.43×10^7 copies/g, 2.02×10^2 copies/L, 3.82×10^5 copies/g, 1.17×10^{10} copies/L, respectively (Fig. 5a). Transposase genes accounted for 83.62%, 88.51% and 61.00% of MGEs absolute abundance in dust, air and water, respectively. In sediment, transposase genes (51.62%) and integrase genes (48.38%) were the major abundant MGEs (Fig. 5e). Considering the relative abundance of MGEs, from our data, dust and air were the two habitats with the significant highest average relative abundance (8.25 and 3.00 copies per 16S rRNA gene copy respectively) ($P < 0.05$, one-way ANOVA with Tukey B post hoc analysis), which were 1–3 orders of magnitude higher than other habitats (Fig. 5b). Transposase genes dominated in dust, air, and sediment (69.15%, 88.17% and 77.37%, respectively), while integrase genes accounted for a large proportion in soil and water (77.00% and 61.04%) (Fig. 5f).

Usage Notes

The data are shared under Creative Commons Attribution 4.0 International (CC BY 4.0) and the license text can be available at <https://creativecommons.org/licenses/by/4.0/>. The data can be downloaded from Scientific Data Bank⁵⁰, where users can also access when clicking 'download' on online platform (<http://sdc.iue.ac.cn/abundance/arg/list>).

The database provides a valuable resource for exploring the patterns of antibiotic resistance and studying the transfer of ARGs facilitated by MGEs. The spatial distribution analysis of ARGs can be used to identify microbial resistance hotspots. The database also provides basic information for studying the risk of antibiotic resistance.

Code availability

No custom code was made during the collection and validation of this dataset.

Received: 5 October 2023; Accepted: 20 February 2024;

Published online: 27 February 2024

References

1. Leung, E., Weil, D. E., Raviglione, M. & Nakatani, H., World Health Organization World Health Day Antimicrobial Resistance Technical Working Group. The WHO policy package to combat antimicrobial resistance. *Bull. World Health Organ.* **89**, 390–392 (2011).
2. Hernando-Amado, S., Coque, T. M., Baquero, F. & Martínez, J. L. Defining and combating antibiotic resistance from One Health and Global Health perspectives. *Nat. Microbiol.* **4**, 1432–1442 (2019).
3. Murray, C. J. L. *et al.* Global burden of bacterial antimicrobial resistance in 2019: A systematic analysis. *Lancet* **399**, 629–655 (2022).
4. O'Neill, J. *Tackling drug-resistant infections globally: Final report and recommendations* (Review on Antimicrobial Resistance, 2016).
5. Finley, R. L. *et al.* The scourge of antibiotic resistance: The important role of the environment. *Clin. Infect. Dis.* **57**, 704–710 (2013).
6. Larsson, D. G. J. & Flach, C.-F. Antibiotic resistance in the environment. *Nat. Rev. Microbiol.* **20**, 257–269 (2022).
7. Baker, S., Thomson, N., Weill, F.-X. & Holt, K. E. Genomic insights into the emergence and spread of antimicrobial-resistant bacterial pathogens. *Science* **360**, 733–738 (2018).
8. Bengtsson-Palme, J., Kristiansson, E. & Larsson, D. G. J. Environmental factors influencing the development and spread of antibiotic resistance. *FEMS Microbiol. Rev.* **42**, 68–80 (2018).
9. Ezzariai, A. *et al.* Human and veterinary antibiotics during composting of sludge or manure: Global perspectives on persistence, degradation, and resistance genes. *J. Hazard. Mater.* **359**, 465–481 (2018).
10. Singh, R., Singh, A. P., Kumar, S., Giri, B. S. & Kim, K.-H. Antibiotic resistance in major rivers in the world: A systematic review on occurrence, emergence, and management strategies. *J. Cleaner Prod.* **234**, 1484–1505 (2019).
11. Blair, J. M. A., Webber, M. A., Baylay, A. J., Ogbolu, D. O. & Piddock, L. J. V. Molecular mechanisms of antibiotic resistance. *Nat. Rev. Microbiol.* **13**, 42–51 (2015).
12. Peterson, E. & Kaur, P. Antibiotic resistance mechanisms in bacteria: Relationships between resistance determinants of antibiotic producers, environmental bacteria, and clinical pathogens. *Front. Microbiol.* **9**, 2928 (2018).
13. Kunhikannan, S. *et al.* Environmental hotspots for antibiotic resistance genes. *Microbiologyopen* **10**, e1197 (2021).
14. Perron, G. G. *et al.* Functional characterization of bacteria isolated from ancient arctic soil exposes diverse resistance mechanisms to modern antibiotics. *PLoS One* **10**, e0069533 (2015).
15. Harkins, C. P. *et al.* Methicillin-resistant *Staphylococcus aureus* emerged long before the introduction of methicillin into clinical practice. *Genome Biol.* **18**, 130 (2017).
16. Gao, F.-Z. *et al.* Swine farming elevated the proliferation of *acinetobacter* with the prevalence of antibiotic resistance genes in the groundwater. *Environ. Int.* **136**, 105484 (2020).
17. Lee, G. W. *et al.* Changes of antibiotic prescribing pattern and its resistance to *E. Coli* in South Korea: A 12-year retrospective observational study. *Sci. Rep.* **11**, 5658 (2021).
18. Larsen, J. *et al.* Emergence of methicillin resistance predates the clinical use of antibiotics. *Nature* **602**, 135–141 (2022).
19. Pruden, A., Arabi, M. & Storteboom, H. N. Correlation between upstream human activities and riverine antibiotic resistance genes. *Environ. Sci. Technol.* **46**, 11541–11549 (2012).
20. Wang, F. *et al.* Influence of soil characteristics and proximity to antarctic research stations on abundance of antibiotic resistance genes in soils. *Environ. Sci. Technol.* **50**, 12621–12629 (2016).
21. Su, J.-Q. *et al.* Metagenomics of urban sewage identifies an extensively shared antibiotic resistome in China. *Microbiome* **5**, 84 (2017).
22. Zhu, Y.-G. *et al.* Continental-scale pollution of estuaries with antibiotic resistance genes. *Nat. Microbiol.* **2**, 16270 (2017).
23. Li, J. *et al.* Global survey of antibiotic resistance genes in air. *Environ. Sci. Technol.* **52**, 10975–10984 (2018).
24. Asghari, F. B. *et al.* Performance evaluation of ozonation for removal of antibiotic-resistant *Escherichia coli* and *Pseudomonas aeruginosa* and genes from hospital wastewater. *Sci. Rep.* **11**, 24519 (2021).
25. Nguyen, A. Q. *et al.* Monitoring antibiotic resistance genes in wastewater treatment: Current strategies and future challenges. *Sci. Total Environ.* **783**, 146964 (2021).
26. Wang, F. *et al.* Antibiotic resistance in the soil ecosystem: A One Health perspective. *Curr. Opin. Environ. Sci. Health* **20**, 100230 (2021).
27. Lee, G. & Yoo, K. A review of the emergence of antibiotic resistance in bioaerosols and its monitoring methods. *Rev. Environ. Sci. Biotechnol.* **21**, 799–827 (2022).
28. Karkman, A., Do, T. T., Walsh, F. & Virta, M. P. J. Antibiotic-resistance genes in waste water. *Trends Microbiol.* **26**, 220–228 (2018).
29. Zainab, S. M., Junaid, M., Xu, N. & Malik, R. N. Antibiotics and antibiotic resistant genes (ARGs) in groundwater: A global review on dissemination, sources, interactions, environmental and human health risks. *Water Res.* **187**, 116455 (2020).
30. Zhang, Z. *et al.* Assessment of global health risk of antibiotic resistance genes. *Nat. Commun.* **13**, 1553 (2022).
31. Ellabaan, M. M. H., Munck, C., Porse, A., Imamovic, L. & Sommer, M. O. A. Forecasting the dissemination of antibiotic resistance genes across bacterial genomes. *Nat. Commun.* **12**, 2435 (2021).
32. Zhang, Z. *et al.* Advancements in detection and removal of antibiotic resistance genes in sludge digestion: A state-of-art review. *Bioresour. Technol.* **344**, 126197 (2022).
33. Liu, B. & Pop, M. Ardb-antibiotic resistance genes database. *Nucleic Acids Res.* **37**, D443–D447 (2009).
34. Alcock, B. P. *et al.* CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids Res.* **48**, D517–D525 (2020).
35. Yin, X. *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. *Bioinformatics* **34**, 2263–2270 (2018).
36. Zheng, D. *et al.* Global biogeography and projection of soil antibiotic resistance genes. *Sci. Adv.* **8**, eabq8015 (2022).
37. Hendriksen, R. S. *et al.* Global monitoring of antimicrobial resistance based on metagenomics analyses of urban sewage. *Nat. Commun.* **10**, 1124 (2019).
38. Liu, X., Xiao, P., Guo, Y., Liu, L. & Yang, J. The impacts of different high-throughput profiling approaches on the understanding of bacterial antibiotic resistance genes in a freshwater reservoir. *Sci. Total Environ.* **693**, 133585 (2019).
39. Waseem, H. *et al.* Contributions and challenges of high throughput qPCR for determining antimicrobial resistance in the environment: a critical review. *Molecules* **24**, 163 (2019).
40. Willms, I. M. *et al.* Distribution of medically relevant antibiotic resistance genes and mobile genetic elements in soils of temperate forests and grasslands varying in land use. *Genes* **11**, 150 (2020).

41. Delgado-Baquerizo, M. *et al.* The global distribution and environmental drivers of the soil antibiotic resistome. *Microbiome* **10**, 219 (2022).
42. Zheng, J. *et al.* High-throughput profiling and analysis of antibiotic resistance genes in East Tiaoxi River, China. *Environ. Pollut.* **230**, 648–654 (2017).
43. Xiang, Q. *et al.* Seasonal change is a major driver of soil resistomes at a watershed scale. *ISME Commun.* **1**, 17 (2021).
44. An, X.-L. *et al.* Tracking antibiotic resistome during wastewater treatment using high throughput quantitative PCR. *Environ. Int.* **117**, 146–153 (2018).
45. Zheng, J. *et al.* High-throughput profiling of seasonal variations of antibiotic resistance gene transport in a peri-urban river. *Environ. Int.* **114**, 87–94 (2018).
46. Zhou, Z.-C. *et al.* Spread of antibiotic resistance genes and microbiota in airborne particulate matter, dust, and human airways in the urban hospital. *Environ. Int.* **153**, 106501 (2021).
47. Ouyang, W.-Y., Huang, F.-Y., Zhao, Y., Li, H. & Su, J.-Q. Increased levels of antibiotic resistance in urban stream of Jiulongjiang River, China. *Appl. Microbiol. Biotechnol.* **99**, 5697–5707 (2015).
48. Looft, T. *et al.* In-feed antibiotic effects on the swine intestinal microbiome. *Proc. Natl. Acad. Sci. USA* **109**, 1691–1696 (2012).
49. Zhang, A.-N. *et al.* An omics-based framework for assessing the health risk of antimicrobial resistance genes. *Nat. Commun.* **12**, 4765 (2021).
50. Xu, W. *et al.* A database on the abundance of antibiotic resistance genes across China. *Science Data Bank* <https://doi.org/10.57760/sciencedb.09803> (2023).
51. Zhang, Q.-Q., Ying, G.-G., Pan, C.-G., Liu, Y.-S. & Zhao, J.-L. Comprehensive evaluation of antibiotics emission and fate in the river basins of China: Source analysis, multimedia modeling, and linkage to bacterial resistance. *Environ. Sci. Technol.* **49**, 6772–6782 (2015).
52. Huang, F., An, Z., Moran, M. J. & Liu, F. Recognition of typical antibiotic residues in environmental media related to groundwater in China (2009–2019). *J. Hazard. Mater.* **399**, 122813 (2020).
53. Zhang, Q. *et al.* A dataset of distribution of antibiotic occurrence in solid environmental matrices in China. *Sci. Data* **9**, 276 (2022).
54. Zhou, Z.-C. *et al.* Antibiotic resistance genes in an urban river as impacted by bacterial community and physicochemical parameters. *Environ. Sci. Pollut. Res.* **24**, 23753–23762 (2017).
55. Zhou, S.-Y.-D. *et al.* Conurbation size drives antibiotic resistance along the river. *Sci. Total Environ.* **823**, 153822 (2022).

Acknowledgements

We would like to acknowledge the help of Hong Chen for providing parts of the data. This work was supported by The International Partnership Program of Chinese Academy of Sciences (132C35KYSB20200012), the Informatization Plan of Chinese Academy of Sciences (CAS-WX2021SF-0302), National Natural Science Foundation of China (32161143016, 42177362, 42250410328), National Basic Science Data Center “Environment Health DataBase” (NO. NBSDC-DB-21), and Belarusian Republican Foundation for Fundamental Research and Ministry of Education of the Republic of Belarus (GPNI GR no. 20212317).

Author contributions

Data collection and writing, Wenjuan Xu and Zhizhen Pan; Data collection, Xin-Li An; Data integration and web design, Yangyu Wu and Weiyi Wang; Review and revision, Boris Adamovich; Writing, review and editing, Yong-Guan Zhu, Jian-Qiang Su, and Qiansheng Huang.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41597-024-03084-8>.

Correspondence and requests for materials should be addressed to J.-Q.S. or Q.H.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024